CODEN: JASMAN

ISSN: 0001-4966

The Journal of the Acoustical Society of America

Vol. 112, No. 5, Pt. 1 November 2002

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 112, NO. 5, PT. 1, NOVEMBER 2002

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 112, NO. 5, PT. 1, NOVEMBER 2002

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

Amy Donohue—For contributions to program development and public understanding of hearing research.

Leon M. Keer—For contributions to structural dynamics of continuous and distributed bodies.

Sean F. Wu—For contributions to structural acoustics and noise control.

ASA members receive Navy Awards

D. Vance Holiday, Principal Scientist and Director of Research and Technology at BAE Systems in San Diego, CA, was awarded the Department of the Navy Meritorious Public Service Award on 3 June 2002 at the Pittsburgh meeting of the Acoustical Society of America. The award was presented by Jeffrey Simmen, ONR Ocean Acoustics Program Manager, on behalf of The Chief of Naval Research (see Fig. 1).

FIG. 1. Jeffrey Simmen (r) presents the Department of the Navy Meritorious Public Service Award to D. Vance Holliday (1).

The award citation read: ''For sustained meritorious public service to the Office of Naval Research and the Navy for the past 40 years. Dr. Holliday is the dean of investigators in the area of marine bioacoustics in the United States. Applying his extensive background in pure and applied acoustics, he has pioneered the application of new technology instrumentation to study and measure life at trophic levels in the marine ecosystem ranging from marine mammals and fish to marine zooplankton. His research work in support of the Navy has spanned the range of environments from temperate seas to the Arctic Ocean. Dr. Holliday played a key role in the development and operation of the Sensor Accuracy Check Site at Long Beach Naval Station, providing the fleet with a new capability to calibrate and test sonar and radar systems prior to deployment. He led the search for a new location for conducting noise trials on submarines which resulted in the discovery, survey and measurement of the acoustic properties of the Behm Canal. Dr. Holliday developed and implemented the Tracor Acoustic Profiling System (TAPS) which is today the most widely used multi-frequency bioacoustic sampling system in ocean science. He has led his team on countless Office of Naval Research funded scientific experiments ranging from mid-water reverberation studies in the Indian Ocean to cutting-edge research on Thin Layers and Benthic Acoustics. Dr. Holliday's exceptional scientific contributions have had a significant impact on the Office of Naval Research and Navy and are deeply appreciated.''

Van Holliday is a Fellow of the Acoustical Society of America and a founding member of the Technical Committee on Acoustical Oceanography.

Frederick D. Tappert was awarded the Department of the Navy Superior Public Service Award. The award was presented to Dr. Tappert on 4 October 2001 at the Rosensteil School of Marine and Atmospheric Science in Miami, FL. The award citation read: ''To Frederick D. Tappert, Professor of Applied Marine Physics at the University of Miami, for his seminal contributions to the development of parabolic-equation-based underwater-acoustic propagation models. In 1974 Professor Tappert's introduction of the

 0001-4966/2002/112(5)/1713/7/$19.00

parabolic equation (PE) propagation model started a revolution in the underwater acoustics modeling community. Since that time Prof. Tappert has remained at the forefront of further developing and improving the model. He has made fundamental contributions to both theoretical and numerical developments, and has used the model in a wide variety of Naval applications. It is, in large part, a tribute to Prof. Tappert's efforts that today the PE model is the de facto standard full wave propagation model in underwater acoustics.''

Fred Tappert died of pancreatic cancer on 9 January 2002. He was a Fellow of the Acoustical Society of America and will be posthumously awarded the ASA Pioneers of Underwater Acoustics Medal at the Cancun meeting in December 2002.

ASA Awards Presented at International Science and Engineering Fair (ISEF)

ASA was again privileged to present awards to four high school students during the annual ISEF, held this year in Louisville, Kentucky under the corporate sponsorship of the Intel corporation. The fair has been held since 1950 to simulate interest in scientific and engineering careers among high school students. This year's fair featured 1238 entrants from the United States and 40 other nations exhibiting projects covering 14 technical disciplines. Grand Awards judges awarded prizes from $5000 to $500 to first-through-fourth place finishers in each of the 14 disciplines, and Special Awards judges made awards on behalf of each of over 70 professional associations, including ASA. Our Society presents a first place award of $500 and three noncash Honorable Mention certificates. Each awardee also receives a free one-year subscription to JASA on CD ROM.

Our first prize award went to 17-year-old Naveen Sinha from Los Alamos High School in Los Alamos, New Mexico for his physics project entitled ''Bubble-Based Resonance-Doppler Sensor for Liquid Characterization,'' an excellently presented exhibit on a process for identifying unknown liquids by the response of introduced bubbles to the viscosity, surface tension, and other physical properties of the liquid. Our judges did not know until after the awards that Naveen had also won ASA's first place award last year for a project on the interaction of sound and liquid crystals, and was an ASA Honorable Mention the year before that. What a wonderful distinction for this young man to have won an ASA award three-years running, and the top prize in the last two years! Some acoustically-related college or university needs to give this student a full scholarship!

ASA also awards three Honorable Mention prizes annually, and these were also of such high quality that we had difficulty assigning them to less than first place. This year, one award went to 18-year-old Yin Nee from Horace Greeley High School in Chappaqua, New York for her Computer Science project on ''Using Hidden Dynamic Models to Predict Pitch in English and Mandarin,'' a technically sophisticated system for computerized voice recognition capable of translating identical words in Mandarin Chinese which change their semantic meanings according to pitch changes in the way they are pronounced. A second award went to 17-year-old Bryce Kittinger Campbell from the Academy of Science and Technology in Conroe, Texas for his Engineering project on ''Vortex-Induced Vibrations on Flexible Cylinders,'' which used both hydrodynamic modeling theory and empirical water basin tests to show that long, flexible cylinders have different vortex-induced vibration characteristics than those predicted based on solid cylinder characteristics, a finding applicable to certain offshore oil drilling structures. A third award went to 15-year-old Giselle M. Torres from the Academia del Perpetuo Socorro in San Juan, Puerto Rico for her Medicine and Health project on the ''Effect of Vibration and Calcium on Bone Density,'' an investigation of the combined use of calcium and mechanical vibrations to reduce bone demineralization in animal bones (measured by bone densitometry), with applications to reducing osteoporosis in elderly humans (see Fig. 1).

FIG. 1. Alvin B. Broderson, Chair of ASA Judges, with winners of ASA awards at 2002 Intel International Science and Engineering Fair. (l to r) Bryce Kittinger Campbell, Giselle M. Torres, Alvin B. Broderson, Yin Nee and Naveen Sinha.

The judging team included Dr. Christopher Richards from the University of Louisville, Department of Mechanical Engineering; Mr. Tim A. Hamel from the General Electric Company in Louisville, and Dr. A. B. Broderson of VESCO Engineering in Louisville, who also presented the awards. The team was pleased that the awards went to a nice mix of young men and women, Americans and foreign nationals, and that they included representatives of the physics, engineering, computer science, and medical disciplines of ASA interest. We were uniformly impressed by the growing technical education and skills of today's high-school students, and feel (based on the entire assembly of exhibits) that our scientific and engineering future is falling into increasingly capable hands.

We found the fair to be technically exhilarating and would recommend that any ASA member asked to judge next year's fair accept the invitation with enthusiasm. We also recommend that ASA consider offering three cash prizes rather than one, with larger cash values, beginning with next year's fair. Many of the students who present at this fair are from poorer school districts and they clearly deserve a more substantial measure of our financial support and recognition.

A. B. BRODERSON, Ph.D., P.E.

Advanced Degree Dissertation in Acoustics

Editor's Note: Abstracts of doctoral and Masters theses will be welcomed at all times. Please note that they must be limited to 200 words, must include the appropriate PACS classification numbers, and formatted as shown below. If sent by postal mail, note that they must be double spaced. The address for obtaining a copy of the thesis is helpful. Submit abstracts to: Acoustical Society of America, Thesis Abstracts, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502, e-mail: asa@aip.org

Dual-pulses for cavitation control in lithotripsy: Shock wave-bubble interactions and bioeffects [43.80.Gx]—Dahlia L. Sokolov, *Department of Bioengineering, University of Washington, July 2002 (Ph.D.).* Cavitation, the growth and collapse of gas/vapor bubbles, appears to play an important role in both stone comminution and tissue injury during shock wave lithotripsy, the clinical treatment in which focused, high amplitude shock pulses are used to comminute kidney stones. The goal of this research was to characterize *in vitro* cavitation activity and stone and cell damage in a novel system that uses converging dual pulses, produced by two identical, confocal lithotripters, to modify the cavitation field. The cavitation bubble dynamics were numerically calculated, and experiments were performed in a research electrohydraulic shock wave lithotripter to determine bubble size, lifetime, and pit depth created in aluminum foils by cavitation collapse. Furthermore, damage to model stones and to red blood cells was measured for both single and dual-pulses. A single shock pulse creates a $\sim 15 \times 100$ mm cloud of bubbles in water. The greatest cavitation activity and stone damage from single-pulses was found to occur 2 cm proximal to the geometric focus, $F2$, where the stone is normally aligned. Therefore, a 2 cm shift in stone alignment may potentially improve stone comminution and reduce tissue injury in clinical treatment. The dual-pulse lithotripter, on the other hand, generates a localized and intensified cavitation field that increased stone comminution efficiency at $F2$ by at least three times the maxi-

mum values achieved by single-pulses. At *F*2, acoustic pressure approximately doubled, as did bubble size, collapse time, and pit depth on foils. A significant reduction in comminution of stones suspended in glycerol indicates that cavitation activity, not the doubling of acoustic pressure, explains the increased comminution. On either side of *F*2, the second delayed pulse mitigated bubble collapse, resulting in little or no pitting on foils and reduced hemolysis, even when compared with single pulses. Numerical calculations of radial dynamics agreed with experimental findings. No significant bubble translation was observed or calculated. The data support the conclusion that cavitation plays an important role in efficient stone comminution and hemolysis. The dual-pulse lithotripter, because of its localized and intensified cavitation field, yields the promise of improved stone comminution efficiency, reduced tissue damage, and accelerated treatment.

Thesis advisor: Lawrence A. Crum

Copies of this thesis written in English can be obtained from the Applied Physics Laboratory Library, University of Washington, 1013 NE 40th Street, Seattle, WA 98105. E-mail address: library@apl.washington.edu.

Report of the Auditor

Published herewith is a condensed version of our auditor's report for calendar year ended 31 December 2001.

Independent Auditors' Report

To the Executive Council
Acoustical Society of America

We have audited the accompanying statements of financial position of the Acoustical Society of America as of December 31, 2001 and the related statements of activity and cash flows for the year then ended. The prior year summarized comparative information has been derived from the organization's 2000 financial statements and, in our report dated May 16, 2001, we expressed an unqualified opinion on those financial statements. These financial statements are the responsibility of the Society's management. Our responsibility is to express an opinion on the financial statements based on our audit.

We conducted our audit in accordance with auditing standards generally accepted in the United States of America. Those standards require that we plan and perform the audit to obtain reasonable assurance about whether the financial statements are free of material misstatement. An audit includes examining, on a test basis, evidence supporting the amounts and disclosures in the financial statements. An audit also includes assessing the accounting principles used and significant estimates made by management, as well as evaluating the overall financial statement presentation. We believe that our audit provides a reasonable basis for our opinion.

In our opinion, the financial statements referred to above present fairly, in all material respects, the financial position of the Acoustical Society of America as of December 31, 2001 and the changes in its net assets and its cash flows for the year then ended in conformity with accounting principles generally accepted in the United States of America.

CONROY, SMITH & CO.
27 June 2002
New York, NY

ACOUSTICAL SOCIETY OF AMERICA
STATEMENT OF FINANCIAL POSITION
AS OF 31 DECEMBER 2001
(With Summarized Totals For 2000)

	2001	2000
Assets:		
Cash and cash equivalents	$1,326,755	$1,301,371
Accounts receivable	136,551	324,410
Marketable securities	5,917,761	6,522,168
Furniture, fixtures and equipment—net	107,286	107,191
Other assets	306,765	349,219
Total assets	$7,795,118	$8,604,359
Liabilities:		
Accounts payable and accrued expenses	$ 429,131	$ 453,544
Deferred revenue	1,264,467	1,449,583
Deferred rent liability	-	41,989
Total liabilities	$1,693,598	$1,945,116
Net assets:		
Unrestricted	$5,055,355	$5,518,982
Temporarily restricted	543,275	565,655
Permanently restricted	502,890	574,606
Total net assets	$6,101,520	$6,659,243
Total liabilities and net assets	$7,795,118	$8,604,359

ACOUSTICAL SOCIETY OF AMERICA
STATEMENT OF ACTIVITY
FOR THE YEAR ENDED 31 DECEMBER 2001
(With Summarized Totals For 2000)

	Unrestricted net assets	Temporarily restricted net assets	Permanently restricted net assets	Total	2000 Total
Support and revenue:					
Contributions	-	$ 25,000	$ 361	$ 25,361	$ 623
Dues	$ 637,075	-	-	637,075	658,172
Publishing	2,134,353	-	-	2,134,353	2,172,282
Standards	259,614	-	-	259,614	418,779
Meetings	410,137	-	-	410,137	396,038
Interest and dividends	232,551	23,823	24,214	280,588	340,555
Unrealized gain (loss)	(793,683)	(81,302)	(82,642)	(957,627)	(613,782)
Other	56,054	-	-	56,054	48,235
Realized gain (loss)	176,244	18,054	18,351	212,649	21,530
Net assets released from restrictions	7,955	(7,955)	-	-	-
Total support and revenue	$3,120,300	($ 22,380)	($ 39,716)	$3,058,204	$3,442,432
Expenses:					
Publishing	$1,524,644	-	-	$1,524,644	$1,630,098
Standards	514,061	-	-	514,061	536,893
Administrative and general	670,364	-	-	670,364	559,803
Meetings	542,939	-	-	542,939	449,117
Other	331,919	-	$ 32,000	363,919	435,623
Total expenses	$3,583,927	-	$ 32,000	$3,615,927	$3,611,534
(Decrease) in net assets	($ 463,627)	($ 22,380)	($ 71,716)	($ 557,723)	($ 169,102)
Net assets, beginning of year	5,518,982	565,655	574,606	6,659,243	6,828,345
Net assets, end of year	$5,055,355	$543,275	$502,890	$6,101,520	$6,659,243

ACOUSTICAL SOCIETY OF AMERICA
STATEMENT OF CASH FLOWS
FOR THE YEAR ENDED 31 DECEMBER 2001
(With Summarized Totals For 2000)

	Total All Funds 2001	2000
Operating Activities		
(Decrease) in net assets	($ 557,723)	($ 169,102)
Adjustments to reconcile net income to net cash provided by operating activities:		
Depreciation and amortization	30,897	32,218
Net realized (gain) on investments	(212,649)	(21,530)
Unrealized loss on marketable securities	957,627	613,782
Changes in operating assets and liabilities:		
Decrease in accounts receivable	187,859	48,426
Decrease (increase) in other assets	42,454	(56,285)
(Decrease) increase in accounts payable and accrued expenses	(24,413)	178,602
(Decrease) in deferred rent liability	(41,989)	(2,203)
(Decrease) in deferred revenue	(185,116)	(75,172)
Loss on disposal of equipment and leasehold improvements	40,257	-
Net cash flows provided by operating activities	$ 237,204	$ 548,736
Investing Activities		
Purchase of furniture, fixtures, equipment and leasehold improvements	($ 73,148)	($ 21,331)
Proceeds from sale of securities	4,270,712	1,815,041
Purchase of securities	(4,409,384)	(2,084,119)
Net cash (used in) investing activities	($ 211,820)	($ 290,409)
Increase in cash and cash equivalents	$ 25,384	$ 258,327
Cash and cash equivalents, beginning of year	1,301,371	1,043,044
Cash and cash equivalents, end of year	$1,326,755	$1,301,371

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.

2002

2–6 Dec. First Pan-American/Iberian Meeting on Acoustics (Joint Meeting: 144th Meeting of the Acoustical Society of America, 3rd Iberoamerican Congress of Acoustics, and 9th Mexican Congress on Acoustics) Cancun, Mexico [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: asa.aip.org/cancun/cancun.html].

2003

13–15 March American Auditory Society Annual meeting, Scottsdale, AZ [American Auditory Society, 352 Sundial Ridge Cir., Dammeron Valley, UT 84783; Tel.: 435-574-0062; Fax: 435-574-0063; E-mail: amaudsoc@aol.com; WWW: www.amauditorysoc.org].

28 April–2 May 145th Meeting of the Acoustical Society of America, Nashville, TN [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: asa.aip.org].

5–8 May SAE Noise & Vibration Conference & Exhibition, Traverse City, MI [P. Kreh, SAE International, 755 W. Big Beaver Rd., Suite 1600, Troy, MI 48084; Fax: 724-776-1830; WWW: http://www.sae.org].

23–25 June NOISE-CON 2003, Cleveland, OH [INCE Business Office, Iowa State Univ., 212 Marston Hall, Ames, IA 50011-2153; Fax: 515-294-3528; E-mail: ibo@ince.org].

5–8 Oct. IEEE International Ultrasonics Symposium, Honolulu, HI [W. D. O'Brien, Jr., Bioacoustics Research Lab., Univ. of Illinois, Urbana, IL 61801-2991; Fax: 217-244-0105; WWW: www.ieee-uffc.org].

10–14 Nov. 146th Meeting of the Acoustical Society of America, Austin, TX [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: asa.aip.org].

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: asa.aip.org].

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].

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; E-mail: asa@aip.org; WWW: asa.aip.org].

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

Ordering information: Orders must be paid by check or money order in U.S. funds drawn on a U.S. bank or by Mastercard, Visa, or American Express credit cards. Send orders to Circulation and Fulfillment Division, American Institute of Physics, Suite 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).

Revision List

New Associates

Athalye, K., ARK Noise Control Engineering Ctr., 73/3/3 ''Namoste-stu,'' Bhakti Marg, Erandwana, Pune, Maharashtra State 411004, India

Azevedo, Maria Augusta D., Acoustical Lab., USG Corporation, 700 North Highway 45, Libertyville, IL 60048

Bedenbaugh, Purvis H., Neuroscience, McKnight Brain Inst., Univ. of Florida, P.O. Box 100244, Gainesville, FL 32610

Biesel, Van B., Mechanical Engineering, Georgia Inst. of Technology, 771 Ferst Dr., Atlanta, GA 30332-0405

Bromilow, Ian D., Bldg. Design Partnership, 16 Gresse St., London W1A 4WD, United Kingdom

Buschert, John R., Physics Dept., Goshen College, 1700 South Main St., Goshen, IN 46526

Davis, Penny R., 8640 SW 212th St., Apt. 307, Miami, FL 33189

Day, Joseph L., U.S. Government, Navy, 4251 Suitland Rd., Washington, DC 20395-5720

Fatemi, Mostafa, Basic Ultrasound Research Lab., Dept. of Physiology and Biophysics, Mayo Clinic, 200 First St. SW, Rochester, MN 55905

Fox, Colin, Mathematics, Univ. of Auckland, PB 92019, Auckland, New Zealand

Frenz, Eric R., Noise and Vibration Division, MTS Systems Corporation, 800 East Whitcomb Ave., Madison Heights, MI 48071

Funatsu, Seiya, Hiroshima Women's Univ., 1-1-71 Ujinahigashi Minami-ku, Hiroshima 734-8558, Japan

Gross, Thomas P., GETRAG Corp., Noise and Vibration, 36200 Mound Rd., Sterling Heights, MI 48310

Hamilton, Michael J., 3680 Alabama St., #7, San Diego, CA 92104

Hamrin, John E., TEAM Industries-Bagley, 105 Park Ave. NW, P.O. Box 539, Bagley, MN 56621

Hollich, George J., Psychological Science, Purdue Univ., West Lafayette, IN 47907

Honda, Ethan P., Applied Research Labs, Univ. of Texas, 10000 Burnet Rd., Austin, TX 78758

Huber, Thomas M., Gustavus Adolphus College, 800 College Ave., Saint Peter, MN 56082
Hughes, William O., Structural Systems Dynamics Branch, M/S 86-10, NASA Glenn Research Ctr., 21000 Brookpark Rd., Cleveland, OH 44135
Johnson, Brian S., BDP Acoustics, P.O. Box 4WD, 16 Gresse St., London W1A 4WD, United Kingdom
Johnson, Thomas E., Grason Stadler, 5225-2 Verona Rd., Madison, WI 53711
Jorgensen, Daren E., Jorgensen Environmental, 10374 Trademark St., Ranco Cucamonga, CA 91730
Kefauver, Alan P., Recording Arts and Sciences, Peabody Inst., The Johns Hopkins Univ., 1 East Mt. Vernon Place, Baltimore, MD 21202
Kim, Jin Chol, H. S. Cooperation, CPO Box 4973, Seoul 100-649, South Korea
Konrad-Martin, Dawn L., Communication Disorders and Sciences, Rush Univ., 1653 West Congree Parkway, Room 1015, Armour, Chicago, IL 60612
Kraft, Barbara J., Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, 24 Colovos Rd., Durham, NH 03820
Krull, Diana V., Linguistics Dept., Stockholm Univ., Univeritetsvagen 10C, Stockholm 10691, Sweden
Li, Haiying, College of Marine Engineering, Noise and Vibration Control Lab., You Yi Western Rd. 127, P.O. Box 58, Xian, Shannxi 710072, People's Republic of China
Liu, Lanbo, Geology and Geophysics, Univ. of Connecticut, 354 Mansfield Rd., U-2045, Storrs, CT 06269-2045
Lu, Zhiqu, National Ctr. for Physical Acoustics, Univ. of Mississippi, Coliseum Dr., University, MS 38677-1848
MacIntosh, Scott, SELF, 43 Brucewood St., Boston, MA 02132
Makoto, Hiroshige, Faculty of Engineering, Hokkaido Univ., North 13, West 8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan
McElveen, James K., Clarifying Technologies, Inc., P.O. Box 6157, Raleigh, NC 27608-6157
Najolia, Dorie A., Acentech, Inc., 33 Moulton St., Cambridge, MA 02138
Nakanishi, Shinsuke, Dept. of Integrated Architecture, Faculty of Infrastructural Technologies, Hiroshima International Univ., Hirokoshingai 5-1-1, Kure, Hiroshima 737-0112, Japan
Nancy, Winfree A., Dominca, 9813 Admiral Dewey Ave. NE, Albuquerque, NM 87111
Ohm, Jens Rainer, Inst. fuer Nachrichteu Technik, Melateuer Strasse 23, Aachen 52072, Germany
Philhong, Lee, Bluetek Co., Ltd., 416 Meatan3-dong, Paldal-gu, Suwon-si, Kyunggi-do 442-742, Korea
Rehder, Douglas E., Rehder Hearing Clinic, 1101 North 27th St., Suite E, Billings, MT 59101
Reilly, Thomas G., Echo Span, 3165 South Pearl St., Englewood, CO 80110-2709
Runge Samuelson, Christina L., Otolaryngology and Communication Sciences, Medical College of Wisconsin, 9200 West Wisconsin Ave., Milwaukee, WI 53226
Schmidt, Val E., Lamont Doherty Earth Observatory, Instrument Lab., Columbia Univ., 61 Route 9W, Palisades, NY 10964
Smith, Martin L., New England Research, Inc., 331 Olcott Dr., Suite L1, White River Junction, VT 05001
Soremekun, Rotimi J., Biofem Nigeria, Ltd., 3A Olorunnimbe St., Wemabod Estate, Ikeja, Lagos, Nigeria
Todd, Neil N. P., Psychology Dept., Univ. of Manchester, Oxford Rd., Manchester M13 9PL, United Kingdom
Tsoflias, Sarah L., U.S. Dept. of the Interior, Minerals Management Service, 1201 Elmwood Park Boulevard, MS 5433, New Orleans, LA 70123-2394
Tyson, Thomas N., 1312 West Whiteside St., Springfield, MO 65807
Valdivia, Nicolas P., Physical Acoustics, Naval Research Lab., Code 7130, 4555 Overlook Ave. SW, Washington, DC 20375
Vera, Michael D., Inst. of Geophysics and Planetary Physics, Scripps Inst. of Oceanography, #0225, La Jolla, CA 92093
Villada, Fernando, Fiberglass Colombia S.A., Cra 9 No. 80-12, Bogota D.C. Cundinamarca, Colombia
Wang, Huigang, Marine Engineering College, Signal & Information Processing Lab., You Yi Western Rd. 127, P.O. Box 58, Xian, Shannxi 710072, People's Republic of China
Yang, Byunggon, English Dept., Dongeui Univ., 24 Kayadong, Pusanjin Gu, Pusan 614-714, South Korea
Young, Frederick J., 800 Minard Run Rd., Bradford, PA 16701
Zawadzki, Ireneusz, NSWC, Carderock Division, 9500 MacArthur Boulevard, Bethesda, MD 20817

New Students

Agwuele, Augustine, 2501 Lake Austin Boulevard, K108, Austin, TX 78703
Ayala, Christopher D., Bowling Green State Univ., Psychology Dept., Bowling Green, OH 43403
Baikachi, Yoko, 3-20-3 Kirigaoka, Midori-Ku, Yokohama, Kanagawa 226-0016, Japan
Bell, Daniel A., 204 East 9th Ave., Johnson City, TN 37604
Blackwell, Ezra L., 4863 West St. Paul Ave., Chicago, IL 60639
Block, Gareth, Applied Research Labs., Univ. of Texas, 10000 Burnet Rd., Austin, TX 78758
Boes, Joachim, Nider-Ramstaedter Str. 51, Darmstadt, Hessen D-64283, Germany
Bravo Pinto, Ernesto M., Chanks 239, Nunoa, Santiago, Chile
Brill, Stefan M., Inst. for Applied Physics, Univ. of Innsbruck, Technikerstr. 25, Innsbruck A-6020, Austria
Brouqueyre, Laurent, 5401 Spring Creek Ln., Dunwoody, GA 30350
Budd, Sara R., 0615 SW Palatine Hill Rd., Burgin, Omid C., OMID Studios, Rua Girassol 965, Sao Paulo 05433-002, Brazil
Caner, Edward M., Case Western Reserve Univ., 10900 Euclid Ave., Cleveland, OH 44106
Chipley, Michael R., 1145 Trentini Ave., Wake Forest, NC 27587
Cretton, Christine, Ruelle du Fornel 2, Neuchatel 2000, Switzerland
Crooks, Gary E., G. C. Productions, 4040 South Bell St., Tacoma, WA 98418
Deshmukh, Om D., 5616 Hamilton Manor Dr., #6, Hyattsville, MD 20782
Di Angelo, Adam, 1147 South 11th St., Phildelphia, PA 19147
Dow, Tracy C., 11615 132nd St., Apt. 12, Edmonton, Alberta T5M 1G5, Canada
Du, Yafei, Speech and Hearing Science, Ohio State Univ., 1070 Carmack Rd., Columbus, OH 43221
Eberle, Geoffrey, 3/107 Hudsons Rd., Spotswood, VIC 3015, Australia
Farny, Caleb H., 44 Quint Ave., Apt. 11, Allston, MA 02134
Griffin, Sarah, 8 Laburnham House, Lordship Terrace, Stoke Newington, London N16 0JB, United Kingdom
Guan, Jingfeng, Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105
Howarth, Eric, 1055 Aloha Dr., Encinitas, CA 92024
Kayan, Thomas E., 212 Old Springfield Rd., Stafford Springs, CT 06076
Kim, Ji Eun, 2215 Cornerstone Ln., #2054, Arlington, TX 76013
Kim, Sahyang, Linguistics Dept., Univ. of California at Los Angeles, 3125 Campbell Hall, Box 951543, Los Angeles, CA 90066
Knezek, Kathleen M., 1712 Coturnix Dr., Austin, TX 78758
Kong, Ying-Yee, 6409 Adobe Cir., Irvine, CA 92612
Kwak, Sangyeop, 11257 Key West Ave., Unit 4, Northridge, CA 91326
Lancaster, Aaran D., 2193 East Long Lake Rd., Troy, MI 48085-3575
Lerner, Elizabeth A., Psychology, Wright State Univ., 3640 Colonel Glenn Highway, Dayton, OH 45435
Loubeau, Alexandra, 131 Sowers St., #F-4, State College, PA 16801
Maehr, Michael D., 316 West John St., Champaign, IL 61820
McGinnes, Courtney, 3515 SW 39th Boulevard, Apt. 12D, Gainesville, FL 32608
Miller, Nathaniel S., Psychology, Bowling Green State Univ., Bowling Green, OH 43403
Moondra, Manmohan S., Mechanical Engineering, Wayne State Univ., 5050 Anthony Wayne Dr., Detroit, MI 48202
Okamura, Kyoko, #308 Campus View Apts., Bloomington, IN 47408
Philipp, Norman H., 721 West Elm, Olathe, KS 66061
Posdamer, Stephanie H., Physics Dept., Washington Univ., One Brookings Dr., Campus Box 1105, St. Louis, MO 63130
Purwanto, Agus, Dept. of Physics and Astronomy, Brigham Young Univ., N377 ESC, Provo, UT 84602
Richlen, Michael F., 2469 168th Place NE, Bellevue, WA 98008
Ross, David A., 81 Main St., #45, Branford, CT 06405
Scelo, Thomas, Acoustics Research Centre, Univ. of Auckland, Private Bag 92019, Auckland 1020, New Zealand
Silva, Glauber T., Ultrasound Research Lab., Mayo Foundation, 200 1st St. SW, Rochester, MN 55905
Sippo, Kurra, Taitoniekantie 9 h 14, Jyvaskyla, Keski-Suomi, Finland

Whitehead, Jeff J., New Hampshire Technical Inst., 11 Institute Dr., Concord, NH 03224
Yasuo, Nakajima, 372-16 Yokokawa-cho, Hachioji-shi, Tokyo 193-0823, Japan
Yunusova, Yana S., 1005 Grant St., #4, Madison, WI 53711
Zanetti, Paolo, 134 St. Mary's St., Apt. 5, Boston, MA 02215

New Electronic Associates

Chen, Shigao, Mayo Clinic, 200 First St. SW, Rochester, MN 55905
Deacon, Crystal K., P.O. Box 36, West Decatur, PA 16878
DiLorenzo, Anthony, U.S. Dept. of Commerce, U.S. Patent and Trademark Office, 2100 Crystal Dr., Arlington, VA 22202
Eldredge, Jeff D., Dept. of Engineering, Cambridge Univ., Trumpington St., Cambridge CB2 1PZ, United Kingdom
Furukawa, Taro, 21-15 Koriencho, Hirakata, Osaka-fu 573-0086, Japan
Kordik, Alex, 54 Rockwood Ave., Dayton, OH 45405
Lemay, Gerald A., 43 Metcalf Dr., Belleville, IL 62223
Reynolds, Scott M., Calsonic Kansei, One Calsonic Way, ESDC B-6, Shelbyville, TN 37160
Toda, Minoru, Measurement Specialties, Inc., 950 Forge Ave., Norristown, PA 19403
Walden, Alan K., Siemens Westinghouse Power Corp., 4400 Alafaya Trail, MC DV 322, Orlando, FL 32826-2399
Wassink, Alicia B., Dept. of Linguistics, Univ. of Washington, Box 354340, Seattle, WA 98195-4340

New Corresponding Electronic Associates

Hernandez-Diaz Huici, Maria E., Centro Estudios de Electronica y Tecnologias de la Informacion, Universidad Central Las Villas, C. Camajuani Km 5 1/2, Santa Clara, Villa Clara 54830, Cuba
Joshi, Nandini V., Audiology, Medinova Diagnostic Centre, J. M. Rd., Opp. Modern School, Maharashtra 411005, India
Konopka, Andrzej, Oticon Polska Sp. z o.o., Pl. Trzech Krzyzy 4/6, Warsaw MZ 00-499, Poland
Malherbe, Francois, Carinus St. 207, Meyerspark, Gauteng 0184, South Africa
Plazaola, Carlos R., Aptdo. 6-4787 Dorado, 19717, Panama
Stickar, Pablo G., Peron 635 Bloque 1 Unidad 28, Campana, Buenos Aires 2804, Argentina

Members Elected Fellows

R. H. Campbell, L. H. Carney, B. D. Cornuelle, A. M. Donahue, G. E. Ioup, L. M. Keer, H. J. McDermott, C. M. McKay, W. C. Moss, P. A. Nelson, M. H. Orr, J. E. Randorff, S. F, Wu, N. Xiang

Associates Elected Members

S. A. Adams, D. J. Bierenbaum, I. C. Bruce, R. S. Bryant, Y. A. Chen, B. Delgutte, B. W. Drinkwater, R. A. Dyba, G. S. Edelson, A. J. Fenwick, M. V. Golden, P. Hanes, W. Hui, A. I. Iliev, M. Krajnc, J. T. Kunio, I. Maddieson, B. Marn, S. W. Martin, W. A. Pritchard, G. Ravn, M. Ruzzene, M. L. Saunders, P. L. Schmidt, D. Wartzok, K. R. Waters, D. M. Wellsted, N. J. Williams

Associate to Corresponding Electronic Associate

C. Padmanabhan

Reinstated

A. Sharma, R. V. Murray—*Members*
T. A. Pyper, D. Rollow, T. M. Young—*Associates*
R. M. Ead, M. S. Hnilo, L. S. Ryden, R. E. Stone, Jr., S. M. Unger—*Electronic Associates*
W. H. Cho—*Corresponding Electronic Associate*

Resigned

R. Guedes—*Associate*

Deceased

O. L. Angevine, R. E. Apfel, A. Glorig, W. D. Neff—*Fellows*
J. E. Ancell, A. L. Anderson, F. Daniel, S. S. Kushner, D. W. Nicoletti—*Members*

Fellows	900
Members	2614
Associates	2776
Students	961
	7251

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).]*

The Noise Manual, Fifth Edition

Elliott H. Berger, Larry H. Royster, Dennis P. Driscoll, Julia Doswell Royster, and Martha Layne, Editors

American Industrial Hygiene Association, Fairfax, Virginia, 2000. Xxii+796 pp. Price: $74.00 (hardcover) ISBN: 1931504024.

The last *Noise Manual* (Fourth Edition) was published in 1986. Certainly a new edition of this important reference book was long overdue. The latest edition includes new material in ten and complete revisions in two of its 18 chapters. The new areas that have been added are room noise criteria, speech and signal detection, relating hearing loss to noise exposure, community noise, and more detailed discussion of regulations and standards. The remaining six chapters have been revised and updated. The professional with interests in noise and its effects might well consider upgrading his/her professional library to the fifth edition. For those without any edition of the noise manual, the fifth edition will serve as a well-referenced and indexed handbook. The authors stated the manual is aimed to appeal to novice industrial hygienists while serving as a reference source for experienced practitioners. These are somewhat conflicting goals. While in some of the chapters the authors have accomplished these objectives reasonably well, I believe that many novices will find some of the chapters overly detailed and difficult to follow. The novice needs to understand they need not read everything in the book.

There are some small errors or discrepancies in the manual, most of which need not be mentioned. For example, on p. 137, Figs. 5.3 and 5.4 are said to represent the total qualitative knowledge about noise exposure in humans. Then Figs. 5.5 and 5.6 show an additional model of hearing loss. There are also many places where the particular author shows a distinct bias without showing that there is not unanimity on the issue. This is apparent with a couple of authors. A good example can be found on p. 141, guideline 4, where it is stated, ''the 3-dB exchange rate will result in the inclusion of a significantly greater number of the noise-exposed population with very low TWAs that would not have to be included under OSHA criteria'' (TWA is time weighted average and OSHA is the Occupational Health and Safety Administration). The authors go on to say the inclusion of the populations is not justifiable. Is this the advice to tell the novice industrial hygienist? No mention is given with respect to the other side of the issue, which is simply that the population that is not protected by the 5-dB rule is also in this low TWA population. For example, 109 dB for 15 min is an OSHA TWA of 84 dB, an exposure I certainly would think is worthy of a hearing conservation program.

The one chapter that was deleted in the fifth edition from the fourth edition is the one on visual evaluation of the ear and the eardrum. I would have liked to see this fact mentioned in the preface, and in a reference as well, where such information could best be obtained. Before discussing each of the 18 chapters in detail, I would like to point out that each chapter was peer reviewed by both the editors and from a pool of 21 other professionals.

Section I: Fundamentals of Sound, Vibration, and Hearing

Chapter 1, Noise Control and Hearing Conservation: Why Do It? This new chapter is a delightful discussion on the need for the preservation of hearing. Too often we can become so involved in details that the purpose of our actions is forgotten. This chapter poignantly reminds us of the importance of hearing.

Chapter 2, Physics of Sound and Vibration. This is a down to earth and easy to understand discussion of the mechanisms of sound and the necessary considerations for proper measurements.

Chapter 3, Sound Measurement: Instrumentation and Noise Dosimeters. I find this chapter well written and up to date. The advice provided is simple and practical. The author is to be commended.

Chapter 4, Anatomy and Physiology of the Ear: Normal and Damaged Hearing. This chapter was originally the work of Dix Ward, who passed away in December 1996. It was honorable that the co-author updated the chapter, keeping Dix as the first author. Where would we be without Dix's famous definitions of Presbycusis, Sociacusis, and Nosoacusis? However, this chapter does contain some strong biases besides the 3-dB/5-dB rule controversy discussed earlier. The section ''0dBHL: Typical Hearing or Normal Hearing?'' misses the point that these numbers are statistical means in any case (HL stands for hearing level in decibels). Because 0 dBHL is the median hearing level of a young normal-hearing population, many young individual workers can be expected to hear better than 0 dBHL. Yet no mention is made about the common practice on only testing to 0 dBHL. Thus a worker may have a shift from −10 to +10 dBHL and still be ignored with respect to STS (standard threshold shift).

Section II: Elements of a Hearing Conservation Program

Chapter 5, Auditory and Nonauditory Effects of Noise. I have already mentioned that two of the three authors are biased against the 3-dB trading relation. However, do not let this detract from the chapter. The chapter is otherwise well written and to the point.

Chapter 6, Program Overview and Administration. This is a somewhat dry but necessary chapter that describes the essential elements to the administration of a hearing conservation program (HCP). The author strongly believes that there must be a key individual in each facility that has responsibility for ''bird-dogging'' the entire HCP. Without this person, the program will fail. A totally contracted program will thus be ineffective. Why? Read the chapter.

Chapter 7, Noise Surveys and Data Analysis. A reasonable discussion of the current thinking in performing noise surveys, up to the point that statistics are introduced. In general, I feel that the use of statistics such as described in the 18 pages after p. 218 is unwarranted. The increased accuracy provided by such a refinement would normally be lost in the expected inaccuracies in the basic measurement.

Chapter 8, Education and Motivation. An important chapter that anyone working in the field of hearing conservation should find valuable. The authors are authorities in this area and bring their experience and wisdom to the reader. They do this with a flare in this well-written chapter.

Chapter 9, Noise Control Engineering. I found this chapter to have a good balance of practical advice and reference material. I believe, however, it was not made clear enough that one is not going to be a noise control engineer from just reading this chapter. And there were no suggestions of asking a professional noise control engineer for help. This is a major oversight.

Chapter 10, Hearing Protection Devices. A useful and informative chapter that is one of the highlights of the manual. I would have liked to see the National Institute for Occupational Safety and Health's (NIOSH) new criteria for derating hearing protection put into this chapter. Certainly there was time to do this as the criteria document itself was published in 1998 and the document itself is referenced on the bottom of p. 452 in the noise manual.

Chapter 11, Audiometric Monitoring Phase of the HCP (Hearing Conservation Program). Julia Royster has conveyed this topic effectively. She has plenty of examples showing the calculation of STS and how to change the baseline. I would like to see a little more on the measurement of thresholds below a hearing level(HL) of 0. Her paragraph in the middle of p. 495 touches on the subject, but leaves the reader to decide if it should be a concern. An illustration could be given of an individual whose hearing threshold shifts from −10 to 10 dB, but because the audiometer only measures to 0 dB, it does not show STS. I believe any audiometer that only allows measurements to 0 dB has no place in an industrial setting.

Chapter 12, Evaluating Hearing Conservation Program Effectiveness. This chapter describes the need and benefit of performing an audiometric database analysis in order to determine the performance of a hearing conservation program. The authors treat this complex subject with enthusiasm, providing what I think is one of the best chapters in the book. Their writing is clear, to the point, and includes many good examples.

Section III: Noise Interference and Annoyance

Chapter 13, Room Noise Criteria. A chapter in the manual that focuses on the acceptance of background levels in spaces such as offices or meeting spaces. This may prove very valuable to some industrial hygienists who are asked to do evaluations other than for hazardous noise. In any case, the chapter is well written and might well be used as a reference by professionals other than industrial hygienists. As might be expected, Blazier introduces his new RC Mark II noise rating.

Chapter 14, Speech Communications and Signal Detection in Noise. Missing from the earlier versions of *The Noise Manual*, this chapter should prove useful for the assessment of workers hearing and communicating in noisy occupations, especially while wearing hearing protection devices. Signal detection, signal localization, and noise masking are discussed. The special problems of workers with existing hearing loss are also included. Noise masking of speech is discussed in detail. All in all, this is a nicely done chapter that needs to be understood by the industrial hygienist.

Chapter 15, Community Noise. The idea here is industrial hygienists often need to advise about community noise issues or find other opportunities in the area. The problem with this concept is putting one chapter on community noise in this manual is like putting one chapter of hearing conservation in a community noise handbook. It is hard to do the subject justice. The authors have done a credible effort, but this chapter would be enhanced if the number of recent standards on the community noise issue were recommended. Being the ASA Standards Director, I am perhaps somewhat prejudiced, but the industrial hygienist would do well to look at the ANSI 12.9 series of six standards on community noise.

Section IV: Regulations, Standards and Laws

Chapter 16, Standards and Regulations. This chapter, coupled with the standards listed in the references for good practice (pp. 740–742), lists most of the key standards for which the industrial hygienist should be aware. I have found especially useful the summary to regulations throughout the world. Standards often go out of date or are revised. The reader is encouraged to periodically visit the ASA or ANSI web sites for the latest information.

Chapter 17, Prediction and Analysis of the Hearing Characteristics of Noise-Exposed Populations or Individuals. The topics introduced and discussed in the latter part of this chapter are for the serious industrial hygienist. The novice would do well just to focus on the earlier part of this chapter, such as Table 17.1, which shows the growth of Noise Induced Permanent Threshold Shift (NIPTS) with exposure level. Likewise, tables 17.2 and 17.3 show the effects of race and gender rather dramatically. These tables also show the effect of screening the database in order to eliminate hearing profiles that do not meet some norm, such as the heavily screened data of table 17.2.

Chapter 18, Workers Compensation. This chapter makes a reasonable attempt to review the practices in North America. Readers will quickly determine that the only consistency among the states and providences of North America is the lack of consistency. It would have been useful to review the compensation practices throughout the world. Maybe the next edition will include world wide practices.

Appendices

There are four appendices. The U.S. Department of Labor Occupational Noise Exposure Standard is one of them. The references for good practice, a useful list of only the best references, is the second. The Properties of Materials and Engineering Conversions is the third. The last is the National Hearing Conservation Association (NHCA) Professional Guide for Audiometric Baseline Revision. This last appendix is especially important if the haring conservation program managers are ever going to bring a degree of uniformity to this practice.

Conclusion

The Noise Manual meets with flying colors it's goal to be ''comprehensive in scope and practical in content.'' The authors that contributed to the manual did so out of love for the subject. I don't believe that anyone interested in conserving hearing could go wrong with the purchase of *The Noise Manual*.

DANIEL L. JOHNSON[a)]
4719 Mile High Drive
Provo, Utah 84604
[a)]Deceased

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.
Reviews of Acoustical Patents

Reviewers for this issue:

GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*
MARK KAHRS, *Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261*
DAVID PREVES, *Micro-Tech Hearing Instruments, 3500 Holly Lane No., Suite 10, Plymouth, Minnesota 55447*
DANIEL R. RAICHEL, *2727 Moore Lane, Fort Collins, Colorado 80526*
CARL J. ROSENBERG, *Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
ERIC E. UNGAR, *Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*

6,367,329

43.20.Hq ACOUSTIC TIME OF FLIGHT AND ACOUSTIC RESONANCE METHODS FOR DETECTING ENDPOINT IN PLASMA PROCESSES

Edward Alois Reitman and Gardy Cadet, assignors to Agere Systems Guardian Corporation
9 April 2002 (Class 73/597); filed 25 August 2000

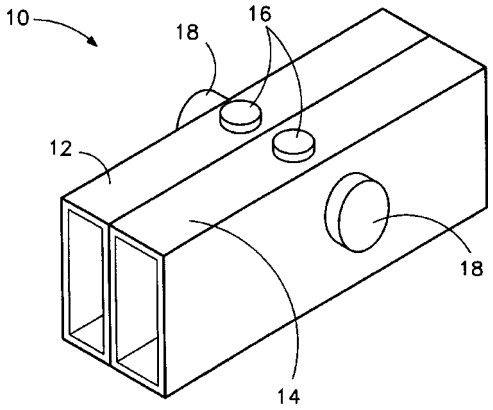
In microelectronic device fabrication, plasma etching is used to form a pattern in a layer of material that has been deposited on a substrate. This etching typically is done in a reaction chamber through which flows a plasma (a partially ionized gas) that interacts with the material that is to be etched. Once the plasma has begun to etch the substrate, the composition of the gas that escapes from the etched area changes. The speed of sound in the gas that leaves the reaction chamber is monitored in order to determine when the deposited layer has been removed. Because plasma etching is done in a vacuum, the gas is compressed before it is introduced in a chamber that is fitted with an acoustic source and receiver.—EEU

6,378,372

43.20.Ks ACOUSTIC RESONANCE ANALYSIS OF GAS MIXTURES

Lawrence J. Karr, Santa Monica, California
30 April 2002 (Class 73/579); filed 15 May 2001

Determination of the composition of binary gas mixtures, such as those used in diving or anesthesia, is accomplished by measuring the resonant frequency of the mixture in a quarter-wavelength volume. The device described in this patent consists of two open-ended half-wavelength tubes, side by side and parallel, with a sending transducer **18** and a receiving



transducer **16** inserted through the tube wall at each tube's midpoint, so as to form a symmetric quarter-wave resonator. The two sending transducers are operated out of phase, in order to minimize the effects of external sounds.—EEU

6,361,747

43.25.Yw REACTOR WITH ACOUSTIC CAVITATION

Jean-Luc Dion and Kodjo Agbossou, assignors to Sonertec Incorporated
26 March 2002 (Class 422/128); filed in Canada 26 May 1998

Acoustic cavitation induced by ultrasound is used for physicochemical treatment of fluid in a flexible tubular enclosure. Electroacoustic transducers that act radially are distributed uniformly around the circumference of the enclosure. Each transducer consists of a driver and an essentially wedge-shaped waveguide, the narrow end of which is coupled to the tubular enclosure via a lubricant film. This arrangement is intended to produce cavitation in the fluid in the central region of the tube without damaging the tube walls or the transducers.—EEU

6,372,195

43.35.Ei MIXED GAS MICROBUBBLE COMPOSITIONS

Ernest G. Schutt *et al.*, assignors to Alliance Pharmaceutical Corporation
16 April 2002 (Class 424/9.52); filed 17 January 1997

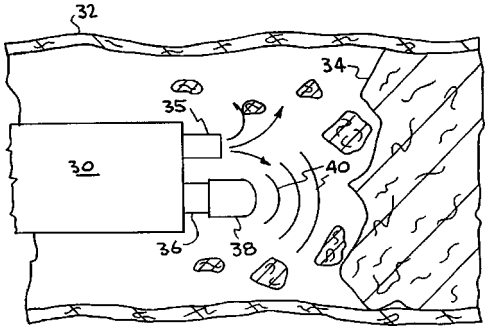
The patent covers a method for preparing stable long-lived microbubbles for ultrasound contrast enhancement or other uses and also for controlling the composition of the prepared bubbles. The method is based on the fact that virtually any microbubble formulation can be prolonged by including an entrapped gas or gas mixture that is selected on the basis of its partial pressure of gases inside and outside of the bubble and on the resulting difference in gas osmotic pressure that opposes bubble shrinkage.—DRR

6,379,325

43.35.Sx OPTO-ACOUSTIC TRANSDUCER FOR MEDICAL APPLICATIONS

William Benett *et al.*, assignors to The Regents of the University of California
30 April 2002 (Class 604/22); filed 9 July 1999

The device is an optically activated transducer for generating acoustic vibrations in a biological medium. The transducer is located at the end of a fiber optic cable that may lie within a catheter. Energy for operating the transducer is provided optically by laser light transmitted to the transducer through the fiber. Pulsed laser light is absorbed in the working fluid of the



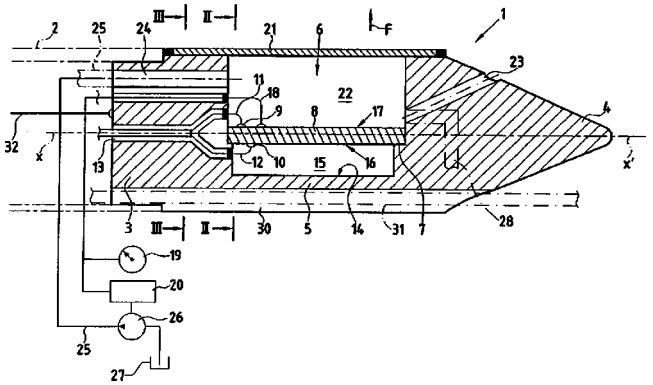
transducer to generate a thermal pressure and consequent adiabatic expansion of the transducer head such that it performs work against the acoustic medium. The transducer returns to its original state through thermal cooling. The motion of the transducer within the ambient medium couples acoustic energy with the medium.—DRR

6,379,320

43.35.Wa ULTRASOUND APPLICATOR FOR HEATING AN ULTRASOUND ABSORBENT MEDIUM

Cyril Lafon *et al.*, assignors to Institut National de la Santa et de la Recherche Medicale I.N.S.E.R.M.
30 April 2002 (Class 601/3); filed in France 11 June 1997

This ultrasonic heat applicator includes an application head having a longitudinal axis and including at least one ultrasonic transducer having a planar emission face and an opposite face, a leakproof membrane which overlays and is spaced from the emission face, and which is transparent to



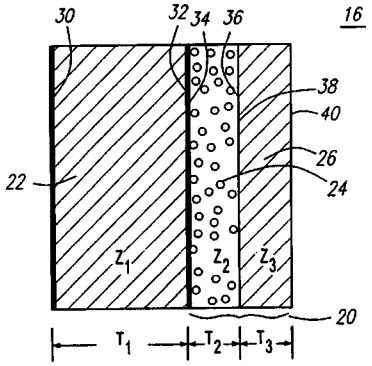
ultrasound, and means for preventing propagation of ultrasound from the opposite face, means for remotely connecting the transducer to an electricity generator, and means for providing ultrasound coupling with the membrane. The planar emission face emits ultrasound waves through the membrane and in a direction that is substantially perpendicular to the emission face, such that the waves emitted are essentially planar and do not diverge in the vicinity of the transducer and the application head.—DRR

6,371,915

43.35.Yb ONE-TWELFTH WAVELENGTH IMPEDANCE MATCHING TRANSFORMER

James D. Koger and Robert W. Crowley, assignors to Scimed Life Systems, Incorporated
16 April 2002 (Class 600/445); filed 2 November 1999

The patent relates generally to ultrasound imaging systems and, more particularly, to an ultrasound transducer that includes a one-twelfth wavelength impedance matching transformer and methods of manufacturing and using the same. The purpose of the transformer is to minimize acoustic



signal attention between a transducer element and an operating medium. The patent also argues for the advantages of a reduced physical size, enhanced diagnostic images, standardized impedance values, and decreased manufacturing costs.—DRR

6,378,376

43.35.Yb ULTRASOUND TRANSDUCER MOUNTING ASSEMBLY

Richard Derman *et al.*, assignors to Life Imaging Systems, Incorporated
30 April 2002 (Class 73/606); filed 25 August 1999

This mounting assembly connects an ultrasonic transducer to an ultrasound machine and provides a means of determining special relationships between a succession of 2D image slices of a target inside a subject. In addition to provisions for mounting a transducer and for engaging a surface

in the proximity of the target, the assembly includes sensing means, in communication with the means to engage a surface, to measure the movement during the acquisition of a succession of 2D image slices. This indicates the necessity for incorporating a tilt sensor and a displacement sensor.—DRR

6,360,609

43.35.Zc METHOD AND SYSTEM FOR INTERPRETING AND UTILIZING MULTIMODE DISPERSIVE ACOUSTIC GUIDED WAVES

Shi-Chang Wooh, assignor to Massachusetts Institute of Technology
26 March 2002 (Class 73/602); filed 15 December 2000

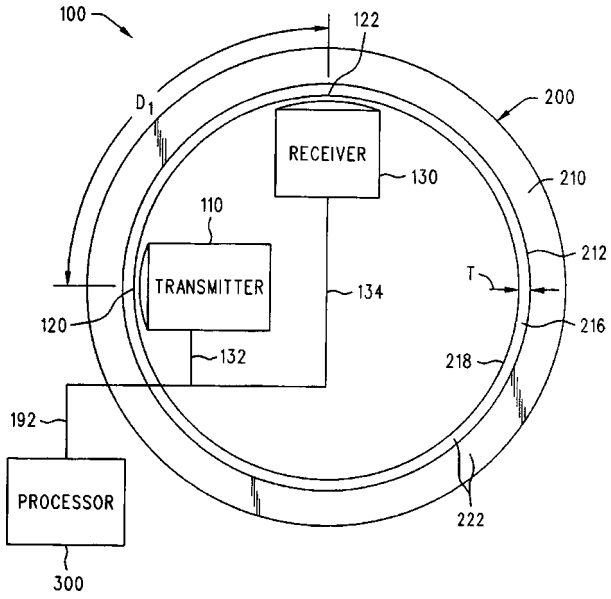
The invention is directed at nondestructive acoustic testing of plates and pipes. To quote from the patent, ''This invention features a method of locating discontinuities in a multimode dispersive medium ...'' Other features include a method of distance measurement, a method of locating a source, and a method of thickness measurement. These are accomplished through fairly sophisticated temporal and spectral computer analysis.—GLA

6,363,787

43.35.Zc APPARATUS AND METHOD FOR MEASURING THE THICKNESS OF A COATING

Nancy M. Carlson *et al.*, assignors to Bechtel BWXT Idaho LLC
2 April 2002 (Class 73/579); filed 13 December 1999

If you build an internal combustion engine with aluminum cylinders, it won't last long. However, if the cylinder bores **212** are coated with a very thin layer of steel **216**, then the engine can be both lightweight and rugged. Obviously the coating must be uniform and exactly the right thickness. The



patent describes an acoustic test procedure in which transmitter **120** induces ultrasonic surface waves into the coating. These are picked up by receiver **130** and the phase velocity is then calculated. A lookup table is used to determine the thickness of the coating.—GLA

6,364,977

43.35.Zc TUNING MECHANISM AND METHOD FOR VIBRATION WELDING

William P. Simon, assignor to Sonics & Materials, Incorporated
2 April 2002 (Class 156/64); filed 5 June 2000

Efficient vibration welding requires operation at the resonant frequency of the welder platen in contact with the workpiece. The object of this patent is to accomplish automatic tuning of the welder to this resonance, doing away with the common manual procedure in which the operator searches for a response maximum. The welder's driving frequency is swept upward and the frequency at which the output of a sensor reaches a preset value is recorded. The process is then repeated, but with downward sweeping of the frequency. The resonant frequency is taken as the arithmetic average of the two recorded frequencies and the welder is set to operate at this frequency.—EEU

6,368,482

43.35.Zc PLATING PROCESSES UTILIZING HIGH INTENSITY ACOUSTIC BEAMS

Richard C. Oeftering and Charles Denofrio, assignors to the United States of America as represented by the Administrator of the National Aeronautics and Space Administration, Washington, DC (US)
9 April 2002 (Class 205/91); filed 19 September 2000

In order to enhance the uniformity of the plating layer or to control its thickness distribution in a desired way, a high intensity acoustic beam is produced in the liquid electrolyte and directed at the object to be plated. The beam, which preferably is at frequencies between 1 and 100 MHz (in order to plate small details as desired), is appropriately focused by shaped faceplates attached to the transducers and/or by use of arrays of transducers.—EEU

6,372,520

43.35.Zc SONIC ASSISTED STRENGTHENING OF GATE OXIDES

Kang-Jay Hsia *et al.*, assignors to LSI Logic Corporation
16 April 2002 (Class 438/4); filed 10 July 1998

Damaged oxide film formed in a semiconductor device is repaired by ''sonic annealing,'' in which the wafer is vibrated or exposed to sound at a predetermined frequency and amplitude for a predetermined period. This type of annealing avoids the damaging side effects associated with common high temperature annealing.—EEU

6,386,038

43.35.Zc ACOUSTIC APPARATUS AND INSPECTION METHODS

Carl Edwin Lewis III, Arlington, Texas and Daric William Escher, Huntsville, Alabama
14 May 2002 (Class 73/587); filed 24 November 1999

The apparatus described here ''listens'' in a passive mode to ultrasound emitted by an object being monitored, so as to detect a signal that could be the result of damage. If it detects such a signal, the apparatus switches to an active mode. In this mode it probes the object with an induced signal to determine whether the sensed signal actually is associated with damage and,

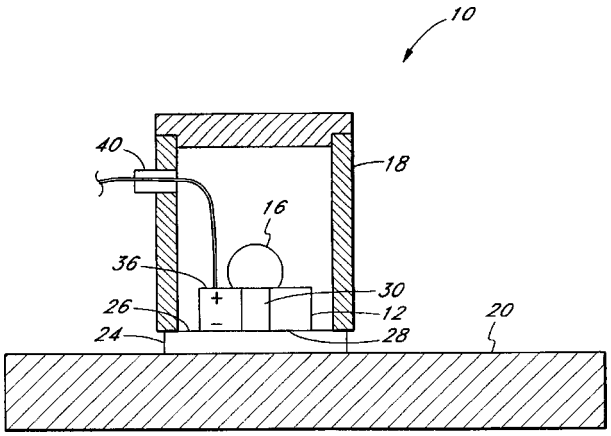
if so, to identify the damage. It then returns to the passive mode, thereby saving wear and energy in an apparatus intended for long-term monitoring.—EEU

6,360,608

43.38.Fx TRANSDUCER FOR MEASURING ACOUSTIC EMISSION EVENTS

Harold L. Dunegan, assignor to Dunegan Engineering Consultants, Incorporated
26 March 2002 (Class 73/587); filed 6 March 2000

Nondestructive acoustic testing of faults in platelike structures has for many years involved sensing and analyzing out-of-plane (OOP) waves produced in response to an impulse source. More recently, the measurement of in-plane (IP) waves above 10 kHz or so has become important. The patent describes a sensing transducer that provides good sensitivity to both types of



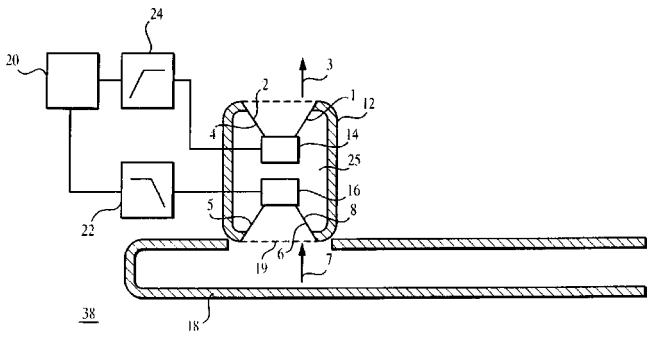
waves. Piezoelectric cylinder **12** is coupled to test sample **20** through ceramic disk **24**. Partial mass loading is provided by ball bearing **16** attached to the top of the cylinder, forcing that portion of the crystal to react in a displacement mode. The patent is clearly written and includes a fair amount of useful background information.—GLA

6,363,157

43.38.Ja MULTIPLE ELEMENT ELECTROACOUSTIC TRANSDUCING

Geoffrey C. Chick, assignor to Bose Corporation
26 March 2002 (Class 381/89); filed 28 August 1997

The illustration is misleading as drawn; ignore high pass filter **24**. At low frequencies, loudspeakers **4** and **5** are intended to operate as a compound transducer driving waveguide **18**. At high frequencies, speaker **4** continues to radiate sound whereas the signal to speaker **5** is attenuated by filter **22**. Known prior art includes a similar arrangement using a vented box instead of a waveguide.—GLA

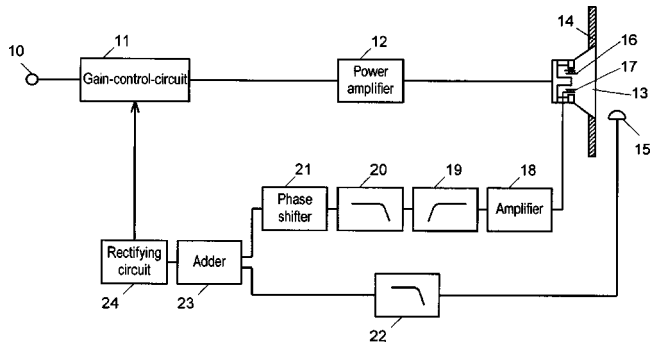


6,370,253

43.38.Ja SOUND REPRODUCING DEVICE

Masahide Onishi *et al.*, assignors to Matsushita Electric Industrial Company, Limited
9 April 2002 (Class 381/98); filed in Japan 1 July 1998

Loudspeaker **16** simultaneously performs the functions of sound reproduction and noise cancellation. To improve upon prior art it includes detect-



ing coil **17**, whose output accurately describes the motion of the loudspeaker cone.—GLA

6,373,955

43.38.Ja LOUDSPEAKERS

Anthony Hooley, assignor to 1... Limited
16 April 2002 (Class 381/335); filed in the United Kingdom 31 March 1995

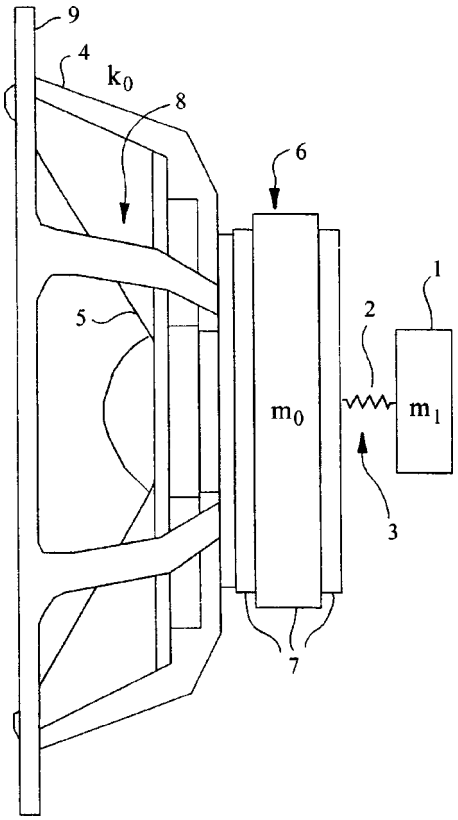
A digital loudspeaker can take the form of an array of individual radiating pistons. The number of pistons actuated during a given clock cycle then determines the resulting sound pressure. The patent describes such an array in which each speaker is independently driven by a ''unary'' digital signal at a common sampling rate. The patent also suggests multiple-layer arrays in which sound energy from rear layers travels through gaps between forward-layer elements. Associated electronic processing includes signal conversion, pulse shaping, and a nondigital volume control. As with most other digital loudspeaker designs, there is no indication that a working model has yet been built.—GLA

6,373,956

43.38.Ja METHOD AND ARRANGEMENT FOR ATTENUATING MECHANICAL RESONANCE IN A LOUDSPEAKER

Ari Varla and Aki Mäkivirta, assignors to Genelec Oy
16 April 2002 (Class 381/353); filed in Finland 12 November 1997

Outboard mass **m1** is coupled to loudspeaker magnet assembly **m0** through lossy spring **2**, thus counteracting the reaction forces transmitted to



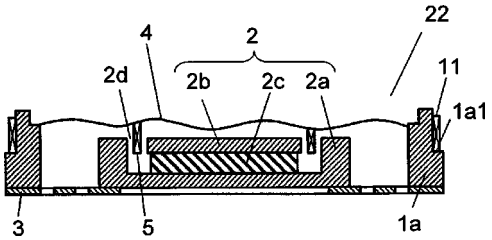
the speaker enclosure. This is certainly not a new idea but may well be the first time it has been applied to loudspeakers.—GLA

6,373,958

43.38.Ja ELECTRICAL/MECHANICAL/SOUND CONVERTER AND APPARATUS OF ELECTRICAL/MECHANICAL/SOUND CONVERSION

Mitsutaka Enomoto *et al.*, assignors to Matsushita Electric Industrial Company, Limited
16 April 2002 (Class 381/406); filed in Japan 29 June 1998

The user of a cell phone or pocket pager may want to choose between an audible ring signal (beep) and a tactile signal produced by vibrating the case of the instrument. Since the resonant frequencies of the two modes vary substantially, it is possible to use a single transducer for both functions by



driving it at the appropriate frequency. However, the actual resonant frequencies may drift slightly in response to environmental conditions. This patent describes a simple transducer having a diaphragm **4** driven by moving coil **5**. An additional sensing coil **11** allows the associated signal generator to zero in on the desired resonant peak.—GLA

6,361,439

43.38.Md GAME MACHINE AUDIO DEVICE AND INFORMATION RECORDING MEDIUM

Yoshinori Kawamoto, assignor to Namco Limited
26 March 2002 (Class 463/35); filed in Japan 21 January 1999

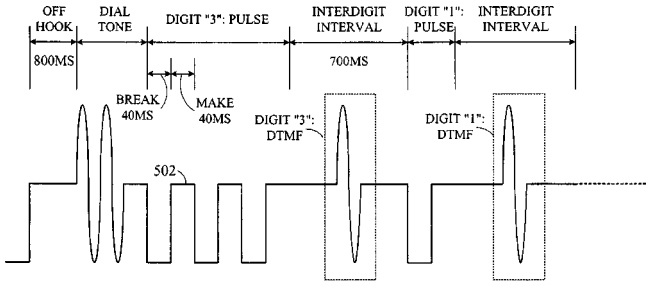
If one listens to two otherwise similar acoustic sources (say, gunshots), the nearer one will be louder. In a gaming environment, louder is better. But this means that the farther source will not produce the desired level of impact and excitement. The invention solves this dilemma by altering spectral response instead. For the farther source, high frequencies are attenuated and lows are accentuated, thus ''...hardly altering the sound volume.''—GLA

6,366,669

43.38.Si SYSTEM AND METHOD FOR PROVIDING UNIVERSAL ACCESS TO INTERACTIVE VOICE RESPONSE SYSTEMS

Iosif Meynekhdrun, East Brunswick, New Jersey
2 April 2002 (Class 379/359); filed 26 April 2000

This patent states that an estimated 60% to 70% of the world's telephones use rotary (DP) rather than tone (DTMF) dialing. Thus, they cannot use interactive voice (IVR) systems. The current pulses produced by a DP phone are generally received only as a series of clicks by the IVR system, making conversion at that point unreliable. Voice recognition support is



expensive and unreliable. Since it would be impractical to change all user telephones, the converter described here is to be installed in the subscriber lines at the central office. It detects the current pulses from a DP phone and inserts the corresponding DTMF tones into the line signal.—DLR

6,366,679

43.38.Vk MULTI-CHANNEL SOUND TRANSMISSION METHOD

Frank Steffen and Matthias Domke, assignors to Deutsche Telekom AG
2 April 2002 (Class 381/356); filed in Germany 7 November 1996

How many discrete channels are needed to envelope a listener in a really convincing three-dimensional sound field? Various experimenters have come up with different answers, but the number is certainly greater than four. Even with seven or eight channels, a fair amount of psychoacoustic trickery must be used. The patent describes a way to improve the stability of phantom sound sources and minimize ''pulling'' toward the nearest loudspeaker.—GLA

6,370,256

43.38.Vk TIME PROCESSED HEAD RELATED TRANSFER FUNCTIONS IN A HEADPHONE SPATIALIZATION SYSTEM

David Stanley McGrath, assignor to Lake DSP Pty Limited
9 April 2002 (Class 381/310); filed in Australia 31 March 1998

In theory, a convincing three-dimensional sound field can be conveyed through two signal channels to a listener wearing headphones. However, to maintain a stable sound field in this virtual environment, the listener's head movements must be tracked so that the two signals can be appropriately modified in real time. This is not a simple process. The inventor proposes a matrix of filters and delay elements to simulate changing head-related transfer functions.—GLA

6,370,964

43.40.Le DIAGNOSTIC LAYER AND METHODS FOR DETECTING STRUCTURAL INTEGRITY OF COMPOSITE AND METALLIC MATERIALS

Fu-Kuo Chang and Mark Lin, assignors to The Board of Trustees of the Leland Stanford Junior University
16 April 2002 (Class 73/862.046); filed 23 November 1999

Critical components such as airplane wings must be regularly inspected, perhaps x-rayed, to detect possible cracks and delamination. Several electronic monitoring schemes have been patented. Most of these utilize arrays of actuators and sensors, either surface-mounted or embedded in the structure itself. This patent describes an improved diagnostic layer that can be incorporated into existing or new structures, both metallic and nonmetallic. It can even monitor the curing process of a laminate material during manufacture. The layer may include separate actuators and sensors, or piezoelectric devices can perform both functions. A stress wave generated by a given actuator/sensor is picked up by neighboring actuator/sensors and their signals are then compared with known signals representing an undamaged condition.—GLA

6,382,079

43.40.Tm VIBRATION SPECIFIC BELLOWS ASSEMBLY

Thomas J. Webb, Hamden, Connecticut
7 May 2002 (Class 92/41); filed 7 October 1999

This patent relates to a sealed bellows assembly as used in aircraft pneumatic bleed pressure controllers. In order for the assembly to survive the vibration environment to which it is exposed when it is used on an aircraft, it is provided with an internal spring to tune its resonances so that they fall outside the band of severe excitation frequencies.—EEU

6,360,486

43.40.Yq WINDOW OBJECT DETECTION SYSTEM WITH VIBRATION COMPENSATION

Joseph Tyckowski, assignor to Meritor Light Vehicle Technology, LLC
26 March 2002 (Class 49/26); filed 16 June 2000

This patent relates to systems that sense obstructions in power-operated windows in terms of changes in the torque, speed, or power of the actuating motor that occur as a window encounters an obstruction. Vibrations of a vehicle traversing a rough road surface may cause the motor to experience changes that exceed a preset threshold limit, thus resulting in the detection of ''false obstructions'' and preventing the window from closing. A system according to the present patent that monitors the motor speed, for example, would note that the motor speed increases during the portion of a cycle in which the window's inertia facilitates closing and would then discount as a false alarm the corresponding decrease in the motor speed during the portion of a cycle in which the inertia force opposes the closing.—EEU

6,360,580

43.40.Yq DEVICE AND METHOD FOR TESTING VEHICLE SHOCK ABSORBERS

Roland Müller, Gersheim, Germany
26 March 2002 (Class 73/11.04); filed in Germany 10 January 1994

The tire of a vehicle is placed on a fork-shaped support and is raised via an arrangement that includes a force and displacement measuring system, permitting determination of the suspension system's spring constant. A cam with adjustable eccentricity, attached to a variable speed motor, is used to apply an oscillatory force to the support via a spring. The motor speed is swept to determine the frequency at which the tire/suspension system resonates and is then kept constant as the cam eccentricity is increased. The suspension's characteristics are determined from measurement of the resulting force and motion.—EEU

6,360,593

43.40.Yq METHOD AND APPARATUS FOR REDUCING VIBRATIONS TRANSMITTED TO A VEHICLE FROM A WHEEL UNIT

Harald Schoenfeld, assignor to Schenck RoTec GmbH
26 March 2002 (Class 73/146); filed in Germany 3 March 1999

This patent describes a machine for balancing automotive wheel/tire assemblies under load. The wheel assembly is mounted on a balancing spindle, whose supports are fitted with sensors to determine the magnitude and location of an unbalance. The tire is rotated at a speed that simulates its general usage while a load that simulates its operating conditions is applied to it via an array of rollers. This roller array is pressed against the tire by a spring and shock absorber arrangement and may be oriented also to load the tire off axis.—EEU

6,361,065

43.40.Yq AIR BAG MODULE AS VIBRATION DAMPER

Ralph Frisch, assignor to TRW Automotive Safety Systems GmbH & Company KG
26 March 2002 (Class 280/728.2); filed in Germany 5 February 1999

This patent describes mounting of an air bag module on a motor vehicle steering wheel elastically, so that the module in essence works as a dynamic absorber.—EEU

6,366,531

43.40.Yq METHOD AND APPARATUS FOR ACOUSTIC LOGGING

Georgios L. Varsamis *et al.*, assignors to Dresser Industries, Incorporated
2 April 2002 (Class 367/26); filed 22 September 1998

This patent relates to a tool used in hydrocarbon exploration. Like many similar tools, it uses acoustic sources and receivers to determine acoustic properties of geologic formations traversed by the tool. A salient feature of the tool described here is the arrangement of transducers around the circumference of the tool's tubular casing; the transducers are arranged in diametrically opposite pairs and are phased to act as dipoles, quadrupoles, or sextupoles.—EEU

6,367,324

43.40.Yq DETECTING DEVICE

Ryouei Nozawa and Akira Sugiyama, assignors to Fuji Photo Film Company, Limited
9 April 2002 (Class 73/290 V); filed in Japan 31 March 1999

This device, intended to detect the level of liquid in a tank in order to prevent its overflowing, consists of a casing in the shape of an inverted ''U.'' One leg of the U contains a sound generator, and the other, a receiver. The device is mounted in the tank so that the liquid covers the sound generator and receiver when it reaches a given level, resulting in a great increase in the received sound signal. The tank fill control is activated on the basis of this increased signal.—EEU

6,377,900

43.40.Yq MEASURING SYSTEM FOR TRANSFER FUNCTION MATRIX OF A SYSTEM TO BE CONTROLLED IN MULTI-DEGREE OF FREEDOM VIBRATION CONTROL

Kazuyoshi Ueno and Yoshikado Yamauchi, assignors to IMV Corporation
23 April 2002 (Class 702/109); filed in Japan 12 May 1998

This patent is relevant to multi-degree-of-freedom vibration testing and to control of vibration of a test object excited by multiple vibrators. A sinusoidal signal is applied to multiple vibrators with random phase shifts. The auto-spectra and cross-spectra of the excitation signals and of the response signal at a selected response point are calculated. This step is repeated two or more times and the spectra are averaged arithmetically and used to calculate the transfer functions. The foregoing steps are repeated at each frequency.—EEU

6,386,042

43.40.Yq METHOD AND APPARATUS FOR TWO-DIMENSIONAL VIBRATION ANALYSIS

Michael Wörtge and Matthias Schüssler, assignors to Polytec GmbH
14 May 2002 (Class 73/657); filed in Germany 16 February 1998

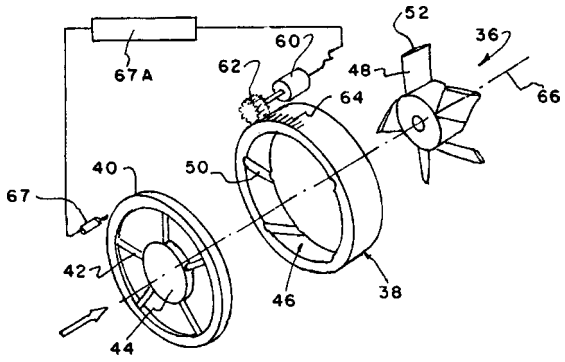
In earlier methods for contactless scanning of a two-dimensional vibrating object, a video image of the object is recorded and a rectangular grid of measurement points is superposed on the image by a computer. The object then is set into vibration and the laser beam of an interferometer is directed toward the grid points under control of a computer. The vibration spectrum for each of these points is recorded. The spectra are analyzed and the object's vibration shapes at selected frequencies are displayed on an image screen. The method described in this patent is more versatile in that it permits selection of measurement points at all desired locations, not just on a rectangular grid, and permits consideration of points that are not in the scanned plane.—EEU

6,375,416

43.50.Gf TECHNIQUE FOR REDUCING ACOUSTIC RADIATION IN TURBOMACHINERY

Kevin J. Farrell and Walter S. Gearhart, both of State College, Pennsylvania
23 April 2002 (Class 415/119); filed 15 July 1993

This technique is aimed at reducing the blade-rate frequency (BRF) tones that radiate from turbomachinery. The inner surface of the shroud surrounding the blades is contoured circumferentially so that the tip clearance between each blade and the shroud's inner surface generates a periodic unsteady pressure field which is substantially equal to, and out of phase with, an existing periodic pressure field resulting from nonuniform flow into the blades.—DRR



6,381,834

43.50.Gf METHOD FOR DAMPENING ACOUSTICAL NOISE IN A DRY-TYPE TRANSFORMER

Scott Edwards May, assignor to Square D Company
7 May 2002 (Class 29/606); filed 22 October 1997

Inserting sound absorbing pads between the low voltage coil and the transformer's iron core lowers the noise from a dry-type power transformer. The pads function as baffles to lower the acoustical noise caused by magnetostriction and other forces when the transformer is energized.—CJR

6,386,316

43.50.Gf DEVICE FOR CONTROLLING THE POROSITY OF A HELICAL, FLEXIBLE, SOUND-ABSORBING AIR-CONDUCTING TUBE

Edmond Pol Jean Lepoutre, Curitiba-Parana, Brazil
14 May 2002 (Class 181/224); filed in Brazil 20 April 1999

This patent in essence pertains to tubing that transmits air in automotive air conditioning installations. Such tubing generally consists of inner and outer tubes, each of which is constructed from a helix made of a strip of porous material with an approximately sinusoidal cross section. In order to control the porosity of the assembly, a strip of nonporous material is wound around the inner tube in the shape of a helix with a relatively large pitch.—EEU

6,386,317

43.50.Gf SOUND-ABSORBING DUCT STRUCTURE

Katsumi Morohoshi *et al.*, assignors to Nissan Motor Company, Limited
14 May 2002 (Class 181/252); filed in Japan 21 December 1998

This is a sound-absorbing duct structure for the intake to an automotive engine. The expanded part of the duct has small holes that open to the area for passage of the gas. This sound-absorbing section of the duct includes a piezoelectric material, and an electrically conductive material in contact with the piezoelectric material to reduce additional (and low frequency) noise.—CJR

6,374,747

43.50.Jh SOUND ABATEMENT METHODS AND CONVEYOR TROLLEYS

Deepak Devnani *et al.*, assignors to Jervis B. Webb Company
23 April 2002 (Class 104/162); filed 5 May 2000

In a moving trolley, some points of impact generate more noise than do others. This sound abatement method capitalizes on this fact by judiciously applying a sound deadening material to a noisy part of the trolley.—CJR

6,376,396

43.55.Ev SOUNDPROOFING MATERIAL AND THE USE THEREOF

Udo Thorn *et al.*, assignors to Beloh Beteiligungsgesellschaft mbH
23 April 2002 (Class 442/136); filed in Germany 28 February 1997

This sound absorptive element is made of nonwoven materials that contain thermoplastic fibers. The fibers are permanently compacted both by mechanical means and by pressure/heat treatment to have a specific flow resistance. To take advantage of the flow resistance properties of the material, it must be spaced away from a hard wall surface by some distance.—CJR

6,382,350

43.55.Ev MOLDED ACOUSTIC AND DECORATIVE MATS AND METHODS FOR FORMING THE SAME

Michael C. Jezewski *et al.*, assignors to Collins & Aikman Products Corporation
7 May 2002 (Class 181/290); filed 2 April 2001

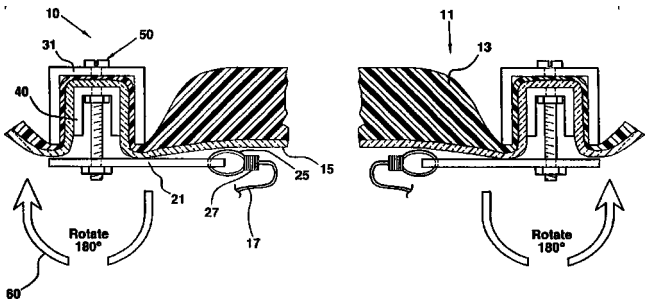
This multi-layered, acoustically absorptive mat includes a base layer (with holes), a face layer, and an acoustic absorbing layer between the two. The acoustic absorbing layer spans across the holes of the base layer. This may be used to absorb noise inside an automobile.—CJR

6,386,263

43.55.Ev SUSPENDED FLEXIBLE SOUND ABSORBING BANNER

Carroll Finley Tomlinson, Charlotte, North Carolina and Eric Von Fange, Fort Mill, South Carolina
14 May 2002 (Class 160/328); filed 24 May 2000

This banner is made of glass fiber or other absorptive material. It is



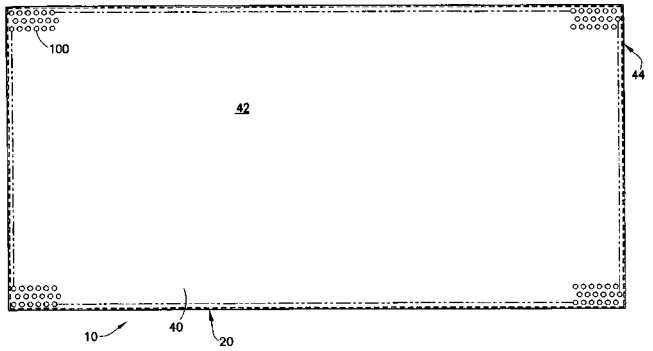
covered with a protective material that is stretched under tension when held in a frame.—CJR

6,382,351

43.55.Ti SOUND REDUCING PANEL FOR ANIMAL HOUSING ROOMS

Philip A. Lastowski, assignor to Lab Products, Incorporated
7 May 2002 (Class 181/295); filed 5 November 1999

This panel comprises a housing, an acoustic baffle, and acoustic netting. The baffle is covered with a waterproof protective casing, which protects it from dirt and moisture without interfering with its acoustic properties



and allows it to be cleaned with fluids. The housing comprises an inner and outer pan, which are hinged together. The housing protects the baffle.—CJR

6,387,470

43.55.Ti SOUND DEADENING AND STRUCTURAL REINFORCEMENT COMPOSITIONS AND METHODS OF USING THE SAME

Chin-Jui Chang and Gerald Fitzgerald, assignors to Sika Corporation
14 May 2002 (Class 428/121); filed 5 November 1998

The patent describes an expandable, lightweight, high strength sealant. The material can seal hollow structural members of vehicles, substantially decreasing the noise that travels along the length of those members as well as strengthening those members with minimal increases to their weight. The compositions can be formed into free-standing, self-sustaining parts or into U-shaped members supported on lattice-type nylon supports.—CJR

6,370,958

43.58.Kr METHOD OF MEASURING THE VIBRATION DAMPING CAPABILITY

Robert P. Uhlig, assignor to DaimlerChrysler Corporation
16 April 2002 (Class 73/662); filed 11 December 2000

This is the latest of a series of patents intended mainly to predict the noise level (‘‘squeal’’) that an automobile brake rotor is likely to generate. A rotor can behave like a bell, and its squeal factor may well be related to the decay of various bell-mode harmonics. ‘‘It is believed that by subjecting the test part to a range of different frequency input signals, the computed Q-factor will be a reliable representation of the damping capability of the test part.’’ Why seemingly identical rotors should have significantly different internal damping is not addressed in the patent document.—GLA

6,381,270

43.58.Vb CABLE EQUALIZER FOR AES DIGITAL AUDIO DATA

Donald S. Lydon and Charles S. Meyer, assignors to ADC Telecommunications, Incorporated
30 April 2002 (Class 375/230); filed 7 January 1999

An electrical signal is equalized at the receiving end of a transmission path by applying the signal both to a frequency-dependent emphasizer and to one port of a mixer and applying the output signal of the emphasizer to another port of the mixer. The mixer combines the signals according to a mixer control signal. Another control signal is generated from the mixer output signal, whose amplitude depends on the absolute value of the derivative of the mixer output during a defined bit cell interval.—DRR

6,381,333

43.58.Vb SOUND PROCESSING CIRCUIT

Tatsuya Suzuki, assignor to Matsushita Electric Industrial Company, Limited
30 April 2002 (Class 381/18); filed in Japan 20 January 1997

This sound processing circuit distributes low-frequency signal components of multichannel audio signals. The circuit filters low-frequency signal components from the digital audio signals of m ($m \leqslant n$) specified channels using an exclusive channel having low frequency signals and n ($n \geqslant 1$) multiple independent channels, and outputs the low-frequency signal components as part of a low-frequency signal channel. The sound processor circuit consists of m high-pass filters to receive digital audio signals from m specified channels and allows the signal components having a frequency higher than a cut-off value to pass.—DRR

6,385,239

43.58.Vb ADAPTIVE EQUALIZING CIRCUIT

Toshinori Okamoto and Youichi Ogura, assignors to Matsushita Electric Industrial Company, Limited
7 May 2002 (Class 375/232); filed in Japan 2 February 1999

The device is a digitally implemented adaptive equalization circuit suitable for equalization of a nonlinear signal. The device performs high-order partial response equalization so that the equalization error of a reproduced signal having nonlinear distortions is minimized when it is reproduced from a high-density recording medium.—DRR

6,370,083

43.60.Bf METHOD AND APPARATUS FOR DETECTING ACOUSTIC SIGNALS FROM A PLURALITY OF ACOUSTIC SIGNAL SENSORS

Carl R. Foreman, assignor to the United States of America as represented by the Secretary of the Navy
9 April 2002 (Class 367/126); filed 10 August 2000

Each of several channels includes a sensor and circuitry for amplifying the received signal, removing the dc component, and low-pass filtering. The outputs of all channels are summed to form a single signal, which is processed to provide a differential signal if at least one channel senses an acoustic signal. This relatively simple arrangement is claimed to provide redundancy to increase the probability that acoustic signals will still be detected by the system even if it experiences component failures.—EEU

6,379,314

43.66.Sr INTERNET SYSTEM FOR TESTING HEARING

Gerald Horn, assignor to Health Performance, Incorporated
30 April 2002 (Class 600/559); filed 19 June 2000

A computer presents over the Internet a series of screens that ask for entry of name, age, social security number, etc. and instructs how to respond to sounds generated by the speaker in the user's computer or via headphones plugged into the user's computer. The user indicates detecting a sound by a mouse click or by depressing a key on the keyboard. Calibration of sounds presented is achieved by reference to a table of pulsed pure tone sound levels just barely heard by normal hearing individuals using five most commonly used PC sound cards. The effect of different levels of background noise in the test environment is said to be minimized by setting a ''0 dB'' level for which the sound is just barely heard.—DAP

6,381,308

43.66.Ts DEVICE FOR COUPLING HEARING AID TO TELEPHONE

Charles H. Cargo *et al.*, all of Idaho Falls, Idaho
30 April 2002 (Class 379/52); filed 2 December 1999

Normally when coupling signals from a telephone receiver to a hearing aid using inductive means, the induction coil inside the hearing aid senses the electromagnetic signal radiated from a telephone handset. In this patent, however, for persons with moderate to severe hearing loss, telephone signals are amplified and coupled inductively to a hearing aid via an induction loop or other wireless link, while enabling the wearer to hear the audio signal simultaneously, as is done normally, from the telephone handset. This implementation allows the telephone signal to be coupled via wireless means into both hearing aids of a binaural fitting. The device described is inserted between the telephone handset and the telephone base and has a manually activated power switch and power-saving automatic shutoff circuitry.—DAP

6,385,323

43.66.Ts HEARING AID WITH AUTOMATIC MICROPHONE BALANCING AND METHOD FOR OPERATING A HEARING AID WITH AUTOMATIC MICROPHONE BALANCING

Fred Zoels, assignor to Siemens Audiologische Technik GmbH
7 May 2002 (Class 381/313); filed in Germany 15 May 1998

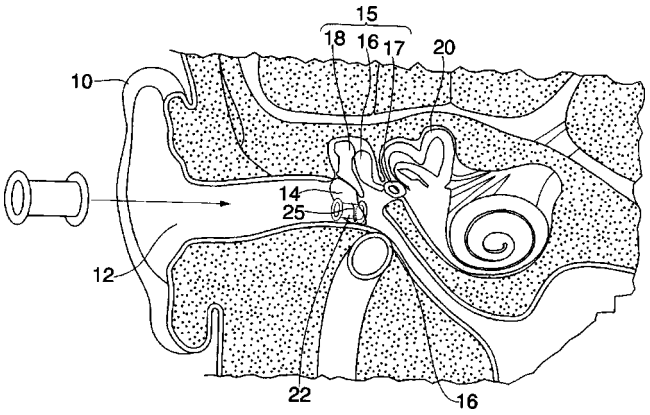
Hearing aids having directional microphone systems made by subtracting the processed outputs of two omnidirectional microphones matched in sensitivity and phase may have decreased directional performance due to temperature effects and microphone aging. The described system subtracts the average values of the microphone output signals and forms a correction signal based essentially on the deviation of the actual outputs from the average outputs. Correction circuitry controls the amplification characteristics of one or both microphones until the deviation from average is reduced to zero.—DAP

6,387,039

43.66.Ts IMPLANTABLE HEARING AID

Ron L. Moses, assignor to Ron L. Moses
14 May 2002 (Class 600/25); filed 4 February 2000

A biocompatible, flanged, tubular capsule containing components for an air conduction hearing aid is inserted partially through the eardrum following an outpatient surgical procedure. Unlike traditional hearing aids, more than one speaker or receiver is used. The part of the capsule containing the hearing aid microphone inlet projects into the ear canal while the portion



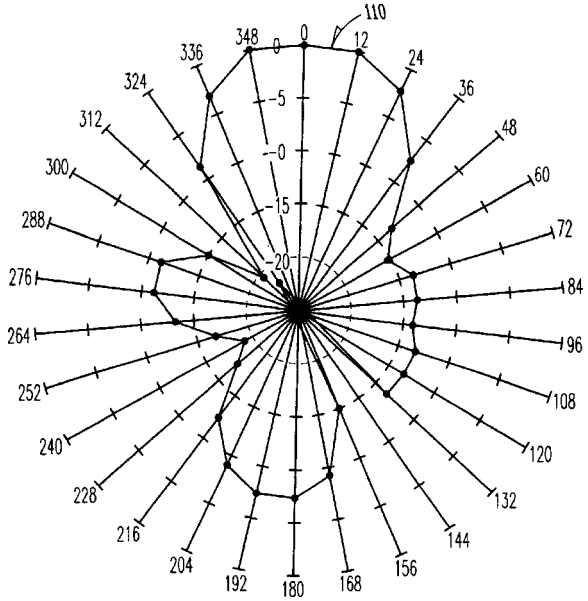
of the capsule housing the outlets of the hearing aid receivers projects into the middle ear. At least one other hearing aid receiver is said to be in direct contact with the eardrum. Amplified sound from the capsule also flows into the middle ear volume and vibrates the ossicular chain. In one embodiment, a rechargeable power source lasts about as long as the capsule is retained in place in the eardrum, typically 1–2 years. Thereafter, the hearing aid is either discarded or recharged via wireless means. Alternatively, the hearing aid components are powered in real time from electromagnetic energy received from a remote transmitter. Capability is provided for wireless ear-to-ear communication between hearing aids on left and right ears. Parameters in the hearing aid circuitry may also be adjusted via a wireless remote control.—DAP

6,389,142

43.66.Ts IN-THE-EAR HEARING AID WITH DIRECTIONAL MICROPHONE SYSTEM

Lawrence T. Hagen and David A. Preves, assignors to Micro Ear Technology
14 May 2002 (Class 381/313); filed 31 March 1998

Two sides of an in-the-ear (ITE) bilateral routing of signals (BICROS) fitting are described in which each side can operate separately under wearer control in either a first-order gradient directional mode or an omnidirectional mode. The outputs of the two microphone systems are summed to form a



second-order directional BICROS fitting when the two sides are operated in directional mode. The polar directivity pattern provided may be varied by adjusting amplitude and phase delay. Connection between the two sides may be made via either hard-wired or by wireless means such as a rf link.—DAP

6,386,314

43.66.Yw FLEXIBLE EARPHONE ASSEMBLY FOR USE DURING HEARING SCREENING

Neil J. Sheehan *et al.*, assignors to Natus Medical, Incorporated
14 May 2002 (Class 181/129); filed 14 September 1999

The disclosed earphone assembly includes an annular baffle and an interior wall which defines a central chamber. A transparent flexible lid is attached to the baffle along the outer surface, completely covering the central chamber and forming an acoustic seal with the baffle. It is claimed that the flexibility of the earphone is improved by including hinges in the lid and by providing cuts, or other discontinuities, in the baffle.—DRR

6,359,988

43.70.Dn PROCESS FOR INTRODUCE REALISTIC PITCH VARIATION IN ARTIFICIAL LARYNX SPEECH

Carol Espy-Wilson, assignor to Trustees of Boston University
19 March 2002 (Class 381/70); filed 29 August 2000

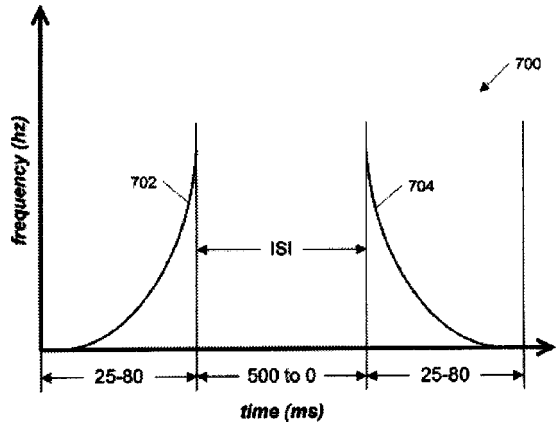
A typical external laryngeal excitation device is held against the skin of the neck and emits a tone which excites the vocal tract, allowing some speech. Even when such devices include a frequency control, users find it cumbersome and much inferior to normal fundamental (*F*0) frequency control. This short patent describes the perceptual improvement obtained by replacing the laryngectomee's *F*0 pattern with that from a speaker with normal speech. The two claims cover the cepstral convolution process used to do the substitution. A method for computing a supposedly suitable replacement *F*0 pattern is mentioned, but is not included in the claims. How such an *F*0 generator could work without any phonetic or syntactic information is not addressed.—DLR

6,358,056

43.71.Ky METHOD FOR ADAPTIVELY TRAINING HUMANS TO DISCRIMINATE BETWEEN FREQUENCY SWEEPS COMMON IN SPOKEN LANGUAGE

William M. Jenkins *et al.*, assignors to Scientific Learning Corporation
19 March 2002 (Class 434/185); filed 21 June 2000

Yet another in a long series of devices to aid in overcoming auditory perceptual deficiencies related to speech perception, this one tests for perception of a sequence of up/down frequency sweeps. The figure shows a two-sweep pattern with correct answer ''up–down.'' Sequences are typically



longer than this and are presented in a video-game-like scenario. Correct subject responses result in shorter tone durations and/or shorter interstimulus intervals (shown as **ISI** in the figure) as the game proceeds.—DLR

6,364,666

43.71.Ky METHOD FOR ADAPTIVE TRAINING OF LISTENING AND LANGUAGE COMPREHENSION USING PROCESSED SPEECH WITHIN AN ANIMATED STORY

William M. Jenkins *et al.*, assignors to Scientific Learning Corporation
2 April 2002 (Class 434/156); filed 30 June 1998

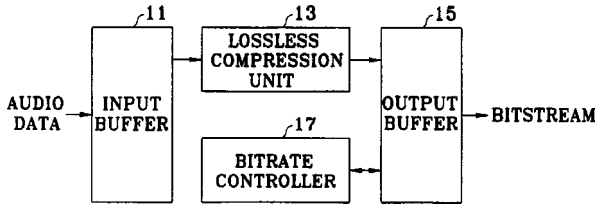
Like United States patent 6,358,056, reviewed above, this one covers another device to help restore auditory perceptual deficiencies. Venturing into harder problems, this device tests for ''increasingly difficult sentence structures.'' A story is narrated and questions are asked via a synthesizer which allows manipulation of the formant frequency sweeps. The subject responds according to whether words are understood to refer to the correct character or object in the story. Synthesis details are adjusted according to the subject's responses.—DLR

6,385,587

43.72.Gy CONSTANT BITRATE REAL-TIME LOSSLESS AUDIO ENCODING AND DECODING BY MOVING EXCESS DATA AMOUNTS

Jae-Hoon Heo, assignor to Samsung Electronics Company, Limited
7 May 2002 (Class 704/503); filed in the Republic of Korea 6 May 1998

A lossless audio encoding and decoding apparatus is designed for real time application by utilizing an input buffer and lossless compression unit in



the encoder and a lossless input buffer and lossless restorer in the decoder. A controller combines data amounts exceeding a maximum bitrate with data amounts less than the maximum bitrate.—DAP

6,366,879

43.72.Ja CONTROLLING INTERACTIVE VOICE RESPONSE SYSTEM PERFORMANCE

Philip Randall Coxhead *et al.*, assignors to International Business Machines Corporation
2 April 2002 (Class 704/201); filed in the United Kingdom 5 October 1998

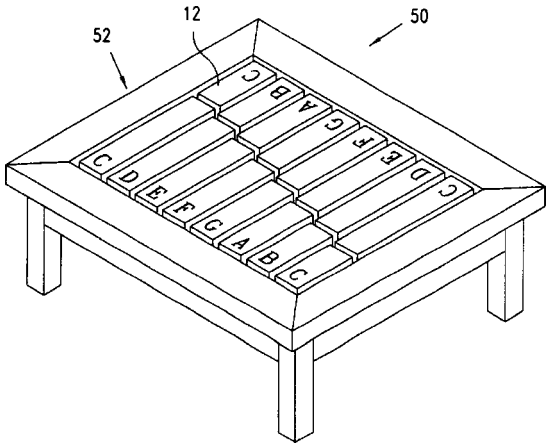
This patent describes a system for managing the playback of speech segments over the phone for use in an interactive voice response system. Multiple playback channels are supported in this multiline system. An efficient buffer management scheme is intended to eliminate gaps in the playback and special recovery actions are taken if such a ''buffer underrun'' condition occurs.—DLR

6,376,757

43.75.Kk FURNITURE INCORPORATING A PERCUSSION INSTRUMENT AND METHOD FOR FORMING THE SAME

Tor C. Clausen, Olympia, Washington
23 April 2002 (Class 84/402); filed 26 October 2000

The idea is simple enough: why not use a chair bottom as a xylophone? Or a bench, or a table, or possibly a door window... Of course, the



keys must be of differing lengths or thickness. Unfortunately, notions of stress seem to be completely missing from the discussion.—MK

6,380,473

43.75.Mn MUSICAL INSTRUMENT EQUIPPED WITH SYNCHRONIZER FOR PLURAL PARTS OF MUSIC

Haruki Uehara, assignor to Yamaha Corporation
30 April 2002 (Class 84/609); filed in Japan 12 January 2000

The Yamaha Disklavier™ again. This time, the problem is how to provide an automatic accompaniment to a player who has a different tempo than the perfect computer. By inserting ''cue flags'' into the score, the machine can coarsely resync. Given the current state of the art, this is crude technology.—MK

6,380,472

43.75.St ELECTRIC TUTOR FOR DIRECTLY INDICATING MANIPULATORS TO BE ACTUATED, MUSICAL INSTRUMENT WITH BUILT-IN ELECTRIC TUTOR, METHOD FOR GUIDING FINGERING AND INFORMATION STORAGE MEDIUM FOR STORING PROGRAM REPRESENTATIVE OF THE METHOD

Nobuo Sugiyama *et al.*, assignors to Yamaha Corporation
30 April 2002 (Class 84/478); filed in Japan 25 December 1998

Teaching any instrument requires feedback from the student's performance. Here, Yamaha proposes using electromagnetic controls to partially depress the correct keys, thereby leading the student. Mechanically, this can be viewed as a follow-on to the large number of Disklavier™ patents.—MK

6,388,183

43.75.Wx VIRTUAL MUSICAL INSTRUMENTS WITH USER SELECTABLE AND CONTROLLABLE MAPPING OF POSITION INPUT TO SOUND OUTPUT

Stephen M. Leh, assignor to Leh Labs, L.L.C.
14 May 2002 (Class 84/645); filed 7 May 2001

The origin of the virtual baton dates back to Max Mathews at Bell Labs. The proposed improvement is to divide the 2D space into different instrumental control spaces. Furthermore, this can be customized by the user since it's just a soft XY map.—MK

6,386,201

43.80.Ka APPARATUS FOR PREVENTING SNORING

Bijan Golriz Fard, Hannover, Germany
14 May 2002 (Class 128/848); filed 4 May 2001

This device for preventing snoring includes a pillow divided internally into separate air chambers, a sound sensor, an air pressure source, an air pressure reducing unit, and a monitoring unit. The monitoring unit is activated by snoring noises sensed by the sound sensor. This establishes the position of the sleeper's head on the pillow by sensing pressure increase on the affected air chambers of the pillow resulting from the weight of the sleeper's head. In response, the air pressure source and air reduction sources are coordinated on selected pillow chambers to effect a change in the position of the sleeper's head on the pillow, which may cause an attenuation or cessation of snoring.—DRR

6,377,693

43.80.Nd TINNITUS MASKING USING ULTRASONIC SIGNALS

Arnold S. Lippa and James A. Nunley, assignors to Hearing Innovations Incorporated
23 April 2002 (Class 381/71.6); filed 23 June 1994

This patent involves a method and apparatus for treating tinnitus. A noise signal is generated to mask the ringing or buzzing in the ears caused by tinnitus and is transposed into the ultrasonic frequency range. It is claimed that the signal effectively masks the tinnitus noise without interfering with the subject's perception of normal sounds such as human speech. In another embodiment, the human speech is transduced into electrical signals, transposed to the ultrasonic frequency range, and physically applied to the patient while tinnitus masking signals in the auditory range are applied to the patient.—DRR

6,388,949

43.80.Nd MARINE TURTLE ACOUSTIC REPELLENT/ALERTING APPARATUS AND METHOD

Martin L. Lenhardt, assignor to Sound Technique Systems LLC
14 May 2002 (Class 367/139); filed 29 August 2000

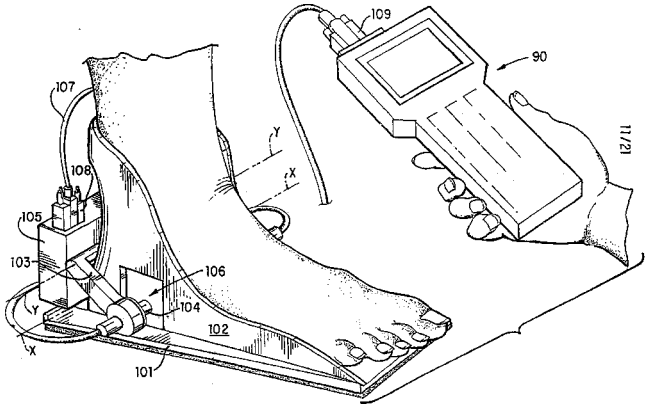
This turtle alerting system and method utilizes an acoustic signal that might be in a low-frequency range, a mid-frequency range, and a high-frequency range. One or more of those acoustic signals may be accompanied and/or modulated by either an ultrasonic signal or a MHz signal, to provide a carrier for the acoustic signal. Moreover, a beam of light may be used as another type of alerting sensor. Sonar may also be used to steer the system in a direction where turtles may be located.—DRR

6,371,916

43.80.Qf ACOUSTIC ANALYSIS OF BONE USING POINT-SOURCE-LIKE TRANSDUCERS

Joe P. Buhler *et al.*, assignors to Metra Biosystems, Incorporated
16 April 2002 (Class 600/449); filed 3 September 1999

In the attempt to improve the measurement of the characteristic behavior of an acoustic wave in the bone of a patient, a preferred embodiment of this apparatus includes two transducers and a way to mount them in a spatial relationship with respect to the bone. One transducer transmits acoustic energy over a broad solid angle, thereby behaving as a point source of



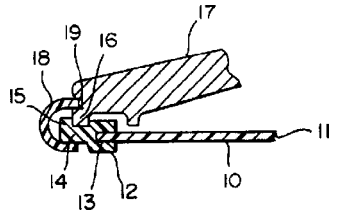
acoustic energy. The other collects acoustic energy over a broad solid angle, thereby acting as a point receiver. A signal processor in communication with the latter transducer provides a measurement that is a function of the transient spectral or temporal components of the signal received.—DRR

6,378,648

43.80.Qf DIAPHRAGM ASSEMBLY FOR STETHOSCOPE CHEST PIECE

Marc S. Werblud, Kirkland, Washington
30 April 2002 (Class 181/131); filed 6 March 2000

A diaphragm assembly consists of a diaphragm adhesively attached to a rim. The rim has a C-shaped cross section with the C-shape facing inward and a membrane extending radially inward from one of the radial portions of the C-shaped section. The diaphragm is attached to the membrane. The



thickness of the membrane is approximately two-tenths of the thickness of the C-shaped section. The groove formed by the C-section engages a rib on a stethoscope chest piece to accommodate the diaphragm assembly.—DRR

6,379,305

43.80.Qf EARLY-FETAL-HEARTBEAT-DETECTION DEVICE AND METHOD

Janet Stoner Eugley, St. George, Maine
30 April 2002 (Class 600/453); filed 5 April 2000

This patent describes a Doppler fetal heartbeat measurement monitor and method for detection of fetal heartbeat and measurement of its rate from seven- to twelve-week gestation. The device is a conventional Doppler base unit incorporating an intravaginal probe. It is claimed that the device would

eliminate the need for the more expensive procedure of transvaginal ultrasonic imaging.—DRR

6,381,350

43.80.Qf INTRAVASCULAR ULTRASONIC ANALYSIS USING ACTIVE CONTOUR METHOD AND SYSTEM

Jon D. Klingensmith *et al.*, assignors to The Cleveland Clinic Foundation
30 April 2002 (Class 382/128); filed 2 July 1999

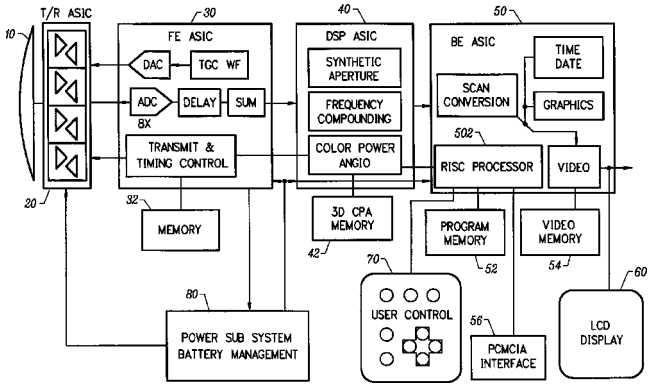
An intravascular ultrasound (IVUS) analysis system and method provides for the determination of luminal and medial-adventitial boundaries of a blood vessel. Ultrasonic data is acquired by a rotating transducer mounted to a tip of a catheter that is inserted into the blood vessel. An intravascular image is reconstructed from the ultrasound data. A boundary contour is obtained, and the properties of the blood vessel are analyzed including determination of the area of the lumen and percent of occlusion caused by plaque.—DRR

6,383,139

43.80.Qf ULTRASONIC SIGNAL PROCESSOR FOR POWER DOPPLER IMAGING IN A HAND HELD ULTRASONIC DIAGNOSTIC INSTRUMENT

Juin-Jet Hwang *et al.*, assignors to SonoSite, Incorporated
7 May 2002 (Class 600/441); filed 22 October 1999

This is a hand-held ultrasonic portable instrument that performs both B mode and Doppler imaging. The instrument includes a transducer array mounted in an enclosure, with an integrated circuit transceiver connection to the elements of the array for the reception of echo signals. A digital signal processing circuit performs both B mode and Doppler signal processing



such as filtering, including wall filtering, matrix and Hartley transform matrix functions, detection and Doppler estimation, as well as advanced functions such as assembly of multiple zone focused scanlines, synthetic aperture formation, depth-dependent filtering, speckle reduction, flash suppression, and frame averaging.—DRR

6,383,141

43.80.Qf ULTRASOUND TRANSDUCER

Hiromu Itoi, assignor to Fuji Photo Optical Company, Limited
7 May 2002 (Class 600/459); filed in Japan 4 March 1999

This is a design for a transducer having its elements arrayed in a polygonal shape on a common plane. The polygonal array of transducer elements includes a core unit consisting of a single transducer element of a polygonal shape and transducer elements arranged in a polygonal array around the core unit. Each one of the transducer elements of the frame units is formed in the shape of a square block and electrically connected to adjacent transducer elements within the same frame unit but insulated from the transducer elements of an adjacent frame unit through the presence of an insulating coat layer.—DRR

6,383,142

43.80.Qf SOUND VELOCITY FOR LUNG DIAGNOSIS

Noam Gavriely, assignor to Karmel Medical Acoustic Technologies Limited
7 May 2002 (Class 600/529); filed 6 October 2000

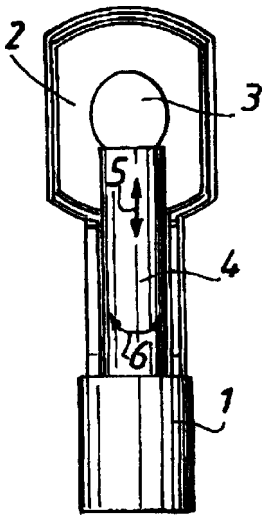
A method and apparatus is disclosed for determining the air pressure in a lung by measuring the velocity of sound in the lung and estimating the pressure based on the measured velocity. The method also includes determining a difference in phase for sound between two positions with respect to the lung.—DRR

6,371,903

43.80.Sh THERAPY PROBE

Emmanuel Blanc *et al.*, assignors to Technomed Medical Systems, S.A.
16 April 2002 (Class 600/2); filed 22 June 2000

This therapy probe includes a therapy transducer movably mounted on a probe body and a partially deformable flexible casing positioned around the transducer with a guard ring that limits deformation of the flexible casing when the transducer moves with respect to the probe body. The transducer can be moved between a therapy position and a retracted position and



the guard ring provides an end-of-travel stop to hold the transducer in a therapy position. One opening is provided for injecting liquid into the flexible casing and another opening on the guard ring allows withdrawing the liquid and also removing bubbles from the flexible casing as soon as they form.—DRR

6,383,152

43.80.Sh APPARATUS FOR PRODUCING SHOCK WAVES FOR TECHNICAL, PREFERABLY MEDICAL APPLICATIONS

Werner Hartmann and Joerg Kieser, assignors to Siemens Aktiengesellschaft
7 May 2002 (Class 601/4); filed in Germany 24 January 1997

The apparatus produces intense pressure waves as shock waves for technical and preferably medical applications, especially those for lithotripsy and/or pain therapy. Acoustic waves of high energy density are produced through the use of pressure pulsations. The pressure waves are pro-

duced by heating a conductive electrolyte, wherein electrical energy is converted directly and largely without losses with the aid of an electric pulse for the purpose of heating up the electrolyte. The device includes two electrodes that enclose the electrolyte and are controlled by a power pulse generator for outputting sound waves into a sound propagation medium.—DRR

6,371,924

43.80.Vj ACOUSTIC WINDOW IDENTIFICATION

Scott Donaldson Stearns, assignor to MedAcoustics, Incorporated
16 April 2002 (Class 600/528); filed 4 November 1999

This device provides for noninvasive detection of the sounds of abnormal blood flow through an array of acoustic sensors. The process involves five steps to determine an acoustic window suitable for passive-acoustic coronary disease evaluation. The steps include (a) positioning a multi-channel acoustic sensor array (consisting preferably of about 9–45 sensors) over a patient's chest, (b) calculating a weighted value for each of the sensor channels in the multi-channel array, (c) determining the location of each sensor in the array, (d) identifying the sensor channels which meet predetermined test criteria (generally the signal-to-noise ratios), and (e) defining a parameter which substantially encompasses the channels identified in step (d), thereby defining an acoustic window suitable for acoustic listening diagnostic procedures.—DRR

6,387,051

43.80.Vj METHOD AND APPARATUS FOR NON-INVASIVELY DERIVING AND INDICATING OF DYNAMIC CHARACTERISTICS OF THE HUMAN AND ANIMAL INTRACRANIAL MEDIA

Arminas Ragauskas and Gediminas Daubaris, assignors to UAB Vittamed
14 May 2002 (Class 600/438); filed 15 September 2000

This is an ultrasonic method of testing intraparenchymal brain tissue. Broadband ultrasound is transmitted from a transducer positioned on one side of a patient's cranium to a receiving transducer located on the other side. The received signal is resolved into narrow-band components. Group delay, phase angle, and attenuation are analyzed to determine the characteristics of the intraparenchymal tissue within the cranium.—DRR

Comment on "On the complex conjugate roots of the Rayleigh equation: The leaky surface wave" [J. Acoust. Soc. Am. 110, 2867 (2001)] (L)

John G. Harris[a] and Jan D. Achenbach
Ctr. QEFP, Northwestern University, 2137 North Sheridan Road, Evanston, Illinois 60208-3020

(Received 26 February 2002; revised 12 July 2002; accepted 19 July 2002)

The authors comment upon the role played by those roots of the Rayleigh equation that are referred to as extraneous, as well as upon the work of the paper given in the title. [DOI: 10.1121/1.1506689]

PACS numbers: 43.20.Bi, 43.35.Pt [JJM]

I. BACKGROUND

When solving Lamb's problem, one of the first steps taken is to define the Riemann sheet upon which the initial calculation will be carried out. One very common choice of this sheet is that defined in Harris.[1] We call this the γ_L^+, γ_T^+ sheet. The $+$ indicates that $\Im(\gamma_I) \geqslant 0$, $\forall \beta$, where $\gamma_I = [(\omega/c_I)^2 - \beta^2]^{1/2}$, and β is the independent variable. The c_L and c_T are the longitudinal and transverse wavespeeds, respectively, and ω is the angular frequency. On this sheet it is known that the Rayleigh equation has only one root,[2] and it is real. However, the equation does have other roots; these correspond to poles lying on the other Riemann sheets, which we call γ_L^+, γ_T^- and γ_L^-, γ_T^+. The $-$ indicates that $\Im(\gamma_I) \leqslant 0$, $\forall \beta$. These poles can sometimes play a role in calculating the total disturbance excited in a homogeneous, isotropic halfspace by a source such as a line source applied to the surface. This is the subject of the paper by Schröder and Scott[3] and upon which we comment.

We agree with these authors that the additional poles that lie on the γ_L^+, γ_T^- sheet can play an important role in calculating the complete disturbance; however, we would give a different interpretation to the results of their calculations. Moreover, it should be noted that the problem can be solved correctly by considering only the Rayleigh pole on the γ_L^+, γ_T^+ sheet and the contributions from the branch cuts, so that, in this sense, the additional poles are extraneous. However, that does not mean that they have been ignored.

In addition to the references cited by Schröder and Scott, the following references deal with the contributions made by the poles on the γ_L^+, γ_T^- sheet: One of the earlier calculations of the wave or pulse associated with these additional poles is that by Gilbert and Laster.[4] They consider the influence that these poles have on a pulse at the surface of a traction-free halfspace, when the source is buried. They then go on to consider the related problem of how these poles affect a pulse at the interface of two elastic half-spaces, in welded contact, when excited by a source buried in one of them. Hudson[5] in his discussion of Rayleigh waves briefly discusses the influence of the additional poles, citing the work of Gilbert and Laster. A thorough discussion of most of the same issues as discussed by Schröder and Scott was given by Tsang,[6] though he limited his calculations to points on the surface.

II. SUMMARY

For the problem of a line source on the surface of an isotropic, homogeneous elastic halfspace, much in the interpretation of the solution depends upon how the complex plane is cut. The cut described by Harris is such that the inverse Fourier transforms will be convergent everywhere on the Riemann sheet γ_L^+, γ_T^+. Therefore, the original contours of integration can be wrapped around the branch cuts picking up the Rayleigh pole as this is done. The resulting expressions describe the overall disturbance as a sum of one discrete eigenmode and two integrals over the continuous eigenmodes, and these expressions converge uniformly in the two coordinates of the halfspace. Though, the branch-cut integrals are by no means easy to evaluate, in a very interesting paper, Maupin[7] has succeeded in writing them as integrals wherein γ_L^+ and γ_T^+ are the integration variables, rather than β.

Other arrangements of the branch cuts are correct and may be advantageous. Schröder and Scott, and Tsang make different choices from that just indicated. In each case the branch cuts are arranged so as to involve the additional poles on the γ_L^+, γ_T^- sheet. As Tsang correctly notes, this rearrangement of the branch cuts is needed if one wants to make an accurate asymptotic calculation of the branch-cut contributions to the overall disturbance in the halfspace.

This does not seem to be precisely the goal of Schröder and Scott, who are more concerned, in their analytic work, with the residue contributions themselves. We believe that it is more appropriate to consider the contributions from these additional poles as part of the overall branch-cut contributions. There are two interrelated reasons for this: (1) If one chooses to adopt the modal interpretation, indicated previously, then the influence of these additional poles would be found in the branch-cut integrals and could be considered as part of the structure of the continuous spectrum. (2) If one chooses to approximate asymptotically the complete solution, as Schröder and Scott do, to derive a ray description of the total disturbance, then, for a limited range of positions of the saddle-point, one must also consider contributions from

[a]Electronic mail: j-harris8@northwestern.edu

the branch cuts. To evaluate these contributions asymptotically one would arrange the branch cuts as Tsang does (and not as Schröder and Scott do); in doing so the additional poles would contribute to the calculation. Those appearing as residue contributions do not satisfy the radiation condition, but they are only one part of the overall branch-cut contribution. The contributions from the modified branch cuts, along with the residue contributions, are also considered in the work of Schröder and Scott, though not asymptotically, and hence the interrelation of the various contributions is not noted.

Tsang makes careful asymptotic estimates of the branch-cut contributions and, as he does so, he is careful to include the influence of the poles on the γ_L^+, γ_T^- sheet so as to give a uniform asymptotic approximation, irrespective of whether or not the poles contribute separate residue contributions. His calculations, thus, also strongly suggest that these contributions are really features of the lateral wave, rather than separate disturbances.

[1] J. G. Harris, *Linear Elastic Waves* (Cambridge U. P., New York, 2001), pp. 52–55 and 96–101.
[2] J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, Amsterdam, 1973), pp. 189–194.
[3] C. T. Schröder and W. R. Scott, "On the complex conjugate roots of the Rayleigh equation: The leaky surface wave," J. Acoust. Soc. Am. **110**, 2867–2877 (2001).
[4] F. Gilbert and S. J. Laster, "Excitation and propagation of pulses on an interface," Bull. Seismol. Soc. Am. **52**, 299–319 (1962).
[5] J. A. Hudson, *The Excitation and Propagation of Elastic Waves* (Cambridge U. P., New York, 1980), pp. 171–172.
[6] L. Tsang, "Time-harmonic solution of the elastic head wave problem incorporating the influence of Rayleigh poles," J. Acoust. Soc. Am. **63**, 1302–1308 (1978).
[7] V. Maupin, "The radiation modes of a vertically varying half-space: a new representation of the complete Green's function in terms of modes," Geophys. J. Int. **126**, 762–780 (1996).

Sound-wave propagation in a membrane–duct (L)

Sunghoon Choi[a] and Yang-Hann Kim
Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Science Town, Taejon 305-701, Korea

(Received 10 May 2001; accepted for publication 1 August 2002)

Theoretical analysis for sound waves in a flexible membrane mounted in an otherwise rigid duct is presented. It has been shown that supersonic fluid-dominant waves and subsonic membrane-dominant waves exist. The latter provides powerful mechanism of transmission loss through the membrane damping and destructive reflections at the edges of the membrane, while the cross-sectional area remains constant. The effect of mean flow and external fluid is also considered.
 [DOI: 10.1121/1.1509761]

PACS numbers: 43.20.Mv, 43.20.Tb, 43.50.Gf [DEC]

I. INTRODUCTION

For years several different passive methods have been used to reduce noise in ducts. Examples are wave reflection by impedance mismatch as in expansion-chamber-type mufflers and viscous dissipation through fibrous duct lining. Each method has unique advantages and disadvantages leading to performance trade-offs during implementation. The expansion-chamber-type mufflers have inherent drawbacks in that they occupy a relatively large space and are effective only for a low frequency range. This paper explores theoretically the possibility of employing a membrane–duct as a compact, broadband noise controller.

First proposed by Huang,[1,2] the membrane–duct system consists of a flexible membrane of finite length mounted in an otherwise rigid duct (see Fig. 1). When a plane wave is traveling in the flexible portion of a duct, the cross-section distensibility renders a wave speed far less than the isentropic sound speed.[3] This mismatch in wave speed leads to reflection and scattering of waves at the edges of the membrane where the cross-sectional area remains constant. The interaction of these waves creates passbands or stop bands in transmission depending on wave patterns. The damping of the membrane further reduces the transmission of waves. At higher frequencies above the first critical frequency of a duct, supersonic waves are found and the sound wave propagates through the duct without any significant dissipation.[4]

In what follows, the frequency characteristics of waves in a membrane–duct of infinite length are first calculated. The results show that both subsonic membrane-dominant waves and supersonic fluid-dominant waves exist. The mean flow in a duct is considered and its effect is found to be negligible as far as the noise reduction is concerned. Also considered is the external air, which eliminates fluid-dominant waves by providing radiation damping. For a practical application, an appropriate use of absorption materials or a shielding screen is necessary to prevent the sound radiation into the external air, the so-called breakout noise.[5,6] In this study, however, the effect of breakout noise has not been considered in calculating the transmission loss of a finite-length membrane–duct. Instead, emphasis has been on the interactions of sound waves inside the duct.

[a]Electronic mail: sch@kaist.ac.kr

II. DISPERSION CHARACTERISTICS

A. A simple analytical method

To investigate the frequency characteristics of waves in a membrane–duct system, we consider a simple two-dimensional duct with its length in the x direction and let the infinitely long membrane be located at $y=\pm h$. Because of the symmetry in the geometrical configuration, the solution can be separated into a symmetric and an antisymmetric one about the $y=0$ plane. The symmetric solution can be obtained by assuming a rigid boundary condition at the $y=0$ plane, while the asymmetric counterpart can be obtained by using a pressure release boundary condition. We will investigate the symmetric solution first. Denoting the wave number in the x direction to be k_x, the harmonic pressure with angular frequency ω can be written as $p(x,y,t)=\tilde{p}(y)\exp[j(\omega t-k_x x)]$, and the corresponding two-dimensional homogeneous wave equation is given by

$$\frac{\partial^2 \tilde{p}(y)}{\partial y^2}+(k^2-k_x^2)\tilde{p}(y)=0, \tag{1}$$

where $k=\omega/c$ is the wave number and c is the isentropic sound speed in a free space. The general solution for $\tilde{p}(y)$ can be written as

$$\tilde{p}(y)=\tilde{A}\ \exp(-ik_y y)+\tilde{B}\ \exp(ik_y y), \tag{2}$$

where $k_y=(k^2-k_x^2)^{1/2}$. The equation for the membrane displacement with harmonic time dependence $\eta(x,t)=\tilde{\eta}e^{i\omega t}$ is given by

$$(Tk_x^2+iD\omega-\rho_m\omega^2)\tilde{\eta}=|\tilde{p}(y)|_{y=h}, \tag{3}$$

where T is the tension applied to the membrane, D is the damping coefficient, and ρ_m is the mass density per unit area. The symmetric boundary conditions are

$$[\partial\tilde{p}/\partial y]_{y=0}=0, \quad [\partial\tilde{p}/\partial y]_{y=h}=\rho\omega^2\tilde{\eta}, \tag{4}$$

where ρ is the mass density of the fluid. Applying these boundary conditions, we obtain the following solution:

$$\tilde{p}(y)=-\omega^2\rho\tilde{\eta}\ \cos(k_y y)/k_y\ \sin(k_y h). \tag{5}$$

Now define the fluid impedance as $Z_f=-\tilde{p}(h)/i\omega\tilde{\eta}$. In this definition the minus sign is used because the positive mem-

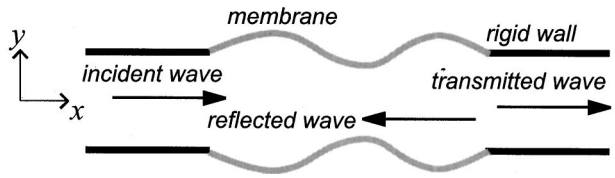


FIG. 1. Illustration of a membrane–duct.

brane velocity induces the negative fluid pressure. Then, we obtain

$$Z_f = \begin{cases} -i\rho\omega \cot(k_y h)/k_y, & \text{for } k > k_x, \\ i\rho\omega \coth(|k_y|h)/|k_y|, & \text{for } k < k_x. \end{cases} \tag{6}$$

Using Eqs. (3) and (6), we obtain the following dimensionless characteristic equation:

$$(\rho_m/\rho h)(c_m/c)^2(k_x h)^2 + i(D/\rho c)(kh) - (\rho_m/\rho h)(kh)^2 = -i(Z_f/\rho c)(kh). \tag{7}$$

Here, $c_m = \sqrt{T/\rho_m}$ is the membrane wave speed *in vacuo*. We seek to calculate the dimensionless wave number $k_x h$ for a given dimensionless frequency kh numerically.

As an example, dispersion relations for a specific case of $T = 50$ N/m, $D = 0$, $h = 0.1$ m, and $\rho_m = 0.1$ kg/m^2 are calculated. This value of ρ_m approximates rubber-like materials of 1-mm thickness. Properties of the air ($\rho = 1.21$ kg/m^3, $c = 340$ m/s are used for the fluid, and the dimensionless parameters are $\rho_m/\rho h = 0.826$ and $c/c_m = 15.2$, respectively. Subsonic and supersonic branches are shown separately in Fig. 2. Also shown are the membrane wave mode *in vacuo* in Fig. 2(a) and rigid duct modes in Fig. 2(b). In most practical cases the membrane wave speed *in vacuo* is far less than the sound speed. Thus, the acoustic impedance given in Eq. (6) implies that the fluid loading on the membrane acts like an added mass, and hence the wave speed decreases as shown in Fig. 2(a). Plotting the pressure profile of this subsonic mode along the cross section, we can observe that the pressure amplitude decays exponentially away from the membrane wall. It means most of its energy is concentrated near the membrane, and hence we will call it a membrane-dominant mode. On the other hand the supersonic branches, so called fluid-dominant modes, are understood as rigid duct modes modified by the vibration of the flexible membrane. The wave speed increases as shown in Fig. 2(b). There are certain critical frequencies below which these fluid-dominant modes become evanescent. The critical frequencies $k_c h$ can be found from Eq. (7) by assuming $k_x \to 0$, that is

$$\cot(k_c h) = (k_c h)(\rho_m/\rho h). \tag{8}$$

The dimensionless critical frequency is a function of the mass ratio $\rho_m/\rho h$ only. Shown in Fig. 3 is a plot of the first critical frequency. The critical frequency decreases as $\rho_m/\rho h$ increases.

For an antisymmetric problem the solution in Eq. (5) becomes

$$\tilde{p}(y) = \omega^2 \rho \tilde{\eta} \sin(k_y y)/k_y \cos(k_y h). \tag{9}$$

The difference between the symmetric and the antisymmetric membrane-dominant mode is found to be negligible. This is

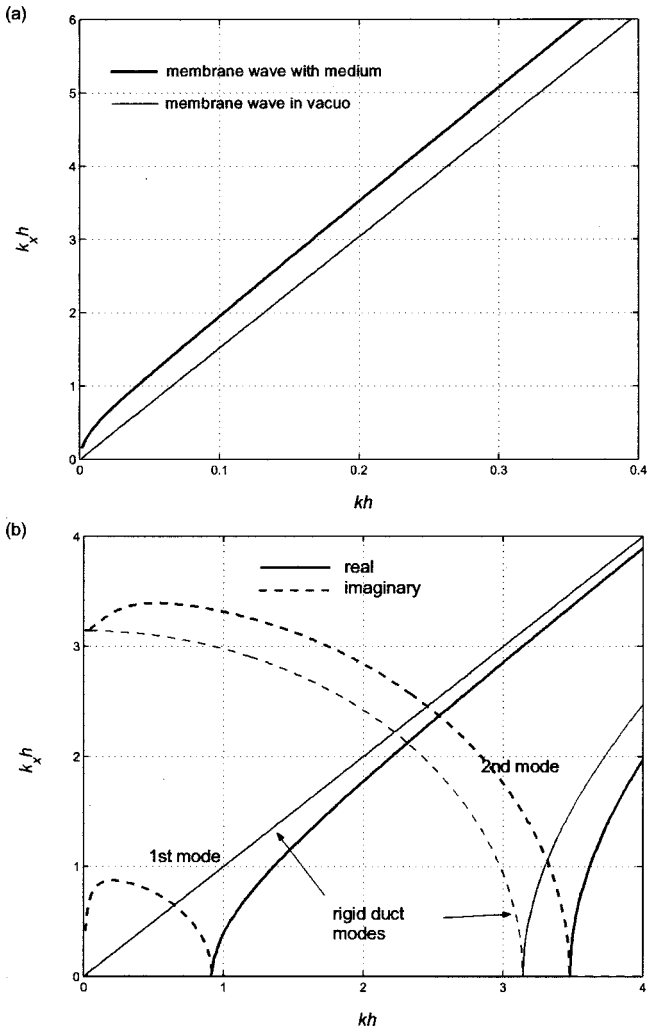


FIG. 2. Dispersion relations for the membrane in contact with fluid with $c/c_m = 15.2$, $\rho_m/\rho h = 0.826$, and $D = 0$: (a) Subsonic branch; the thick line represents the membrane-dominant mode and the thin curve is for the membrane wave *in vacuo*. (b) Supersonic branches; thick lines are for the fluid-dominant waves and the thin lines are for the rigid duct modes. Negative parts of imaginary wave numbers are shown with dashed lines.

because the boundary condition at $x = 0$ has little effect on its exponentially decaying pressure profile. The antisymmetric fluid-dominant modes are similar to those found in a rigid duct.

There will be energy dissipation through the membrane when it has damping. The average input power on the membrane by the fluid waves per unit length is given by

$$P = \langle p\, \dot{\eta} \rangle_t|_{y=h} = \tfrac{1}{2}\mathrm{Re}[\tilde{p}(h)(i\omega\tilde{\eta})^*], \tag{10}$$

where $\langle\ \rangle_t$ represents an average over time, and * represents the complex conjugate. Substituting Eq. (3) into Eq. (10) gives

$$P = \tfrac{1}{2}|\tilde{\eta}|^2 \mathrm{Re}[-i\omega T k_x^2 + D\omega^2]. \tag{11}$$

The first term due to the membrane restoring force becomes zero when k_x is pure real or pure imaginary. Then, the damping in a membrane will be the only cause of energy loss. A sound wave entering into a membrane duct transfers its energy through the membrane-dominant wave and the fluid-dominant wave. Below the first critical frequency of the

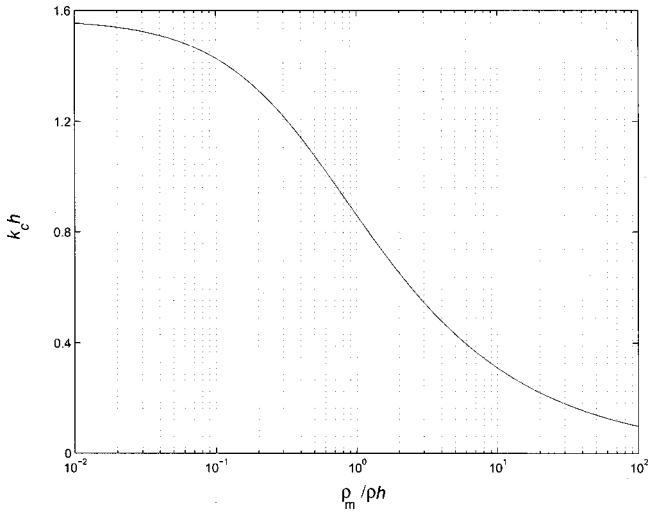


FIG. 3. Plot of the first critical frequency below which the propagating fluid-dominant wave does not exist.

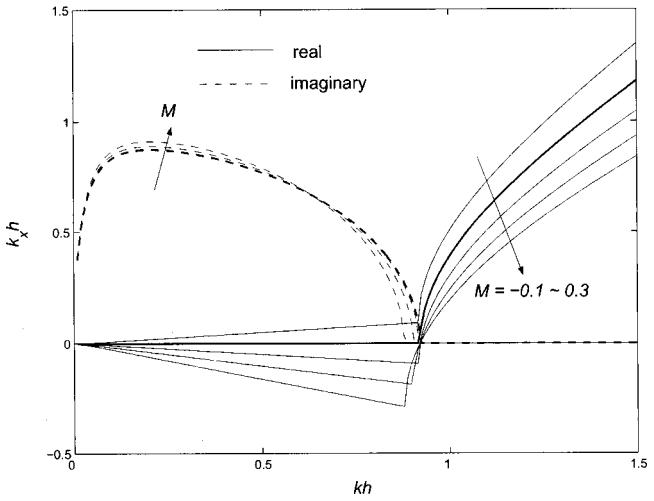


FIG. 4. Effect of mean flow speed on the dispersion of the fluid-dominant mode. ($c/c_m = 15.2$, $\rho_m/\rho h = 0.826$, $D = 0$.)

fluid-dominant wave, the energy is carried by the membrane wave only and may be dissipated through the membrane damping effectively.

B. The effect of mean flow speed

Here, we will consider a duct with mean flow. The wave equation now includes a convective term[7] and the solution is also written as in Eq. (5), in which

$$k_y = [(k - k_x M)^2 - k_x^2]^{1/2}, \tag{12}$$

and M is the Mach number of the mean flow. We assume that the mean flow speed is less than the membrane wave speed to avoid flexible instabilities in the flow.[1] It has been found that the membrane-dominant wave speed decreases slightly as the mean flow speed increases, though the effect is almost negligible. Shown in Fig. 4 are fluid-dominant modes for $-0.1 \leqslant M \leqslant 0.3$. The figure shows that the critical frequency is unchanged and phase speed increases as the mean flow speed increases, which is quite straightforward.

C. The effect of an external fluid

Assume an external pressure field as $p_+(x,t) = \tilde{p}_+(y)\exp[i(\omega t - k_x x)]$; then, $\tilde{p}_+(y)$ satisfies the Eq. (1) and its solution should have $e^{-ik_y y}$ dependence in which $k_y = \pm(k^2 - k_x^2)^{1/2}$. The appropriate sign of the square root is determined by the physics of the model. In addition to the boundary condition at the membrane–fluid interface, the solution must satisfy the far-field radiation condition such that no wave is coming from the infinity and $\tilde{p}_+(y)$ becomes zero as y approaches infinity. Therefore, k_y must be chosen so that $\mathrm{Re}(k_y) \geqslant 0$ and $\mathrm{Im}(k_y) \leqslant 0$. The wave number k_x representing a wave propagating in the positive x direction also satisfies the same relations as k_y does. From these conditions and $k^2 = k_x^2 + k_y^2$, we may easily prove that both k_x and k_y are either pure real or pure imaginary. Thus, k_y is rewritten as

$$k_y = \begin{cases} \sqrt{k^2 - k_x^2}, & \text{for } k > k_x, \\ -i\sqrt{k_x^2 - k^2}, & \text{for } k < k_x. \end{cases} \tag{13}$$

The right-hand side of Eq. (3) becomes $\tilde{p}(h) - \tilde{p}_+(h)$ and the acoustic impedance is written as follows:

$$Z_f = \begin{cases} -i\rho\omega[\cot(k_y h)/k_y + i], & \text{for } k > k_x, \\ i\rho\omega[\coth(|k_y|h)/|k_y| + 1], & \text{for } k < k_x. \end{cases} \tag{14}$$

Since k_x is either pure real or pure imaginary, we can show from Eq. (7) that only subsonic waves exist for a lossless membrane ($D = 0$). For $k > k_x$ the fluid loading becomes resistive and fluid energy is lost due to the radiating waves, and hence freely propagating waves do not exist anymore. For $k < k_x$ the added mass effect of fluid loading increases and further reduces the membrane-dominant wave speed.

III. SOUND WAVE TRANSMISSION

In this section we will consider a flexible membrane occupying $-L/2 \leqslant x \leqslant L/2$ in an otherwise rigid duct. Assume an incident harmonic plane wave $p_i(x,t) = \tilde{p}_i \exp[i(\omega t - k_x x)]$. The incident wave creates vibration of the membrane, which results in additional pressure in the duct through the local radiation impedance. The membrane equation of motion is given by

$$T\frac{\partial^2 \eta}{\partial x^2} - D\frac{\partial \eta}{\partial t} + p_i - Z\frac{\partial \eta}{\partial t} = \rho_m \frac{\partial^2 \eta}{\partial t^2}, \tag{15}$$

where Z is the local radiation impedance. Assuming harmonic time dependence, we can expand the velocity of the membrane, $\partial\eta/\partial t = \tilde{v}(x)e^{i\omega t}$, using the Fourier sine series

$$\begin{aligned} \tilde{v}(x) &= \sum_{n=0}^{\infty} v_n \sin[n\pi(x/L + 1/2)]. \\ v_n &= \frac{2}{L}\int_{-L/2}^{L/2} \tilde{v}(x)\sin[n\pi(x/L + 1/2)]dx. \end{aligned} \tag{16}$$

The sine transform of the membrane equation yields a following set of linear equations:

$$[\mathcal{Z}_{mn}][v_n] = [q_m], \tag{17}$$

where $\mathcal{Z}_{mn}$ and q_m are defined as

$$\mathcal{Z}_{mn}=[i\rho_m\omega+D+(T/i\omega)(n\pi/L)^2]\delta_{mn}+Z_{mn}, \tag{18}$$

$$q_m=2m\pi e^{ikL/2}\overline{p_i}\left[\frac{e^{i(-kL+m\pi)}-1}{(kL)^2-(m\pi)^2}\right], \tag{19}$$

respectively. In Eq. (18) Z_{mn} is the radiation impedance of the mth mode induced by the vibration of the nth mode, which will be given in the following. In Eq. (19) the value in the square bracket becomes $1/2ikL$ when $kL\rightarrow m\pi$.

The pressure perturbation induced by the wall vibration was given by Doak[8] and is rewritten for the two-dimensional configuration as follows:

$$\begin{aligned}\tilde{p}(x,y)=&\frac{\rho}{2h}\sum_{j=0}^{\infty}c_j\psi_j(y)\psi_j(h)\int_{-L/2}^{L/2}\tilde{v}(x')\\&\times[H(x-x')e^{-i\omega(x-x')/c_j}\\&+H(x'-x)e^{i\omega(x-x')/c_j}]dx',\end{aligned} \tag{20}$$

where H is the Heaviside function, ψ_j is the channel modal function defined as

$$\psi_j(y)=\sqrt{\epsilon_j}\cos\,\kappa_j y,\quad \kappa_j=j\pi/h,\quad \epsilon_j=\begin{cases}1, & j=0,\\ 2, & j\geqslant 1,\end{cases}$$

and c_j is the complex modal wave speed defined as

$$c_j=\begin{cases}\omega/\sqrt{\kappa^2-\kappa_j^2}, & \text{for } k>\kappa_j,\\ i\omega/\sqrt{\kappa_j^2-k^2}, & \text{for } k<\kappa_j.\end{cases} \tag{21}$$

The pressure field induced by the nth mode of membrane vibration of unit amplitude can be obtained by replacing $\tilde{v}(x')$ by $\sin[n\pi(x'/L+1/2)]$ in Eq. (20). Then, the radiation impedance Z_{mn} is the mth Fourier component of the resultant pressure field, which is found by

$$Z_{mn}=\rho\sum_{j=0}^{\infty}\epsilon_j c_j\mathcal{I}_j(m,n), \tag{22}$$

where

$$\begin{aligned}\mathcal{I}_j(m,n)=&\frac{mn\pi^2(e^{im\pi}-e^{-ik_jL})(e^{im\pi}+e^{in\pi})}{[(k_jL)^2-(m\pi)^2][(k_jL)^2-(n\pi)^2]}\\&-\frac{ik_jL\,\delta_{mn}}{(k_jL)^2-(m\pi)^2}.\end{aligned} \tag{23}$$

It can be easily shown that $\mathcal{I}_j(m,n)$ is always finite.

The total sound waves transmitted to the downstream is the sum of the incident wave and the wave radiated by the vibrating membrane. In Eq. (20) the first term in the integrand represents waves traveling to the positive-x direction and only those with real modal wave speed will be radiated to the downstream. We will denote the maximum mode number with real c_j as J; by using Eqs. (16) and (20), the radiated pressure is found by

$$\tilde{p}_r(x,y)=\rho c\frac{L}{2h}\sum_{j=0}^{J}(-1)^j\epsilon_j c_j\Gamma_j e^{-ik_j(x+L/2)}\cos(\kappa_j y), \tag{24}$$

where $k_j=\omega/c_j$ and

$$\Gamma_j=\frac{1}{c}\sum_{n=1}^{\infty}nv_n\pi\left[\frac{e^{i(k_jL-n\pi)}-1}{(k_jL)^2-(n\pi)^2}\right]. \tag{25}$$

The fluid particle velocity in the positive-x direction is obtained by using Eq. (24). Then, the power transmitted to the downstream can be calculated by integrating the time-averaged product of the fluid pressure and particle velocity over the channel cross section. Using the orthonormal property of the channel modal functions, the ratio of the transmitted power Π_t to the input power Π_i is given by

$$\frac{\Pi_t}{\Pi_i}=\left|1+\frac{L}{2h}\Gamma_0 e^{-ik_0L/2}\right|^2+\frac{1}{2}\left(\frac{L}{h}\right)^2\sum_{j=1}^{J}\frac{c_j}{c}|\Gamma_j|^2. \tag{26}$$

The transmission loss is

$$\mathrm{TL}=-10\log_{10}(\Pi_l/\Pi_i). \tag{27}$$

When the wave number corresponding to the critical frequency is smaller than that of the first channel mode, the only contribution to the transmission comes from the plane wave mode; hence, only the first term in the right-hand side of Eq. (26) will contribute to the transmission loss.

IV. CONCLUSIONS

Theoretical analysis has been carried out to determine the frequency characteristics and sound reduction performance of a membrane–duct system. At low frequencies, the fluid loading acts like an inertia force and a membrane-dominant wave, slower than the membrane wave *in vacuo*, is found. At higher frequencies above the first critical frequency of a duct, fluid-dominant waves are found and the sound wave propagates through the duct without any significant dissipation. The mean flow in a duct of low Mach number has been found to have little effect on the frequency characteristics of the membrane-dominant wave. The external fluid provides the propagating waves with radiation damping and the fluid-dominant modes become evanescent.

ACKNOWLEDGMENTS

The first author was supported by BK21 fund by the Department of Education, Korean Government.

[1] L. Huang, "A theoretical study of duct noise control by flexible panels," J. Acoust. Soc. Am. **106**, 1801–1809 (1999).
[2] L. Huang, "Experimental study of sound control in a flexible panel," J. Acoust. Soc. Am. **108**, 624–631 (2000).
[3] J. Lighthill, *Waves in Fluids* (Cambridge University Press, Cambridge, 1978).
[4] S. Park and Y.-H. Kim, "Low frequency noise reduction by acoustic coupling with membrane" (in Korean), J. KSNVE **7**, 579–588 (1997).
[5] M. L. Munjal and P. T. Thawani, "Acoustic performance of hoses—A parametric study," Noise Control Eng. J. **44**, 274–280 (1996).
[6] A. Cummings, "Sound transmission through duct walls," J. Sound Vib. **239**, 731–765 (2001).
[7] S. H. Ko, "Sound wave propagation in a two-dimensional flexible duct in the presence of an inviscid flow," J. Sound Vib. **175**, 279–287 (1994).
[8] P. E. Doak, "Excitation, transmission and radiation of sound from source distributions in hard-walled ducts of finite length. I. The effects of duct cross-section geometry and source distribution space-time pattern. See Eq. (123)," J. Sound Vib. **31**, 1–72 (1973).

Object localization from acoustic emissions produced by other sonars (L)

Roman Kuc[a)]
Intelligent Sensors Laboratory, Department of Electrical Engineering, Yale University, New Haven, Connecticut 06520-8284

(Received 15 February 2002; revised 19 July 2002; accepted 1 August 2002)

When multiple active sonars, such as dolphins, bats, or mobile robots, interrogate the environment simultaneously, one view is that each sonar considers the signals produced by the others to be interference, while another view is that these signals contain useful information. A listening sonar can typically detect the time delay between another sonar's interrogation emission and its corresponding echo from an object and possibly also the emission and echo directions. This letter first shows that, even if the directions of these two signals and relative delays are known, determining the object location is an ill-posed problem, supporting the interference view. It then shows the time delay alone provides useful information, indicating that no echo-producing object can be closer than a perceived minimum range. Echoes from specular surfaces provide similar information but also require self echolocation to refine the range estimate. Experimental acoustic waveforms illustrate the approach. The results indicate that traveling in fixed formation and emitting identifiable echolocation sounds faciliate cooperative echolocation by biological sonars. [DOI: 10.1121/1.1508792]

PACS numbers: 43.80.Ka, 43.64.Bt, 43.10.Vx [WA]

I. INTRODUCTION

Consider a pod of dolphins swimming among obstacles, each transmitting echolocation sounds and detecting their own echoes and the emissions produced by others and their corresponding echoes.[1] Are these latter signals disturbances or crosstalk to be eliminated[2,3] or do they contain useful information?[4] Figure 1 shows a listening dolphin L that detects the emission from sonar source dolphin S and the echo from object O. The question is: What does L know about the location of O from the detected signals? The locations of S, O, and L define a plane, allowing the problem to be solved using a two-dimensional argument. Dolphins use the time-of-flight T between their own emissions and the echoes to determine object range, so it is feasible that they can also measure delays between other sounds. Straight-line propagation paths indicate that $T_{SOL}=T_{SO}+T_{OL}\geqslant T_{SL}$. The detected echo is delayed relative to the direct path emission by $\Delta T=T_{SOL}-T_{SL}$. Binaural hearing and pinna shape provide cues for estimating wavefront arrival angles θ_{SL} and θ_{OL}.[5–8] The figure shows these angles with respect to an arbitrary travel direction by L.

II. ANALYSIS

We first show that determining O's location knowing only ΔT, θ_{SL}, and θ_{OL} is an ill-posed problem in that it has an infinite number of valid solutions. To show this, we consider T_{SL} also to be known, which together with θ_{SL} locates S relative to L. Then, knowledge of ΔT forms an ellipse on which O lies,[9] and θ_{OL} determines O's location, as shown in Fig. 2. However, if the location of S is unknown, locating O from ΔT and θ_{OL} becomes problematic because multiple ellipses can exist. For example, if the source is located at S, the object is at O, but if the source is at S2, the object is at O2.

We now show that although O's exact location cannot be determined, L can still extract information about O useful for navigation and collision avoidance. We consider two classes of objects: The first assumes O is an isolated object, such as another dolphin or a fish. A valid reflector in this class is the section of a rough surface, such as coral, that produces the first detected echo. The second class assumes O is the echo-producing section of a large specular surface, such as a sandy sea floor or the sea surface. In the first class, knowing only ΔT, L can still determine that O is no closer than minimum range

$$r_{\min}=c\Delta T/2, \tag{1}$$

where c is the sound speed. This is most easily seen if S and O lie on opposite sides of L. Then the emission by S passes by L and can be considered to be an emission by L. L detects the corresponding echo from O ΔT later, determining $r_{\min}$. This $r_{\min}$ value forms the radius of a circle centered at L that contains no objects. This circle is the ellipse that results when S and L occupy the same point, forming a worst-case assumption of S's position. This circle lies within any ellipse formed by ΔT, independent of the S location. Knowing θ_{OL} indicates that O lies no closer than $r_{\min}$ in that direction.

The second object class contains environments that have specularly reflecting surfaces, corresponding in a visual world to a hall of mirrors, which complicate the problem of obstacle localization, as is well known in robotics.[10] Figure 3 shows that the section on a specular surface at minimum range r from L is not typically the section, designated by O, producing the echo to S's emission. This may cause a collision if L is traveling obliquely toward the surface. Additional

[a)]Electronic mail: roman.kuc@yale.edu

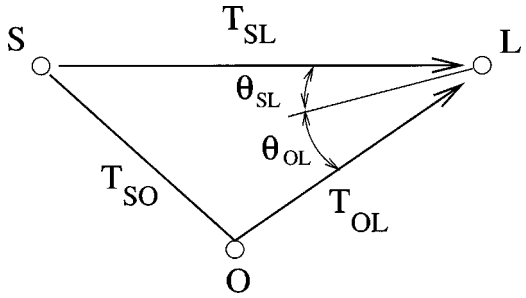


FIG. 1. Sonar L hears echolocation sound from sonar S and echo from object O. L detects delay $\Delta T = T_{SO} + T_{OL} - T_{SL}$ and arrival angles θ_{SL} and θ_{OL}.

information is needed to avoid collision. L's own echolocation system would localize the section at range r on a specular reflector. Hence, in specular environments the information produced by other sonars may not be sufficient, but only complements L's own echolocation capabilities.

In both object classes, L can refine O's location during movement by processing multiple emissions by S. Biosonars typically emit a sequence of pulses.[1,11] If O is the closest object to L, ΔT would decrease as L approaches O. If L travels at constant speed, the decrease in ΔT infers object location using acoustic flow.[12] For example, if L is heading directly toward stationary O, the rate of change of $r_{\min}$ with time would equal L's travel speed.

III. EXPERIMENTAL VERIFICATION

An experiment was performed to mimic a pair of dolphins. Figure 4 shows the experimental configuration. The experiment was performed in an empty squash court, a 9.75-m-long by 5.4-m-wide by 5.64-m-high room with smooth hard surfaces. The listening sonar L is a cardiod microphone (AT9400) located 0.5 m from a side wall at a height of 1.2 m and at least 4 m away from other walls. The microphone was oriented so that its most sensitive direction pointed toward the wall. The sound source S mimicked a dolphin echolocation *click*[1] produced by striking a dampened aluminum rod with a metal ringer. Actual dolphin clicks are similar in waveform but occupy a frequency range around 50 kHz. S

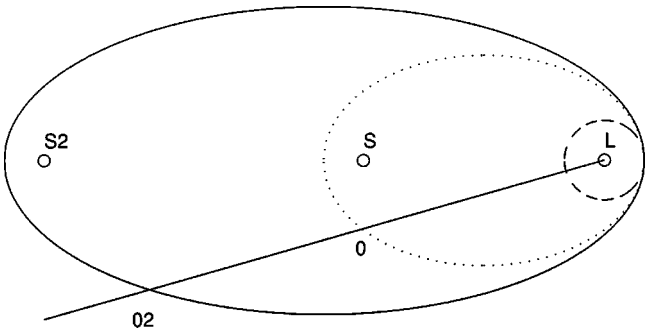


FIG. 2. If the location of S relative to L and all travel times are known, ΔT determines the ellipse on which O lies. Even if L knows arrival angles, localizing O is an ill-posed problem if range to S is unknown. The object lies at O if the source is at S (dotted ellipse), but lies at O2 if the source is S2 (solid ellipse). Knowing only ΔT, L can still determine that O is no closer than $r_{\min} = c\Delta T/2$, where c is the sound speed (dashed circle, which lies within all ellipses formed by ΔT).

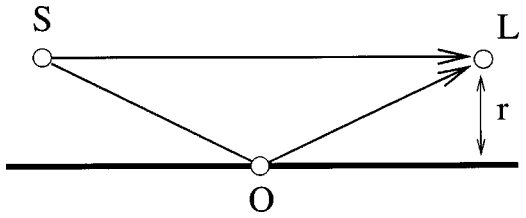


FIG. 3. The section of a specular surface at minimum range, r, to L is not typically the section O producing the echo to the S emission. Echolocation emissions by S would locate O, while self-echolocation emissions by L would locate the minimum range section.

was located 0.5 m behind L at the same height. The waveform was recorded using the microphone input on a laptop computer to generate a wav file with sampling rate 22.05 ksps and 8-bit resolution. The waveform time origin corresponds to the sound generation time.

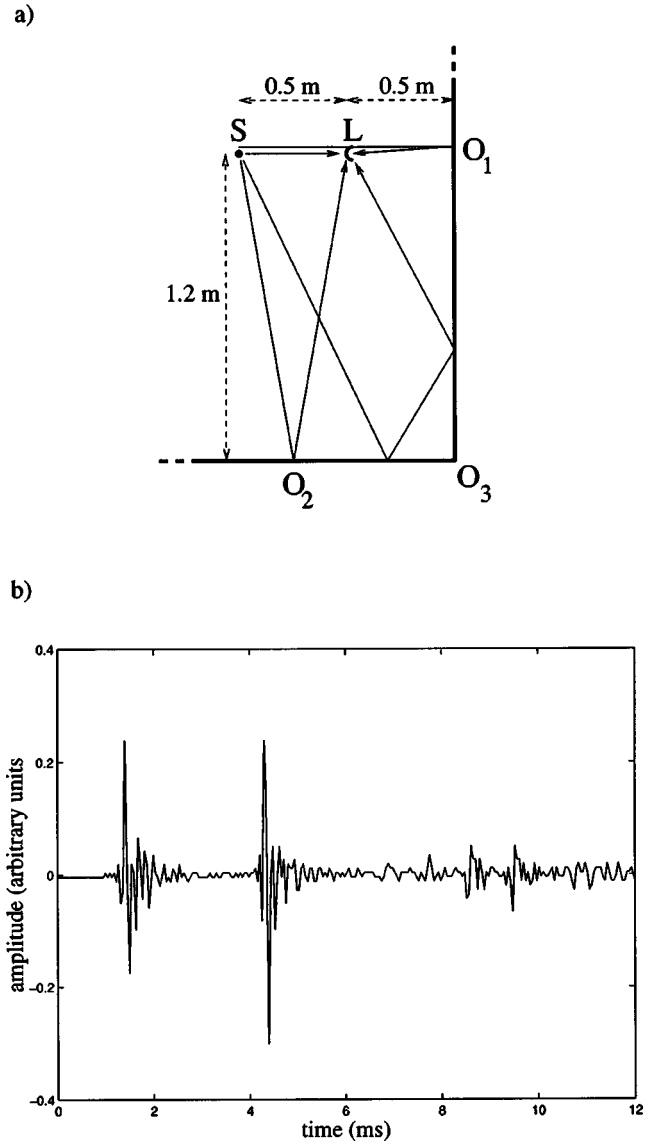


FIG. 4. Experimental results. (a) Experimental configuration showing location of acoustic source S and microphone L within an empty squash court. (b) Microphone waveform showing acoustic pulse emitted by the source at $t = 0$ (unknown to microphone), detected at 1.5 ms, followed by echos from wall in front of the microphone at 4.4 ms, from floor at 8.6 ms, and from corner at 9.6 ms.

The observed waveform clearly shows the direct path emission detected at 1.5 ms, the wall echo detected at 4.4 ms, and a pair of smaller echoes at 8.6 and 9.6 ms. Other small echoes were produced by the microphone stand. L's wall-directed sensitivity pattern attenuates the direct signal from S, enhances the wall echo, yet still detects echoes from the floor and corner. Applying the observed echo delays to Eq. (1) with the speed of sound in air $c=343$ m/s, the first echo with $\Delta T_1=2.9$ ms indicates the closest obstacle O_1 is at range 0.5 m. Echolocation by L would also detect O_1 at 0.5 m. The echo from the floor section O_2 produces $\Delta T_2=7.1$ ms with resulting perceived range 1.22 m, while self-echolocation would indicate 1.2 m. The echo from the corner formed by the wall and floor O_3 produces $\Delta T_3=8.1$ ms or perceived range 1.39 m, while self-echolocation would indicate 1.3 m.

IV. DISCUSSION

These results indicate that sonars can extract useful information from echolocation signals of other sonars or from sounds produced by other external acoustic sources. To exploit this information echolocating sonars could act cooperatively, working toward the common goal of navigating without colliding with obstacles by providing information useful to the group. With multiple sources, $S_1, S_2, S_3, \ldots$, L would need to match each emission with its corresponding echoes, which indicates that each sonar should emit uniquely recognizable signals. It has been reported that bats transmit individualized foraging calls,[13] which would facilitate this association process. The above analysis also indicates that L could localize O if it knew S's position. This could be accomplished if dolphins swam in a fixed formation. Unique emissions and known locations would allow the set of echolocation sounds to provide accurate obstacle localization and possibly object characterization. Animal experiments are needed to verify these speculations and to determine the interactions of other sonars with self-echolocation.

V. SUMMARY

This letter demonstrates that useful information can be extracted from the emissions produced by other sonars and their echoes. Results from other researchers suggest that biosonars exhibit the flexibility to exploit this approach when multiple sonars operate simultaneously.

[1] W. W. L. Au, *The Sonar of Dolphins* (Springer-Verlag, New York, 1993).
[2] W. M. Masters and K. A. S. Raver, "The degradation of distance discrimination in bit brown bats (Eptesicus fuscus) caused by different interference signals," J. Comp. Physiol., A **179**(5), 703–713 (1996).
[3] V. Masek, M. Kajitani, A. Ming, and L. V. Vlacic, "Rapid obstacle sensing using mobile robot sonar," Mechatronics **10**(1-2), 191–213 (2000).
[4] M. J. Xitco and H. L. Roitblat, "Object recognition through eavesdropping: Passive echolocation in bottlenose dolphins," Anim. Learn Behav. **24**(4), 355–365 (1996).
[5] E. A. Lopez-Poveda and R. Meddis, "A physical model of sound diffraction and reflections in the human concha," J. Acoust. Soc. Am. **100**, 3248–3259 (1996).
[6] N. H. Fletcher, *Acoustic Systems in Biology* (Oxford U.P., New York, 1992).
[7] J. Mogdans, J. Ostwald, and H.-U. Schnitzler, "The role of pinna movement for the localization of vertical and horizontal wire obstacles in the greater horseshoe bat, *rhinonlopus ferrumequinum*," J. Acoust. Soc. Am. **84**, 1676–1679 (1988).
[8] B. D. Lawrence and J. A. Simmons, "Echolocation in bats: The external ear and perception of the vertical positions of targets," Science **218**, 481–483 (1982).
[9] R. Kuc and B. Barshan, "Bat-like sonar for guiding mobile robots," IEEE Control Syst. Mag. **12**, 4–12 (1992).
[10] R. Kuc, "A spacial sampling criterion for sonar obstacle detection," IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 686–690 (1990).
[11] J. A. Simmons, M. B. Fenton, and M. J. O'Farrell, "Echolocation and pursuit of prey by bats," Science **203**, 16–21 (1979).
[12] D. N. Lee, F. R. van der Weel, T. Hitchcock, E. Matejowsky, and J. D. Pettigrew, "Common principle of guidance by echolocation and vision," J. Comp. Physiol., A **171**, 563–571 (1992).
[13] K. L. Murray, E. R. Britzke, and L. W. Robbins, "Variation in search-phase call of bats," J. Mammal. **82**(3), 728–737 (2001).

On the existence of flexural edge waves on thin orthotropic plates

Ian Thompson and I. David Abrahams[a)]
Department of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

Andrew N. Norris
Department of Mechanical and Aerospace Engineering, Rutgers University, Piscataway, New Jersey 08854-8058

(Received 11 January 2002; revised 12 July 2002; accepted 15 July 2002)

This paper is concerned with an investigation into the existence of waves propagating along a free edge of an orthotropic plate, where the edge is inclined at arbitrary angle to a principal direction of the material. After deriving the governing equation and edge conditions, an edge wave ansatz is substituted into this system to reduce it to a set of algebraic equations for the edge wave wave number and wave vector. These are solved numerically for several typical composite materials although analytic expressions can be obtained in the case of special values of the material parameters and inclination angle. It is found that a unique edge wave solution, which generally exhibits oscillation as well as decay away from the free edge, exists in all cases, and its wave speed is independent of its direction of propagation along the plate. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1506686]

PACS numbers: 43.20.Gp, 43.20.Jr, 43.40.Dx [JGH]

I. INTRODUCTION

This paper is concerned with the existence of edge waves propagating on the free edge of a thin elastic plate. Such flexural waves are localized about the edge, i.e., have exponential decay in the direction normal to the free edge, and therefore their energy vector is directed along the edge. These solutions show similarity with other localized waves, such as Rayleigh waves in elasticity, beach waves in water wave theory, etc. The existence of edge waves along the free edge of a homogeneous and isotropic semi-infinite thin plate, modeled using Kirchhoff theory, was first noted by Konenkov[1] in 1960. Published in a Russian journal, this result was little known in the West until it was independently rediscovered by Thurston and McKenna[2] and by Sinha[3] in articles published in 1974. Indeed, it was also ''predicted'' by De La Rue[4] in 1972, and surprisingly, continues to this day to be commonly overlooked by researchers (see, e.g., a recent paper by Kauffmann[5]). Konenkov and these other authors established that, for isotropic plates, precisely one edge wave solution exists for all values of the two free parameters, namely the bending stiffness and Poisson's ratio. The edge wave speed is found to be proportional to and slightly less than the speed of flexural (one-dimensional) waves on a plate of infinite extent.

Edge waves are not unique to isotropic Kirchhoff thin plates; Thurston, Boyd, and McKenna[6] and Krylov[7] have independently looked for guided elastic modes on the tip of a slender wedge, and work by Norris[8] in 1994 demonstrated the existence of edge waves in orthotropic Kirchhoff plates. More recently Norris, Krylov, and Abrahams[9] investigated the propagation of edge waves on thick plates, and did this by generalizing the model to incorporate Mindlin plate theory. Their results were consistent with that found by Konenkov, as were those of Lagasse[10] and Burridge and Sabina,[11,12] who employed finite element modeling to examine the existence of localized waves on the edge of a thick elastic structure employing full three-dimensional elasticity. Experiments in 1976 by Lagasse and Oliner[13] have demonstrated the existence of the edge wave with wave speed as predicted by theory, although the behavior of the wave speed at higher frequencies is better modeled using the improved theories, such as Mindlin plate theory.[9] Various authors have concerned themselves with the existence of edge waves along fluid-loaded structures: Krylov[14] has an approximate approach to such problems (for plates of wedge shaped cross section submerged in water) employing a ray theory analysis in the limit of high frequency. His results do not prove that such wave can propagate without loss, but there is some experimental support for their existence[15] even if they do radiate a small amount of energy in the direction perpendicular to the edge. Only recent work by Abrahams and Norris[16] has established rigorously that unattenuated edge waves can exist, within a limited frequency window, on thin elastic plates composed of a range of typical materials, from aluminum to Plexiglas™, submerged in air.

The presence or absence of edge waves on thin elastic structures can be of great importance in the field of nondestructive evaluation (NDE) of material components. Ultrasonic elastic waves are regularly employed in NDE to examine an engineering specimen, such as a rotor blade of an aerojet aircraft, and from measurement of the scattered field any imperfections or inclusions in the body can be found. This is an inverse problem, determining the shape, size and *nature* (e.g., inclusion, void, crack, change of phase, etc.) of defects from the overall measured far field. It is helpful, for

a)Electronic mail: i.d.abrahams@ma.man.ac.uk

the development of efficient inverse routines, to have knowledge of forward problems; that is, the scattering signature produced by a wide variety of common material inhomogeneities. For cracks on thin structures this scattering signature can be greatly altered by the existence or nonexistence of flexural edge waves, especially in the shadow zone where the specular field is absent and so the edge wave may be significantly more energetic than the diffracted wave shed from the crack edges. In the context of NDE the authors[17] are currently investigating the forward problem of the diffraction of flexural waves by a semi-infinite crack in a fiber reinforced thin plate, modeled using orthotropic Kirchhoff thin plate theory. This is an important problem to study because of the now widespread use of fiber reinforced and/or laminated materials in industrial applications. These offer extremely strong and light materials but inspection of critical components is still required because of the possibility of delamination, fiber debonding, or other modes of fracture.

For the above-mentioned diffraction problem[17] a crack can, in the general case, be inclined at arbitrary angle to the principal axes of orthotropy. In view of the relevance of edge waves to energy radiation in such models, a precursor study is needed to establish the existence or nonexistence of such waves along *inclined* edges. Norris[8] proved existence of a unique edge wave on an orthotropic plate, but this result is only given for the case in which the free edge is parallel to a principal direction of the material. Thus, in this article, the possibility of edge waves in anisotropic Kirchhoff thin plates when the free edge is inclined at general angle to a principal direction is investigated. In Sec. II, the governing equation and edge conditions governing such a system are derived, where the notation employed by Timoshenko and Woinowsky-Krieger[18] is loosely followed. An edge wave ansatz (of the form employed by Konenkov) is substituted into the system, in Sec. III, and from this a set of algebraic equations is obtained for the edge wave wave number and wave vector. These equations are a generalization of those found by Norris,[8] and in general do not permit analytic solutions. For several special values of the material parameters and inclination angle, explicit solutions are readily obtainable and these are discussed in Sec. IV. Finally, numerical solutions to the general system of equations are given in Sec. V, along with some concluding remarks. In particular, it will be shown that a single edge wave solution is found to exist in all cases considered; however, in general, it possesses oscillation as well as decay in the direction orthogonal to the free edge, unlike solutions in previous literature, all of which exhibit pure decay.

II. DERIVATION OF EQUATIONS

The classical equations of plate flexure for an orthotropic, homogeneous thin plate are studied in detail by Timoshenko,[18] and summarized by Norris.[8] Employing the latter author's notation, and taking the principal axes of orthotropy, or fiber direction(s), as x and y, the governing equation for flexural (transverse) displacements W is

$$D_x\frac{\partial^4 W}{\partial x^4}+2H\frac{\partial^4 W}{\partial x^2\partial y^2}+D_y\frac{\partial^4 W}{\partial y^4}+\rho h\frac{\partial^2 W}{\partial t^2}=0, \tag{1}$$

where ρ is the plate density, h is the plate thickness, t denotes time, and the bending stiffnesses in the x, y directions are D_x, D_y, respectively. Further $H=D_1+2D_{xy}$, where D_x, D_y, D_1, and D_{xy} can be written as

$$D_x=\frac{h^3}{12}\frac{E_1}{1-\nu_{12}\nu_{21}},\quad D_y=\frac{\nu_{21}}{\nu_{12}}D_x,$$

$$D_1=\nu_{21}D_x,\quad D_{xy}=\frac{h^3}{12}G_{12},$$

and

$$\nu_{21}E_1=\nu_{12}E_2.$$

Here, the suffixes 1 and 2 refer to the x and y directions, respectively, so E_1 is the Young modulus in the x direction, G_{12} is the shear modulus in the $x-y$ plane, and ν_{12} is the Poisson ratio for transverse strain in the y direction caused by stress in the x direction, with similar definitions for E_2 and ν_{21}. Note that the condition of positive definiteness of strain energy density means that the material parameters must satisfy the following conditions:

$$D_x>0,\quad D_y>0,\quad D_{xy}>0,\quad D_xD_y>D_1^2. \tag{2}$$

Note that the third inequality gives

$$H>D_1, \tag{3}$$

while the final inequality reveals that

$$\sqrt{D_xD_y}>D_1>-\sqrt{D_xD_y}, \tag{4}$$

where D_1 takes negative values when ν_{12} and ν_{21} are negative. It is noted that the bending moments M_x and M_y arising from distributions of in-plane normal stresses σ_x and σ_y, and the twisting moment M_{xy} and shear forces per unit length Q_x, Q_y arising from the shear stresses in the plate, can be written in terms of the transverse displacement, respectively, as

$$M_x=-\left(D_x\frac{\partial^2 W}{\partial x^2}+D_1\frac{\partial^2 W}{\partial y^2}\right), \tag{5}$$

$$M_y=-\left(D_y\frac{\partial^2 W}{\partial y^2}+D_1\frac{\partial^2 W}{\partial x^2}\right), \tag{6}$$

$$M_{xy}=-M_{yx}=2D_{xy}\frac{\partial^2 W}{\partial x\partial y}, \tag{7}$$

$$Q_x=-\frac{\partial}{\partial x}\left(D_x\frac{\partial^2 W}{\partial x^2}+H\frac{\partial^2 W}{\partial y^2}\right), \tag{8}$$

and

$$Q_y=-\frac{\partial}{\partial y}\left(D_y\frac{\partial^2 W}{\partial y^2}+H\frac{\partial^2 W}{\partial x^2}\right). \tag{9}$$

As mentioned in Sec. I, a semi-infinite plate is considered where the free-edge is inclined at arbitrary angle to the principal axes of orthotropy. Thus, as shown in Fig. 1, coordinates X, Y are chosen such that the plate is infinite in the X direction, semi-infinite in Y with a free edge along the line $Y=0$, and with the principal direction x inclined at an angle $\theta\in[0,\pi/2]$ to the X axis. Expressions are required for the

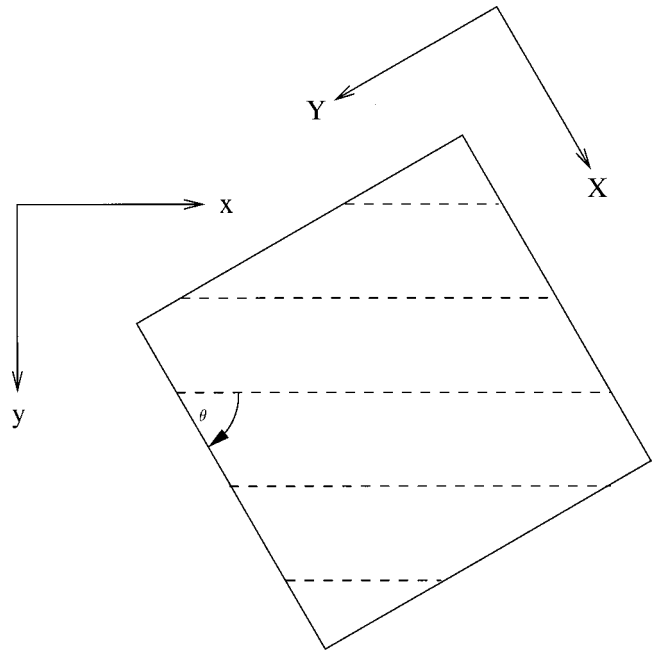


FIG. 1. A square section of the orthotropic plate. The dotted lines represent a principal direction of the material, which is parallel to the x axis, and the free edge lies along $Y=0$.

bending moment M_Y and the generalized Kirchhoff shear force

$$V_Y = Q_Y + \partial M_{YX}/\partial X, \tag{10}$$

where Q_Y is the shear force arising from the shear stress τ_{YZ} (i.e., acting on the edge at $Y=0$ but in the direction normal to the plane of the plate) which can be written as

$$Q_Y = \frac{\partial M_Y}{\partial Y} + \frac{\partial M_{YX}}{\partial X}. \tag{11}$$

The free-edge boundary conditions for a Kirchhoff thin plate stipulate that these quantities vanish along this edge (see Graff[19]). These conditions can be obtained in terms of the bending and twisting moments in the principal directions x,y by considering a small triangular section of the plate. Omitting details, it is easily shown that

$$M_Y = s^2 M_x + c^2 M_y + 2scM_{xy}, \tag{12}$$

$$M_{YX} = (s^2 - c^2)M_{xy} - sc(M_x - M_y), \tag{13}$$

and also

$$Q_Y = -sQ_x + cQ_y, \tag{14}$$

where here and henceforth the notation $c=\cos\theta$ and $s=\sin\theta$ has been introduced for brevity. Using the coordinate transformation for derivatives,

$$\frac{\partial}{\partial x} = c\frac{\partial}{\partial X} - s\frac{\partial}{\partial Y}, \quad \frac{\partial}{\partial y} = s\frac{\partial}{\partial X} + c\frac{\partial}{\partial Y},$$

with Eqs. (1), (5)–(9), (10), (11), and (12)–(14), the entire system is now rewritten in terms of derivatives parallel and perpendicular to the edge of the plate. Thus, the governing equation becomes

$$\left[D_x\left(s\frac{\partial}{\partial Y} - c\frac{\partial}{\partial X}\right)^4 + 2H\left(s\frac{\partial}{\partial Y} - c\frac{\partial}{\partial X}\right)^2\left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right)^2 + D_y\left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right)^4\right]W + \rho h\frac{\partial^2 W}{\partial t^2} = 0, \tag{15}$$

while the bending moment is given by

$$M_Y = -\left[(s^2 D_x + c^2 D_1)\left(c\frac{\partial}{\partial X} - s\frac{\partial}{\partial Y}\right)^2 + (s^2 D_1 + c^2 D_y) \times \left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right)^2 + 4scD_{xy}\left(s\frac{\partial}{\partial Y} - c\frac{\partial}{\partial X}\right) \times \left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right)\right]W, \tag{16}$$

and the Kirchhoff generalized shear force is

$$V_Y = \left\{s\left[-H\left(s\frac{\partial}{\partial Y} - c\frac{\partial}{\partial X}\right)\left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right)^2 - D_x\left(s\frac{\partial}{\partial Y} - c\frac{\partial}{\partial X}\right)^3\right] + c\left[-H\left(s\frac{\partial}{\partial Y} - c\frac{\partial}{\partial X}\right)^2\left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right) - D_y\left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right)^3\right] - \frac{\partial}{\partial X}\left[sc(D_1 - D_x) \times \left(s\frac{\partial}{\partial Y} - c\frac{\partial}{\partial X}\right)^2 + sc(D_y - D_1)\left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right)^2 + 2(s^2 - c^2)D_{xy}\left(s\frac{\partial}{\partial Y} - c\frac{\partial}{\partial X}\right)\left(c\frac{\partial}{\partial Y} + s\frac{\partial}{\partial X}\right)\right]\right\}W. \tag{17}$$

III. EDGE WAVE SOLUTIONS

Following the approach of Konenkov,[1] a solution of the governing equation (1) is sought, of the form

$$W = \Re[(A_1 e^{-\gamma_1 k_0 Y} + A_2 e^{-\gamma_2 k_0 Y})e^{i(\xi k_0 X - \omega t)}], \tag{18}$$

where ω is the frequency and k_0 is a suitable wave number selected so that the parameter ξ lies in the range $(1,\infty)$. For exponential decay away from the free edge, the coefficients γ_1 and γ_2 must lie in the open right half plane. This represents a flexural edge wave solution, i.e., where the energy is confined to the vicinity of $Y=0$, propagating in the positive X direction. Clearly, for a comprehensive study of the existence of such waves, the case of propagation in the $-X$ direction must also be examined. For convenience this is most neatly incorporated by maintaining $\xi\in(1,\infty)$ and instead without loss of generality allowing the orientation angle θ to range over $[-\pi/2,\pi/2]$ (Fig. 1). This symmetry is easily observed by changing $\sin(\theta)$ to $\sin(-\theta)$, and $\partial/\partial X$ to $-\partial/\partial X$, in (15)–(17); the governing equation and boundary conditions remain unaltered.

To simplify the subsequent analysis, it is convenient to write

$$\gamma_j = -i\xi\lambda_j, \tag{19}$$

for $j\in\{1,2\}$, and thus both λ_1 and λ_2 must lie in the open upper half plane. The edge wave now becomes

$$W=(A_1e^{i\xi k_0\lambda_1 Y}+A_2e^{i\xi k_0\lambda_2 Y})e^{i(\xi k_0 X-\omega t)} \tag{20}$$

and substitution into Eq. (15) yields the auxiliary equations:

$$D_x(s\lambda_j-c)^4+2H(s\lambda_j-c)^2(c\lambda_j+s)^2+D_y(c\lambda_j+s)^4 - \frac{R}{\xi^4}=0, \quad j\in\{1,2\}, \tag{21}$$

where

$$R=\frac{\rho h\omega^2}{k_0^4}. \tag{22}$$

This is a quartic polynomial in λ, with real coefficients, which can be rewritten in the form

$$a_1\lambda^4+b_1\lambda^3+a_2\lambda^2+b_2\lambda+\left(a_3-\frac{R}{\xi^4}\right)=0, \tag{23}$$

with

$$a_1=s^4D_x+c^4D_y+2s^2c^2H, \tag{24}$$

$$b_1=4sc[-s^2D_x+c^2D_y+H(s^2-c^2)], \tag{25}$$

$$a_2=6s^2c^2(D_x+D_y-2H)+2H, \tag{26}$$

$$b_2=4sc[-c^2D_x+s^2D_y-H(s^2-c^2)], \tag{27}$$

$$a_3=c^4D_x+s^4D_y+2s^2c^2H. \tag{28}$$

As stated, the edge wave must decay as $Y\to\infty$ and so the two roots λ_1 and λ_2 must have positive imaginary parts; the two remaining roots of (23) are therefore their complex conjugates λ_1^* and λ_2^*. Hence (29) may be cast as

$$a_1(\lambda-\lambda_1)(\lambda-\lambda_1^*)(\lambda-\lambda_2)(\lambda-\lambda_2^*)=0 \tag{29}$$

from which it can be deduced that the constant term is

$$a_3-\frac{R}{\xi^4}=a_1|\lambda_1|^2|\lambda_2|^2. \tag{30}$$

Now, the inequalities in (3), (4) can be employed to show that

$$a_1=s^4D_x+c^4D_y+2s^2c^2H>s^4D_x+c^4D_y+2s^2c^2 2D_1> s^4D_x+c^4D_y-2s^2c^2\sqrt{D_xD_y}=(s^2\sqrt{D_x}-c^2\sqrt{D_y})^2\geqslant 0 \tag{31}$$

and similarly

$$a_3>(c^2\sqrt{D_x}-s^2\sqrt{D_y})^2\geqslant 0 \tag{32}$$

for *all* values of the parameters. Therefore, the right-hand side of (30) is greater than zero and rearranging gives the bound

$$\xi^4>\frac{R}{a_3}. \tag{33}$$

It is thus convenient to choose k_0 so that

$$R=a_3=c^4D_x+s^4D_y+2s^2c^2H, \tag{34}$$

thereby ensuring that $\xi\in(1,\infty)$. Note that this particular value of k_0 is the wave number of a plane wave traveling in the X direction, and the positivity of the normalized wave phase speed $1/\xi$ is a direct consequence of the positivity of the strain energy in the plate.

Now, on using Eqs. (16) and (17), and applying the boundary conditions $M_Y=0$ and $V_Y=0$ along the line $Y=0$, it can be shown that

$$\sum_{j=1}^{2} A_j[D_xs^2(s\lambda_j-c)^2+D_yc^2(c\lambda_j+s)^2 +2Hsc(s\lambda_j-c)(c\lambda_j+s)+D_1]=0, \tag{35}$$

and

$$\sum_{j=1}^{2} A_j[D_xs(s\lambda_j-c)^2(s\lambda_j-2c)+D_yc(c\lambda_j+s)^2(c\lambda_j+2s) +2H(s\lambda_j-c)(c\lambda_j+s)(sc\lambda_j+s^2-c^2)-D_1\lambda_j]=0. \tag{36}$$

The unknown quantities A_1 and A_2 can be then eliminated between these last two equations to give the dispersion relation:

$$\mathcal{F}(\lambda_1,\lambda_2)=\mathcal{F}(\lambda_2,\lambda_1) \tag{37}$$

in which

$$\begin{aligned}\mathcal{F}(\lambda_1,\lambda_2)\equiv&[D_xs^2(s\lambda_1-c)^2+D_yc^2(c\lambda_1+s)^2\\ &+2Hsc(s\lambda_1-c)(c\lambda_1+s)+D_1]\\ &\times[D_xs(s\lambda_2-c)^2(s\lambda_2-2c)+D_yc(c\lambda_2+s)^2\\ &\times(c\lambda_2+2s)+2H(s\lambda_2-c)(c\lambda_2+s)\\ &\times(sc\lambda_2+s^2-c^2)-D_1\lambda_2].\end{aligned} \tag{38}$$

The roots, λ_j, $j=1,2$ of the polynomial (21) can be substituted into (38) so that it is a single equation for the wave number ξ. In the cases treated in earlier literature, both γ_1 and γ_2 were found to be positive real. This corresponds to positive imaginary values of λ_1 and λ_2, but is impossible in general, since setting $\lambda=iL$, with $L\in\mathbb{R}^+$, and taking the imaginary part of Eq. (21) yields

$$sc[D_xs^2-D_yc^2+H(c^2-s^2)]L^2 =sc[D_xc^2-D_ys^2-H(c^2-s^2)], \tag{39}$$

from which it is clear that two distinct real positive values of L cannot exist except in cases where both sides of this equation are zero, which are discussed in the following section. In general, therefore, edge waves must exhibit oscillation as well as decay in the Y direction, and Eq. (37) must be solved numerically; this is performed in Sec. V.

At the beginning of this section the angle of orientation, θ, between the free edge $Y=0$ and the principal direction of orthotropy $y=0$ was taken over $[-\pi/2,\pi/2]$ in order to account for all possible edge wave solutions. It is now convenient to relate the form of the edge wave for negative values of θ (or left traveling waves) to that for positive θ (right traveling waves) and thereby restrict attention henceforth, without loss of generality, to $\theta\in[0,\pi/2]$. If the angle θ is replaced by $-\theta$ in the auxiliary equation (23) then the coefficients a_j, $j=1,2,3$ remain the same while b_1 and b_2 change their signs. Therefore, the four roots are related to the

TABLE I. Coefficient values in each of the special cases.

Case	a_1	a_2	a_3	b
i	D	$2D$	D	D_1
ii	D_y	$2H$	D_x	D_1
iii	D_x	$2H$	D_y	D_1
iv	$(H+D)/2$	$3D-H$	$(H+D)/2$	$(D-H)/2+D_1$

former roots; these are now $-\lambda_1^*$, $-\lambda_2^*$ in the upper half plane and $-\lambda_1$, $-\lambda_2$ in the lower half plane. From this it can be concluded that the change of sign of the orientation angle yields a change in sign in the real part of λ_j, $j=1,2$, while the imaginary part (the decay rate in the Y direction) remains the same. With the changed λ_j the dispersion relation (37) becomes

$$\mathcal{F}(-\lambda_1^*,-\lambda_2^*)|_{\theta\to-\theta}=\mathcal{F}(-\lambda_2^*,-\lambda_1^*)|_{\theta\to-\theta}, \tag{40}$$

where $\mathcal{F}$ is given in Eq. (38). This is just the conjugate equation to Eq. (37) and so yields the same *real* root ξ, if one exists. In conclusion, left and right propagating edge waves have the same phase speed, and decay with identical exponents in the direction Y perpendicular to the edge. However, inspection of Eqs. (35) and (36) reveals that A_j, $j=1,2$ are changed to their complex conjugates as θ changes sign, and for left traveling waves the oscillatory behavior in Y for each of the two exponents in W (20) is exactly out of phase with that for right traveling waves.

IV. SPECIAL CASES

As well as allowing the possibility of real values of γ_1 and γ_2, the vanishing of both sides of Eq. (39) gives rise to four special cases in which an analytical solution is easily obtained. It is noted that, in view of the inequalities (2), the terms of Eq. (23) which are of order zero, two, and four in λ are strictly positive; hence meaningful simplification occurs only when the coefficients of λ^3 and λ, b_1 and b_2, are simultaneously zero. By inspection of Eq. (39), it is evident that there are four cases to consider, namely:

(i) $D_x=D_y=H=D$ (say);
(ii) $\theta=0$;
(iii) $\theta=\pi/2$;
(iv) $D_x=D_y=D$ (say), $\theta=\pi/4$.

Of the above, (i) is that of isotropy, as investigated by Konenkov,[1] Thurston and McKenna,[2] etc., (ii) and (iii) are essentially equivalent and were discussed by Norris,[8] and (iv) appears to be new. In each case, Eq. (23) is reduced to the form

$$a_1\lambda_j^4+a_2\lambda_j^2+a_3\left(1-\frac{1}{\xi^4}\right)=0, \tag{41}$$

and the dispersion relation (37) becomes

$$\lambda_2[a_1\lambda_1^2+b][a_1\lambda_2^2+(a_2-b)] =\lambda_1[a_1\lambda_2^2+b][a_1\lambda_1^2+(a_2-b)]. \tag{42}$$

For each of the above-mentioned cases, the values of a_1, a_2, a_3, and b are given in Table I. It is important to note that, in each case, due to the inequalities (2), the following relations hold:

$$a_1>0,\quad a_3>0,\quad a_2-2b>0,\ a_1a_3>b^2. \tag{43}$$

An example of case (iv) could be, for instance, a fiber reinforced plate with identical fibers aligned in both x and y directions. Clearly, $\theta=\pm\pi/4$ are also lines of symmetry, and so the plate can be viewed as orthotropic with principal axes X and Y when $\theta=\pi/4$. From Table I it is seen that, in this rotated frame, the effective H value is $a_2/2=(3D-H)$ and the bending stiffnesses (a_1 and a_3) are $D_X=D_Y=(H+D)/2$.

Equation (41) may be solved to give the values of λ_1 and λ_2 in terms of the as yet unknown parameter ξ. Thus,

$$\lambda_j^2=\frac{1}{2a_1}\left[-a_2+(-1)^j\left[a_2^2-4a_1a_3\left(1-\frac{1}{\xi^4}\right)\right]^{1/2}\right], \tag{44}$$

which immediately yields the relations

$$\lambda_1^2+\lambda_2^2=-\frac{a_2}{a_1}, \tag{45}$$

and

$$\lambda_1^2\lambda_2^2=\frac{a_3}{a_1}\left(1-\frac{1}{\xi^4}\right). \tag{46}$$

Now, since λ_1 and λ_2 must lie in the upper half plane, it follows that either λ_1^2 and λ_2^2 are real and negative, or [by Eq. (44)], they are complex conjugates. In either case, we have

$$\lambda_1^2=re^{i\phi},$$

for some positive r and $\phi\in[-\pi,0)$. Again using the fact that λ_1 and λ_2 must lie in the upper half plane, this leads to

$$\lambda_1=\sqrt{r}e^{i(\phi/2+\pi)},$$

and

$$\lambda_2=\sqrt{r}e^{-i\phi/2}.$$

Consequently, it is clear from (46) that

$$\lambda_1\lambda_2=-\sqrt{a_3/a_1}\sqrt{1-1/\xi^4}. \tag{47}$$

In view of this, along with Eq. (45), the dispersion relation (42) reduces to

$$a_1a_3\left(1-\frac{1}{\xi^4}\right)+\sqrt{a_1a_3}(a_2-2b)\sqrt{1-\frac{1}{\xi^4}}-b^2=0. \tag{48}$$

Possible roots for ξ are now easily obtained; factoring the quadratic in (48) yields

$$\sqrt{1-\frac{1}{\xi^4}}=\frac{-(a_2-2b)\pm\sqrt{(a_2-2b)^2+4b^2}}{2\sqrt{a_1a_3}},$$

and since the left-hand side is positive, we must choose the "+" sign to obtain a solution. Thus, the unique root for ξ is the positive value which satisfies

$$\frac{1}{\xi^4}=1-\frac{1}{4a_1a_3}[(a_2-2b)-\sqrt{(a_2-2b)^2+4b^2}]^2. \tag{49}$$

TABLE II. Material parameters (modulii units are GPa).

Reinforcing fibers	E_1	E_2	ν_{12}	G_{12}
Glass	54.2	18.1	0.250	8.96
Boron	208	20.8	0.300	6.95
Graphite	208	5.21	0.250	2.59
Bidirectional glass	54.2	54.2	0.250	8.96

Finally, note that a necessary and sufficient condition for pure imaginary values of λ_1 and λ_2 in these cases is obtained by substituting (49) into Eq. (44); the resulting inequality is

$$(a_2-2b)[-a_2+6b+2\sqrt{(a_2-2b)^2+4b^2}]>0.$$

Using the fact that $a_2>2b$, it is straightforward to show that λ_1 and λ_2 lie at distinct points on the imaginary axis provided that

$$a_2>-\tfrac{2}{3}b, \tag{50}$$

otherwise the solution will exhibit oscillation away from the free edge. Note that, since a_2 is greater than $2b$, the inequality (50) can only be violated if b is negative. In addition, a repeated root may arise (i.e., $\lambda_1=\lambda_2$; this occurs when

$$a_2=-\tfrac{2}{3}b. \tag{51}$$

In this case, it can be expected that a solution of the form

$$W=(A_1+A_2Y)e^{-\gamma_1 k_0 Y}e^{i(\xi X k_0-\omega t)}$$

holds instead of the standard edge wave ansatz (18). It is important to note, however, that in general the parameters required to satisfy Eq. (51) will not occur.

V. NUMERICAL SOLUTIONS AND CONCLUDING REMARKS

Numerical solutions of Eq. (21) and (37) were obtained by using a linear search to locate an interval in which the function

$$f(\xi)=\mathcal{F}(\lambda_1,\lambda_2)-\mathcal{F}(\lambda_2,\lambda_1),$$

possesses a zero. The bisector algorithm was then employed to find a value of ξ for which

$$|f(\xi)|<10^{-12}.$$

This was carried out for four sets of parameters, representing epoxy resin plates with a variety of reinforcing fibers (see Table II). The first three sets of data are taken from Norris,[8] and the fourth is included to illustrate special case (iv) discussed in Sec. IV. In each example, precisely one value of ξ was found in the interval $(1,\infty)$; these results are shown in Fig. 2. The ordinate of this figure is $1/\xi$, which is the edge wave phase speed normalized with respect to the wave speed of flexural waves propagating along the X direction, i.e., ω/k_0. Note the proximity of the solutions to $\xi=1$ when $\theta\approx 0$, $\pi/2$, and also that the plot for bidirectional glass is symmetric about $\theta=\pi/4$ as expected since $D_x=D_y$ here. To reveal the *absolute* magnitudes of the wave speeds as θ varies, Fig. 2 is replotted for the nondimensional propagation speed c_n in the various media (Fig. 3), where

$$c_n=\left(\frac{R}{D_x}\right)^{1/4}\frac{1}{\xi}=\frac{\omega^{1/2}}{k_0\xi}\left(\frac{\rho h}{D_x}\right)^{1/4},$$

i.e., this is the phase speed relative to that for a plane flexural wave traveling in the x direction. As one might intuitively expect, the magnitude in the change of c_n with angle of inclination appears to be directly related to the value of $|E_1-E_2|$ (and, consequently $|D_x-D_y|$), with materials exhibiting stronger orthotropy experiencing greater changes in phase speed. Thus, for graphite fibers, with E_1/E_2 around 40, c_n varies from unity to less than 0.4 as θ increases to $\pi/2$.

Values of γ_1 and γ_2 associated with Figs. 2 and 3 were also computed; all three plates with unidirectional fiber support exhibited similar behavior in this respect. Plots of the real and imaginary part of γ_1 and γ_2 for boron fibers are shown in Figs. 4 and 5. The real parts of both γ_1 and γ_2

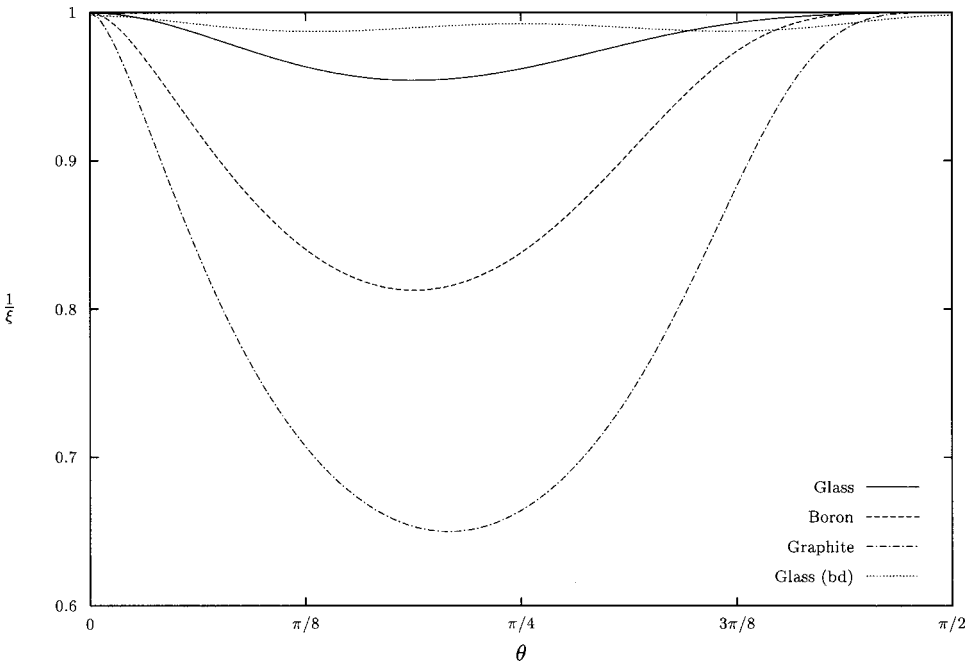


FIG. 2. Phase speed of edge waves in the respective media normalized with respect to flexural waves propagating in the X direction, i.e., parallel to the plate edge.

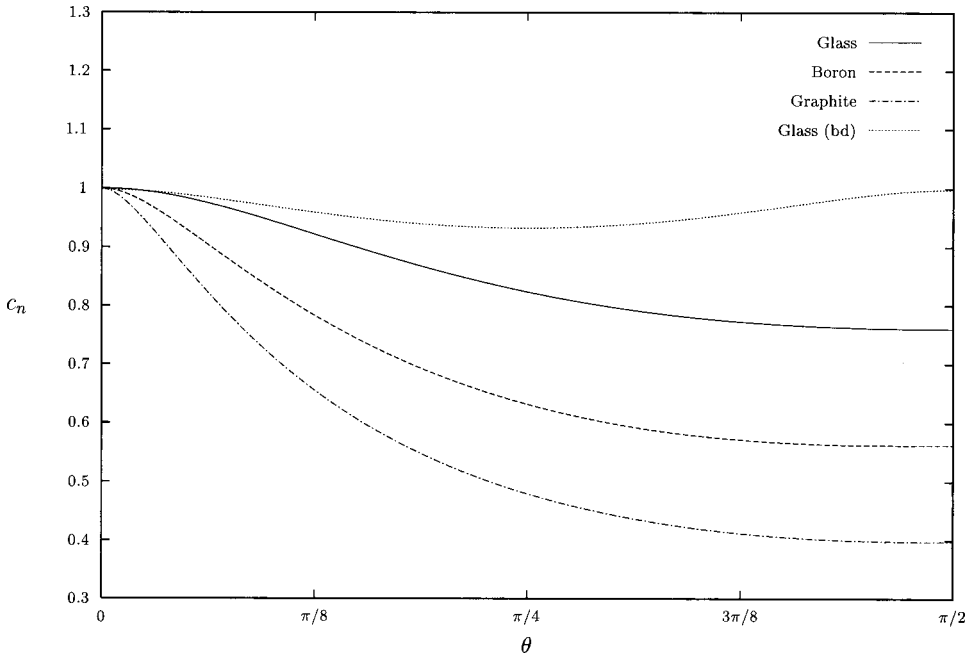


FIG. 3. Phase speed of edge waves in the respective media normalized with respect to flexural waves propagating in the principal x direction.

show a clear maximum, indicating the point of strongest decay for their respective exponent term in W (18), though the maxima do not occur at quite the same angle of inclination. Note also that the occurrences of $\Im(\gamma_1)=\Im(\gamma_2)=0$ agree with the results of Sec. IV. Different behavior was observed in the values of γ_1 and γ_2 for bidirectional glass fibers, Figs. 6 and 7, as might be expected. Here, the curves are symmetric about $\theta=\pi/4$, and both $\Re(\gamma_1)$ and $\Re(\gamma_2)$ exhibit maxima at this angle. In this case, $\Im(\gamma_1)=\Im(\gamma_2)=0$ for $\theta \in\{0,\pi/4,\pi/2\}$, again agreeing with the results of Sec. IV. In all cases, the weakest decay is observed at the limit values $\theta=0$ and $\theta=\pi/2$, since the curves for $\Re(\gamma_1)$ and $\Re(\gamma_2)$ both exhibit minima here. This is to be expected, since, when $\theta\approx 0,\ \pi/2$, the wave number is close to $\xi=1$, i.e., the value for straight crested flexural waves.

Further understanding of the oscillatory behavior associated with complex values of γ_1 and γ_2 can be gained by considering the flexural displacement with nondimensional arguments, defined as

$$\hat{W}=\Re\left\{(A_1+A_2)\sum_{j=1}^{2}\left(\frac{A_j}{A_1+A_2}\right)\exp[-\gamma_j(D_x/R)^{1/4}\hat{Y}]\times\exp[i\xi(D_x/R)^{1/4}\hat{X}-i\omega t]\right\}. \tag{52}$$

Here, the spatial coordinates have been normalized with respect to the wave number of flexural waves in the x direction, i.e.,

$$\hat{X}=\left(\frac{\rho h\omega^2}{D_x}\right)^{1/4}X,$$

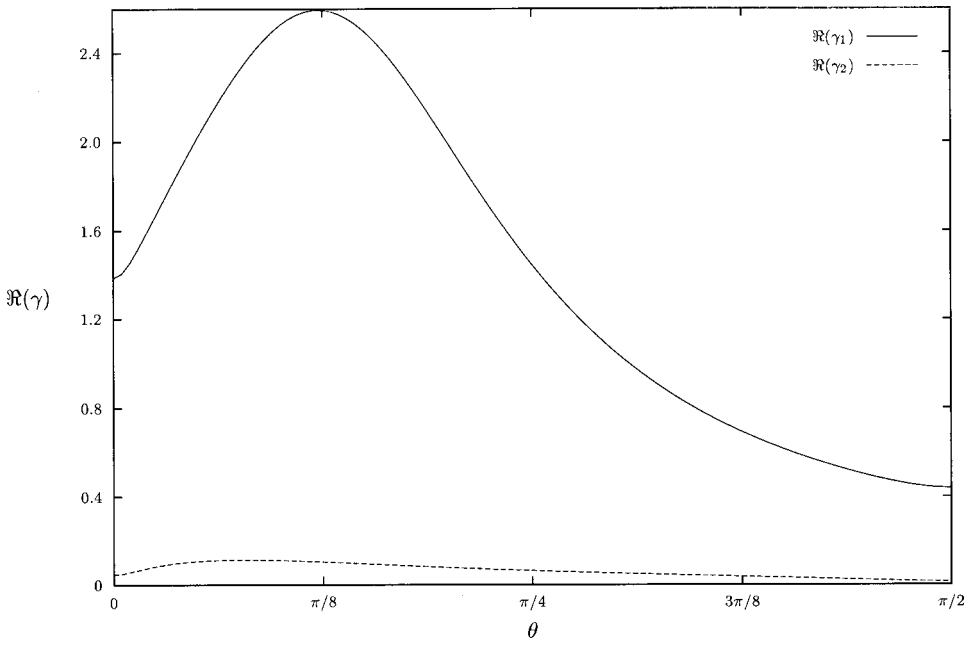


FIG. 4. Values of the decay coefficients $\Re(\gamma_j)$ for boron/epoxy.

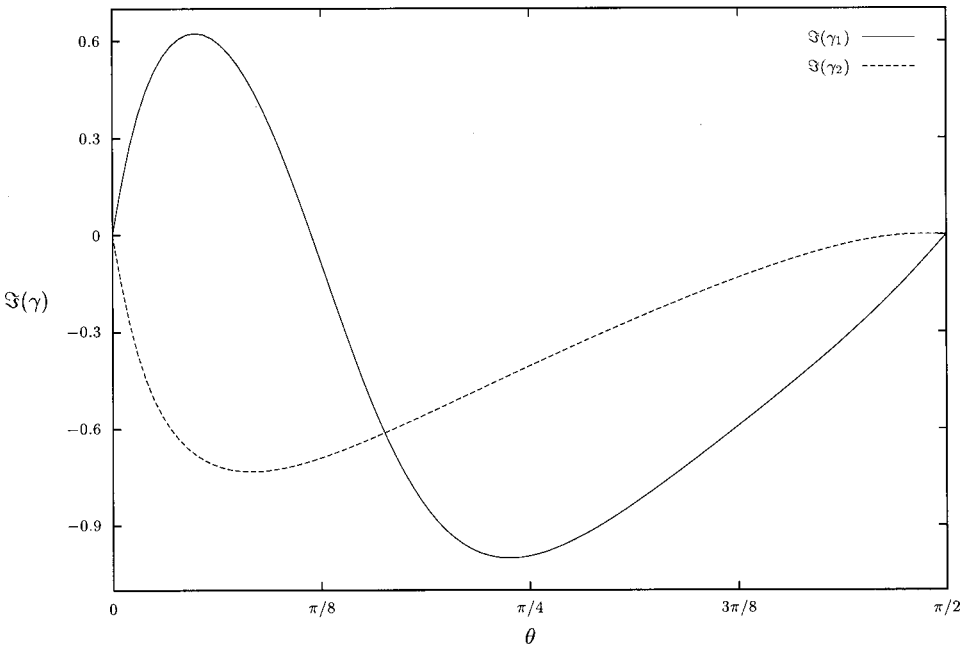


FIG. 5. Values of the oscillation coefficients $\Im(\gamma_j)$ for boron/epoxy.

and

$$\hat{Y}=\left(\frac{\rho h\omega^2}{D_x}\right)^{1/4} Y.$$

A contour plot of the function $\hat{W}$ can then be produced, for a given value of the (arbitrary) constant A_1+A_2 and for any fixed time. This consists (see Fig. 8) of a regular pattern of lobes of alternating sign, separated by "null" lines on which $\hat{W}=0$. In all of these cases investigated herein it is found that $\Re(\gamma_1)\gg\Re(\gamma_2)$ (also $A_1/A_2<0.3$), so the behavior of the wave is almost completely determined by the relative sizes of $\Re(\gamma_2)$ and $\Im(\gamma_2)$, i.e.,

$$\hat{W}\approx\Re\{A_2\exp[(D_x/R)^{1/4}(i\xi\hat{X}-\gamma_2\hat{Y})-i\omega t]\}.$$

By explicitly separating real and imaginary parts, it is easy to show that an accurate estimate for the angle of deviation, ϕ, of the null lines from the direction normal to the free edge, $\hat{Y}$, is given by

$$\phi\approx\arctan\left(-\frac{\Im(\gamma_2)}{\xi}\right),$$

where ϕ is measured positive counterclockwise looking down onto the plate as in Fig. 8. Thus in the limit cases $\theta=0$ and $\theta=\pi/2$, the null lines are perpendicular to the edge, since $\Im(\gamma_2)=0$ here. In general the null lines are inclined to the edge, though one should note that despite efforts by the authors no obvious relation between ϕ and the principal directions could be ascertained. This is borne out by a plot of ϕ against the inclination angle θ given in Fig. 9. For the materials with unidirectional fiber support, $\phi>0$, since $\Im(\gamma_2)<0$, so the null lines tilt away from the direction of propagation. For bidirectional glass fiber support, however, $\phi>0$

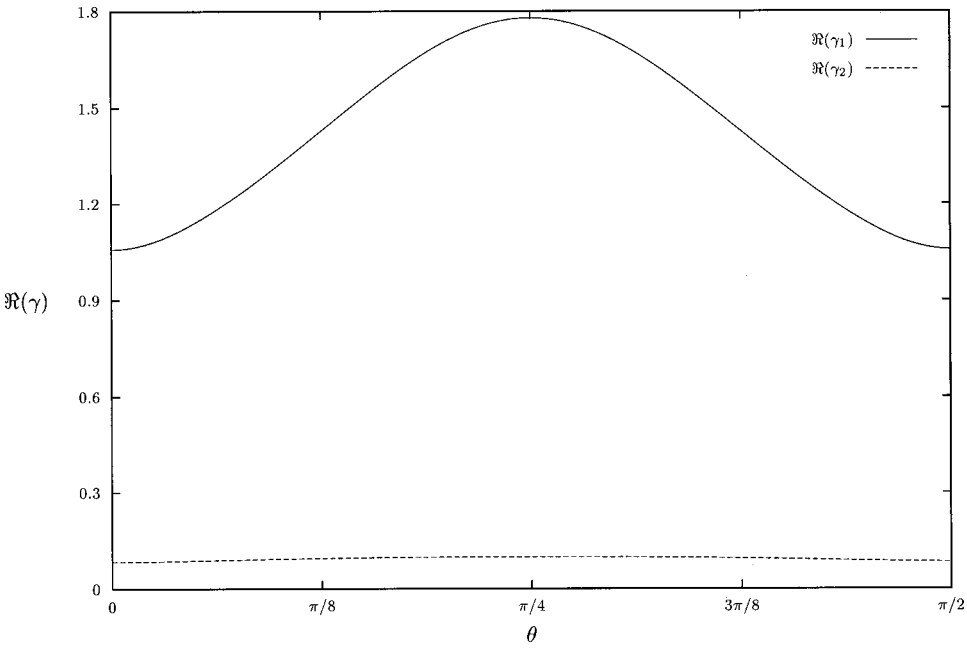


FIG. 6. Values of the decay coefficients $\Re(\gamma_j)$ for bidirectional glass/epoxy.

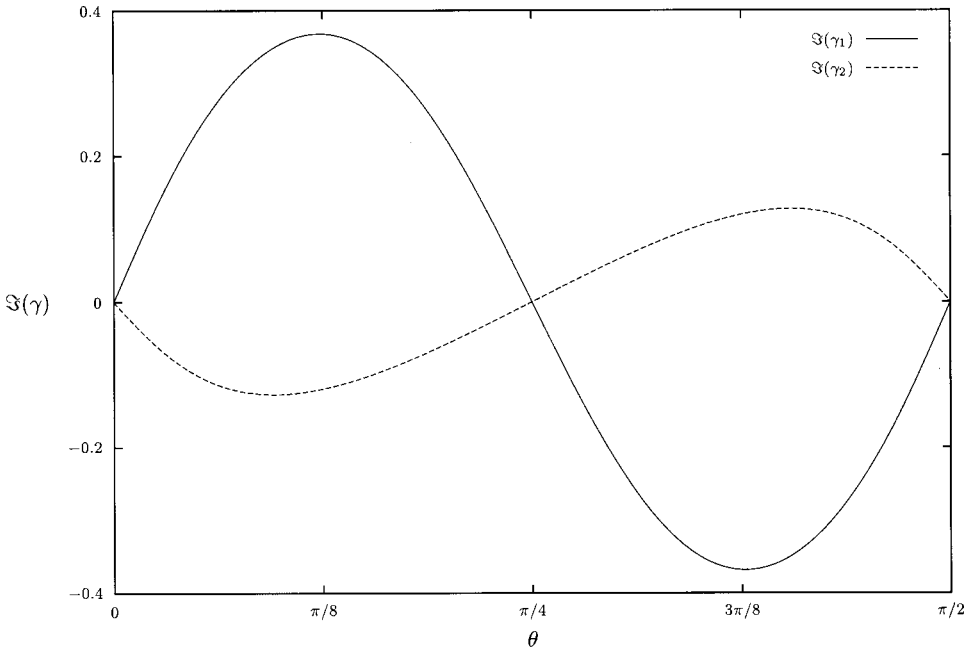


FIG. 7. Values of the oscillation coefficients $\Im(\gamma_j)$ for bidirectional glass/epoxy.

for $0<\phi<\pi/4$, and $\phi<0$ for $\pi/4<\theta<\pi/2$. Due to the inclination of the null lines, a straight line perpendicular to the edge (in the $\hat{Y}$ direction) will pass alternately through regions in which $\hat{W}$ is positive and negative, thereby giving rise to the oscillations shown in earlier figures.

To summarize, numerical evidence suggests that, as for the special cases discussed in Sec. IV and in previous literature, in general a unique edge wave solution exists for every set of material parameters and inclination angle. The wave generally exhibits both oscillation and decay away from the free edge; the oscillation arising due to the "tilting" of the solution pattern which occurs when $\theta\in(0,\pi/2)$ (see Figs. 8

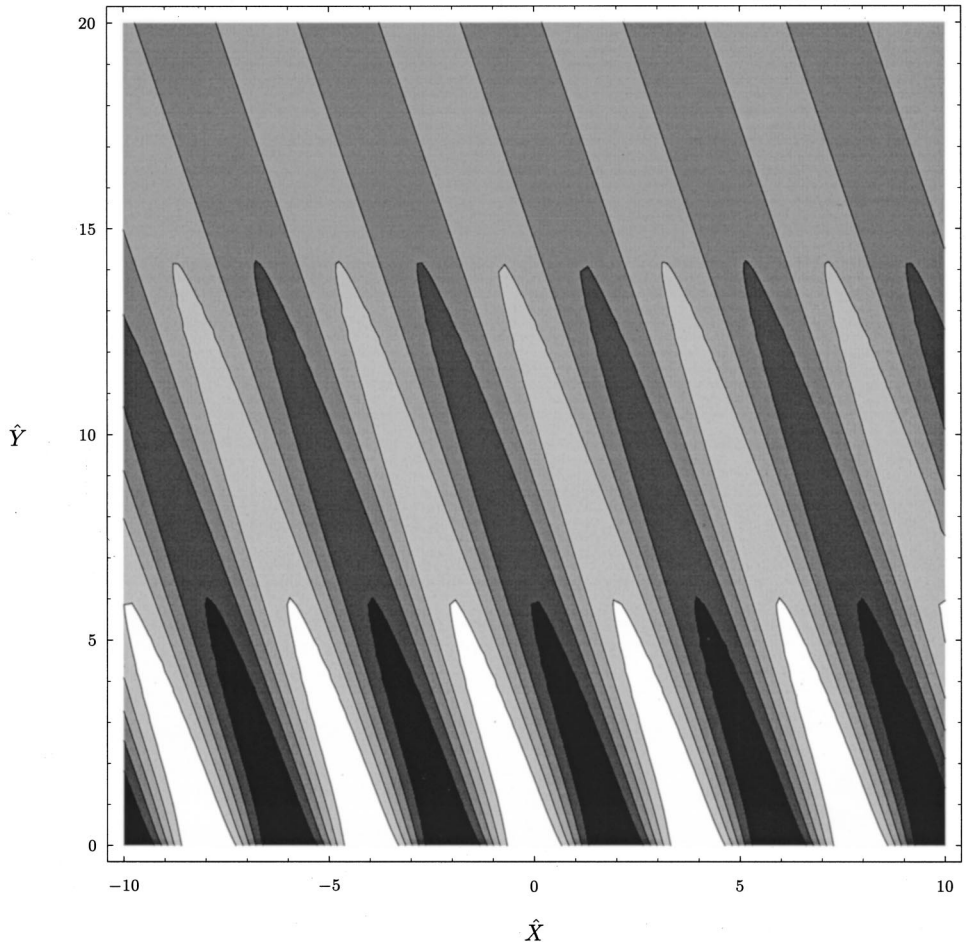


FIG. 8. Contour plot of the nondimensional flexural displacement $\hat{W}$ at time $t=0$, for boron/epoxy in the case $\theta=\pi/4$, with $A_1+A_2=1$. Positive values of $\hat{W}$ occur in regions with lighter shading; dark shading indicates $\hat{W}<0$.

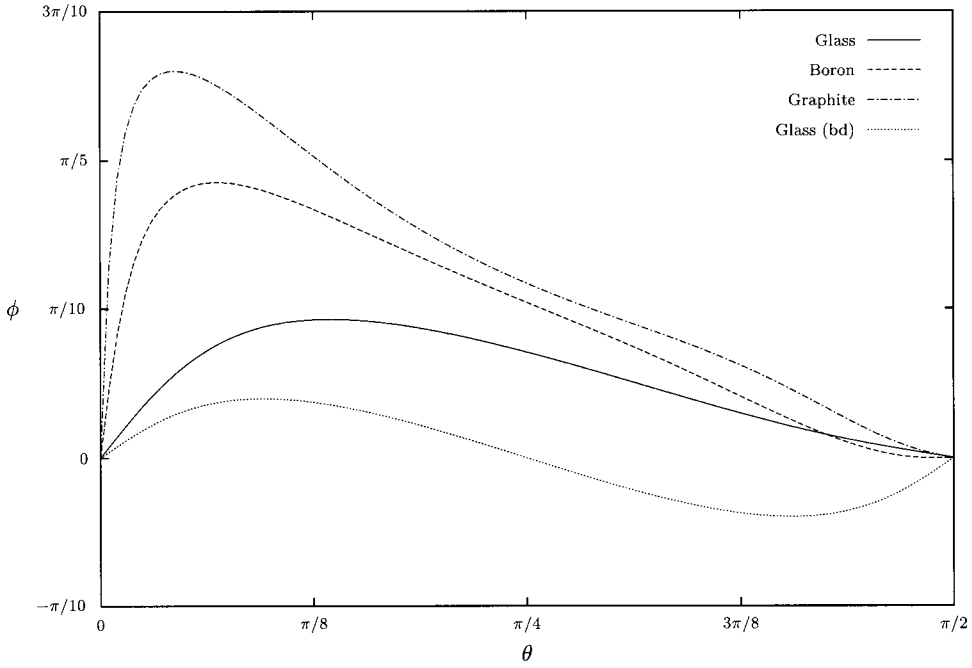


FIG. 9. Deviation angle ϕ of the null lines from the direction normal to the free edge (measured positive counterclockwise).

and 9 and the discussion in the preceding paragraph). Edge waves propagating in the negative X direction travel with the same speed as their counterparts with phase vector pointing in the positive X direction; however, their oscillatory behavior in the orthogonal direction Y is π out of phase. That is, in Eq. (18), when X is changed to $-X$, the complex conjugate of the expression in parentheses is taken. A final important point to note concerns the values of the parameters for which edge waves have pure decay in the Y direction. Norris's previous article[8] might lead a reader to hold the mistaken impression that, for $\theta=0$, $\pi/2$ [i.e., cases (ii) and (iii) of Sec. IV], all parameter values will lead to $\Im(\gamma_j)=0$, $j=1,2$. This is not true, and by inspection of his Eq. (7), or equivalently (45), γ_j, $j=1,2$, certainly cannot be positive real when H (or a_2) is less than or equal to zero. However, there are also positive values of H for which the Y behavior is oscillatory. This can be deduced from the inequality (50), which, from Table I, implies that pure decay away from the edge occurs if and only if

$$H > -\frac{D_1}{3} \tag{53}$$

or equivalently

$$D_{xy} > -\frac{2D_1}{3}, \tag{54}$$

and so for $D_1<0$ this yields a window of positive values of H and D_{xy} for which this condition is violated.

ACKNOWLEDGMENTS

The authors are grateful to the Engineering and Physical Sciences Research Council, UK for the support of a research studentship for Thompson and to the Leverhulme Trust, UK, for a *Grant in Aid of Research* for Abrahams.

[1] Yu. K. Konenkov, "A Rayleigh-type flexural wave," Sov. Phys. Acoust. **6**, 122–123 (1960).
[2] R. N. Thurston and J. McKenna, "Flexural acoustic waves along the edge of a plate," IEEE Trans. Sonics Ultrason. **21**, 296–297 (1974).
[3] B. K. Sinha, "Some remarks on propagation characteristics of ridge guides for acoustic waves at low frequencies," J. Acoust. Soc. Am. **56**, 16–18 (1974).
[4] R. M. De La Rue, "Experimental and theoretical studies of guided acoustic surface wave propagation," Ph.D. thesis, University College, London, 1972.
[5] C. Kauffmann, "A new bending wave solution for the classical plate equation," J. Acoust. Soc. Am. **104**, 2220–2222 (1998).
[6] R. N. Thurston, G. D. Boyd, and J. McKenna, "Plate theory solution for guided flexural acoustic waves along the tip of a wedge," IEEE Trans. Sonics Ultrason. **21**, 178–186 (1974).
[7] V. V. Krylov, "Geometrical-acoustics approach to the description of localized modes of an elastic solid wedge," Sov. Phys. Tech. Phys. **35**, 137–140 (1990).
[8] A. N. Norris, "Flexural edge waves," J. Sound Vib. **171**, 571–573 (1994).
[9] A. N. Norris, V. V. Krylov, and I. D. Abrahams, "Flexural edge waves and comments on 'a new bending wave solution for the classical plate equation,'" J. Acoust. Soc. Am. **107**, 1781–1784 (2000).
[10] P. E. Lagasse, "Higher-order finite-element analysis of topographic guides supporting elastic surface waves," J. Acoust. Soc. Am. **53**, 1116–1122 (1973).
[11] R. Burridge and F. J. Sabina, "Theoretical computations on ridge acoustic surface waves using the finite element method," Electron. Lett. **7**, 720–722 (1971).
[12] R. Burridge and F. J. Sabina, "The propagation of elastic surface waves guided by ridges," Proc. R. Soc. London, Ser. A **330**, 417–441 (1972).
[13] P. E. Lagasse and A. A. Oliner, "Acoustic flexural mode on a ridge of semi-infinite height," Electron. Lett. **12**, 11–13 (1976).
[14] V. V. Krylov, "On the velocities of localized vibration modes in immersed solid wedges," J. Acoust. Soc. Am. **103**, 767–770 (1998).
[15] M. de Billy, "On the influence of loading on the velocity of guided elastic waves in linear elastic wedges," J. Acoust. Soc. Am. **100**, 659–662 (1996).
[16] I. D. Abrahams and A. N. Norris, "On the existence of flexural edge waves on submerged elastic plates," Proc. R. Soc. London, Ser. A **456**, 1559–1582 (2000).
[17] I. Thompson and I. D. Abrahams, "Diffraction of flexural waves by a semi-infinite crack in a fibre reinforced thin elastic plate" (unpublished).
[18] S. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells*, 2nd ed. (McGraw–Hill, New York, 1959).
[19] K. F. Graff, *Wave Motion in Elastic Solids* (Dover, New York, 1991).

Elastic guided waves in a layered plate with rectangular cross section

O. M. Mukdadi
Department of Mechanical Engineering, University of Colorado, Boulder, Colorado 80309-0427

Y. M. Desai
Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

S. K. Datta
Department of Mechanical Engineering, University of Colorado, Boulder, Colorado 80309-0427

A. H. Shah
Department of Civil and Geological Engineering, University of Manitoba, Winnipeg, Manitoba R3T 5V6, Canada

A. J. Niklasson
Department of Mechanics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

(Received 26 June 2001; revised 15 June 2002; accepted 22 June 2002)

Guided waves in a layered elastic plate of rectangular cross section (finite width and thickness) has been studied in this paper. A semianalytical finite element method in which the deformation of the cross section is modeled by two-dimensional finite elements and analytical representation of propagating waves along the length of the plate has been used. The method is applicable to arbitrary number of layers and general anisotropic material properties of each layer, and is similar to the stiffness method used earlier to study guided waves in a laminated composite plate of infinite width. Numerical results showing the effect of varying the width of the plate on the dispersion of guided waves are presented and are compared with those for an infinite plate. In addition, effect of thin anisotropic coating or interface layers on the guided waves is investigated. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1509426]

PACS numbers: 43.20.Mv [SKD]

I. INTRODUCTION

There is considerable recent interest in guided ultrasonic waves in layered anisotropic plates for ultrasonic characterization of material properties. In order to develop/improve such experimental methods, mathematical modeling and numerical simulation are of great importance. Most of the recent studies have dealt with plates having infinite planar dimensions. Theory of guided waves in layered anisotropic plates has been discussed by Nayfeh,[1] Datta,[2] and Rose.[3] An extensive review of the use of guided waves in plates for material characterization has been done by Chimenti.[4]

In this paper, our attention is focused on the dispersion of guided elastic waves in an anisotropic layered plate having finite width. A problem of particular interest in this context is the effect of different width-to-thickness ratios on guided waves in tapes with anisotropic superconducting layers. Another problem of current interest is the propagation of acoustic phonon modes in rectangular bilayered anisotropic wires.[5] The current study is motivated by the desire to develop a fundamental understanding of dispersion of guided wave modes in plates of rectangular cross section. This would aid in the development of ultrasonic material characterization of tapes that are fabricated as layered metallic and superconducting materials for commercial high-current applications. Also, such model study is needed to accurately predict low-energy phonon scattering in nanostructures. Recently,[6–9] the effect of anisotropic superconducting coating and interface layers on the dispersion and mode coupling of guided waves was presented. There, the tapes were considered to be of infinite width. In practice, however, the tapes are of finite width $B(>H)$, where H is the thickness of a tape. In particular, our primary objective is to compare the results for symmetric three-layered plates with those for the homogeneous case when B/H is finite. Another objective is to compare these with dispersion in plates of infinite width $(B/H\rightarrow\infty)$.

A semianalytical finite element method similar to that used by Aalami[10] and Taweel *et al.*[11] has been adopted here. In this method, the rectangular cross section of the plate is discretized into a set of planar elements (Fig. 1). It is a generalization of the one-dimensional finite element (stiffness) method that was used[12,13] for the analysis of dispersion of Rayleigh–Lamb waves in layered anisotropic plates. The stiffness method has been shown to give results that agree well with analytical solutions. Comparison of model predictions with experimental observations has been successfully used to obtain the anisotropic elastic constants of a layer in a layered plate.[14,15]

In Ref. 5, acoustic phonon modes of a free-standing wire were studied by a variational method[16] developed for resonant ultrasonic spectroscopy. Variational methods were used earlier by Kynch[17] and Nigro[18] to study guided wave modes

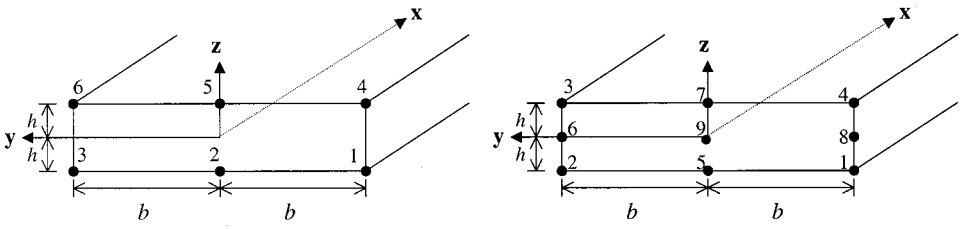


FIG. 1. Geometry of the six-node and nine-node elements.

in elastic isotropic bars of rectangular cross section. A detailed study of guided modes in isotropic bars of square and nearly square cross sections was made by Fraser,[19] who used a collocation method. The methods used in these investigations are suited for homogeneous bars of nearly square cross sections. In contrast, the solution technique adopted here is quite general and is applicable to general anisotropy and cross section of the plate. In the following, an outline of the formulation and solution procedure is presented first. For details, the reader is referred to earlier publications,[2,12,13] and the references therein.

II. FORMULATION

Time-harmonic waves in a multilayered anisotropic plate of rectangular cross section have been considered in this analysis. It is assumed that the tape consists of N parallel, homogeneous, and orthotropic layers, which are perfectly bonded together along the entire length of the plate. Global rectangular Cartesian coordinates (X, Y, Z) with origin on the central axis of the plate are chosen. The X axis coincides with the central axis and the Y- and Z-axes are parallel, respectively, to the width and thickness directions.

Two different finite element discretizations of the cross section have been used in this paper for the purpose of assessing computational efficiency and accuracy. In one, six-noded elements are used such that at each node the traction and displacement continuity is maintained in the thickness direction, while only displacement continuity is enforced in the width direction. Since one of the objectives of this investigation is to study the effect of layering and different aspect ratio of the plate on guided wave dispersion, maintaining traction continuity in the thickness direction would be desirable for analyzing interface stresses. It is a generalization of the one-dimensional finite element (stiffness) method that was used[12,13] for analyzing dispersion of guided waves in infinite layered anisotropic plates. In the other, nine-node quadratic elements that satisfy displacement continuity in both thickness and width directions are used. Relative merits of the two methods will be pointed out in the following.

For each element, the local coordinates (x, y, z) are taken to be parallel, respectively, to the global (X, Y, Z) coordinates. Typical six-node and nine-node elements with thickness $2h$ (in the z direction) and width $2b$ (in the y direction) are shown in Fig. 1.

A. Six-node element formulation

The displacement $\mathbf{u}(x,y,z,t)$ at a point within an element will be written in the product form

$$\mathbf{u}(x,y,z,t)=\mathbf{N}(y,z)\tilde{\mathbf{q}}(x,t), \tag{1}$$

where $\mathbf{N}$ contains the interpolated functions and $\tilde{\mathbf{q}}$ is the array of nodal displacement-stress variables. The stress–strain relations at a point inside an element are given by

$$\boldsymbol{\sigma}=\mathbf{D}\boldsymbol{\varepsilon}, \tag{2}$$

where

$$\boldsymbol{\sigma}=[\sigma_{xx}\ \ \sigma_{yy}\ \ \sigma_{zz}\ \ \tau_{xy}\ \ \tau_{xz}\ \ \tau_{yz}]^T, \tag{3a}$$

$$\mathbf{D}=\begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} & . & . \\ & D_{22} & D_{23} & D_{24} & . & . \\ & & D_{33} & D_{34} & . & . \\ & & & D_{44} & . & . \\ & & & & D_{55} & D_{56} \\ \text{Symm.} & & & & & D_{66} \end{bmatrix}, \tag{3b}$$

$$\boldsymbol{\varepsilon}=[\varepsilon_{xx}\ \ \varepsilon_{yy}\ \ \varepsilon_{zz}\ \ \gamma_{xy}\ \ \gamma_{xz}\ \ \gamma_{yz}]^T. \tag{3c}$$

The elastic coefficient matrix $\mathbf{D}$ is obtained by transforming the stiffness matrix for the material to the element coordinate system shown in Fig. 1. The global displacement components U, V, and W along the global X, Y, and Z directions for an element can be written as

$$\mathbf{U}=[U\ \ V\ \ W\ \ ]^T=\mathbf{B}_1\mathbf{q}_{,x}+\mathbf{B}_2\mathbf{q} \tag{4}$$

where,

$$\mathbf{q}=[q_1\ \ q_2\ \ q_3\ \ q_4\ \ q_5\ \ q_6]^T,$$

$$q_i=[U_i\ \ V_i\ \ W_i\ \ \tau_i\ \ \chi_i\ \ \sigma_i]^T,\ \ i=1,2,\ldots,6, \tag{5}$$

$$\mathbf{B}_j=[\mathbf{N}_{j1}\ \ \mathbf{N}_{j2}\ \ \mathbf{N}_{j3}\ \ \mathbf{N}_{j4}\ \ \mathbf{N}_{j5}\ \ \mathbf{N}_{j6}],\quad j=1,2,$$

where q_i consists of nodal degrees of freedom with $\tau_i=\tau_{yz_i}$, $\chi_i=\tau_{xz_i}$, and $\sigma_i=\sigma_{zz_i}$. These nodal values of the displacement and traction components are functions of x and t. It will be assumed that the time dependence is of the form $e^{-i\omega t}$, ω being the circular frequency. The factor $e^{-i\omega t}$ will be dropped in the sequel. Here

$$\mathbf{N}_{1i}=\begin{bmatrix} . & . & -g_i f_p & . & . & . \\ . & . & . & . & . & . \\ -g_i f_p \dfrac{D_{13}}{D_{33}} & -g_i f_p \dfrac{D_{34}}{D_{33}} & . & . & . & . \end{bmatrix}, \quad i=1,2,\ldots,6,$$

$$\mathbf{N}_{2i}=\begin{bmatrix} g_i f_q & . & . & g_i f_p \dfrac{D_{66}}{\Delta} & -g_i f_p \dfrac{D_{56}}{\Delta} & . \\ . & g_i f_q & h_i f_p & -g_i f_p \dfrac{D_{56}}{\Delta} & g_i f_p \dfrac{D_{55}}{\Delta} & . \\ h_i f_p \dfrac{D_{34}}{D_{33}} & h_i f_p \dfrac{D_{23}}{D_{33}} & g_i f_q & . & . & \dfrac{g_i f_p}{D_{33}} \end{bmatrix}, \quad i=1,2,\ldots,6., \tag{6}$$

with $p=3$, $q=1$ for $i=1,2,3$, and $p=4$, $q=2$ for $i=4,5,6$. Moreover

$$\begin{aligned} &f_1=\tfrac{1}{4}(2-3\xi+\xi^3), \quad f_2=\tfrac{1}{4}(2+3\xi-\xi^3), \\ &f_3=\frac{h}{4}(1-\xi-\xi^2+\xi^3), \quad f_4=\frac{h}{4}(-1-\xi+\xi^2+\xi^3), \\ &g_1=g_4=\frac{\eta}{2}(\eta-1), \qquad g_2=g_5=1-\eta^2, \\ &g_3=g_6=\frac{\eta}{2}(1+\eta), \\ &h_1=h_4=\frac{1}{2b}(3g_1+g_2-g_3), \quad h_2=h_5=\frac{2}{b}(g_3-g_1), \\ &h_3=h_6=\frac{1}{2b}(g_1-g_2-3g_3). \end{aligned} \tag{7}$$

Here

$$\xi=\frac{z}{h}, \quad \eta=\frac{y}{b}, \quad \text{and} \quad \Delta=D_{55}D_{66}-D_{56}^2. \tag{8}$$

The interpolation polynomials f_i and g_i are such that the displacement-traction vector q_i is continuous at the nodes lying on the interface $z=z_i$ and the displacement is continuous at the nodes lying on the interface $y=y_i$. As was shown by Datta *et al.*,[12] the enforcement of traction continuity at the interface between two layers in the plate gives better results at high frequencies with smaller number of nodes in the z direction. Also, this way the traction at the interfaces is determined directly as members of the eigenvector array. Since there are no material discontinuities in the width direction, and when the width is considered to be much larger than the depth, it is economical to use quadratic interpolation in y.

The differential equations of motion for an element can be obtained by applying Hamilton's variational principle. For this purpose, the Lagrangian L for an element can be computed from

$$L=\frac{1}{2}\int_t\int_x\int_{-h}^{h}\int_{-b}^{b}(\rho\dot{\mathbf{U}}\dot{\mathbf{U}}-\overline{\boldsymbol{\varepsilon}}^T\mathbf{D}\boldsymbol{\varepsilon})dy\,dz\,dx\,dt. \tag{9}$$

Here, ρ indicates mass density per unit volume and the overbar denotes complex conjugate. By substituting Eq. (3) in the strain–displacement relations, and the resulting expression in Eq. (9), the expression for L can be computed to be

$$\begin{aligned} L=\frac{1}{2}\int_t\int_x[\,&\omega^2(\overline{\mathbf{q}}^T\mathbf{c}_1\mathbf{q}'+\overline{\mathbf{q}}'^T\mathbf{c}_2\mathbf{q}+\overline{\mathbf{q}}^T\mathbf{c}_2^T\mathbf{q}'+\overline{\mathbf{q}}^T\mathbf{m}\mathbf{q}) \\ &-\overline{\mathbf{q}}''^T\mathbf{e}_1\mathbf{q}''-\overline{\mathbf{q}}''^T\mathbf{e}_2\mathbf{q}'-\overline{\mathbf{q}}'^T\mathbf{e}_2^T\mathbf{q}''-\overline{\mathbf{q}}''^T\mathbf{e}_3\mathbf{q}-\overline{\mathbf{q}}^T\mathbf{e}_3^T\mathbf{q}'' \\ &-\overline{\mathbf{q}}'^T\mathbf{e}_4\mathbf{q}'-\overline{\mathbf{q}}'^T\mathbf{e}_5\mathbf{q}-\overline{\mathbf{q}}^T\mathbf{e}_5^T\mathbf{q}'-\overline{\mathbf{q}}^T\mathbf{e}_6\mathbf{q}]dx\,dt. \end{aligned} \tag{10}$$

Matrices $\mathbf{c}_1$, $\mathbf{m}$, $\mathbf{e}_2$, etc., are defined in the Appendix. Note that the prime stands for $\partial/\partial x$. The total Lagrangian for the plate is obtained by assembling all the elements. Now, by considering variation, the governing equations for the plate is obtained as

$$\begin{aligned} &\mathbf{E}_5\mathbf{Q}+\mathbf{E}_4\mathbf{Q}'+\mathbf{E}_3\mathbf{Q}''+\mathbf{E}_2\mathbf{Q}'''+\mathbf{E}_1\mathbf{Q}^{iv} \\ &\quad=\omega^2(\mathbf{M}\mathbf{Q}-\mathbf{C}_1\mathbf{Q}''-\mathbf{C}_2\mathbf{Q}'). \end{aligned} \tag{11}$$

Matrices $\mathbf{C}_1$, $\mathbf{M}$, $\mathbf{E}_1$, $\mathbf{E}_3$, and $\mathbf{E}_5$ are symmetric, whereas the others are skew symmetric. The assembled matrices are defined in the Appendix. It can be observed that the global degrees of freedom vector $\mathbf{Q}$ satisfies a fourth-order homogeneous ordinary differential equation in X (see Ref. 12). A solution to Eq. (11) is assumed in the form

$$\mathbf{Q}=\mathbf{Q}_o e^{ikX}, \tag{12}$$

where $\mathbf{Q}_o$ represents the amplitude vector. By substituting Eq. (12) into Eq. (11), the following system of linear homogeneous equations is obtained:

$$\begin{aligned} &[k^4\mathbf{E}_1-ik^3\mathbf{E}_2-k^2\mathbf{E}_3+ik\mathbf{E}_4+\mathbf{E}_5 \\ &\quad-\omega^2(\mathbf{M}-ik\mathbf{C}_2+k^2\mathbf{C}_1)]\mathbf{Q}_o=\mathbf{0}. \end{aligned} \tag{13}$$

Equation (13) defines a generalized eigenvalue problem, the solutions to which give the eigenvalues ω for given k. Alternatively, the equation can be recast into the form

$$[k^4\mathbf{E}_1-ik^3\mathbf{E}_2-k^2\mathbf{E}_6+ik\mathbf{E}_7+\mathbf{E}_8]\mathbf{Q}_o=\mathbf{0}, \tag{14}$$

where

$$\mathbf{E}_6=\mathbf{E}_3+\omega^2\mathbf{C}_2, \quad \mathbf{E}_7=\mathbf{E}_4+\omega^2\mathbf{C}_1, \quad \mathbf{E}_8=\mathbf{E}_5-\omega^2\mathbf{M}.$$

Equation (14) provides the eigenvalues k for given ω.

B. Nine-node quadratic element formulation

As in Eq. (1), the displacement $\mathbf{u}(x,y,z,t)$ at a point within the element is written in the product form

$$\mathbf{u}(x,y,z,t)=\mathbf{N}(y,z)\mathbf{u}_e(x,t). \tag{15}$$

Using linear elasticity, the strain tensor is derived from the kinematic equation

$$\boldsymbol{\varepsilon}=(\mathbf{L}_x+\mathbf{L}_{yz})\mathbf{u}, \tag{16}$$

where $\mathbf{L}_{yz}$, $\mathbf{L}_x$ are elastic strain operators defined as

$$\mathbf{L}_x=\begin{bmatrix} \partial_x & . & . & . & . & . \\ . & . & . & \partial_x & . & . \\ . & . & . & . & \partial_x & . \end{bmatrix}^T,$$

$$\mathbf{L}_{yz}=\begin{bmatrix} . & . & . & \partial_y & \partial_z & . \\ . & \partial_y & . & . & . & \partial_z \\ . & . & \partial_z & . & . & \partial_y \end{bmatrix}^T.$$

The stress–strain relations at a point inside an element are given by Eqs. (5) and (6). Thus, the strain vector can then be written as

$$\boldsymbol{\varepsilon}=\mathbf{B}_1\mathbf{u}_{e,x}+\mathbf{B}_2\mathbf{u}_e \tag{17}$$

$$\mathbf{B}_1=\begin{bmatrix} \mathbf{N} & . & . & . & . & . \\ . & . & . & \mathbf{N} & . & . \\ . & . & . & . & \mathbf{N} & . \end{bmatrix}^T,$$

$$\mathbf{B}_2=\begin{bmatrix} . & . & . & \mathbf{N}_y & \mathbf{N}_z & . \\ . & \mathbf{N}_y & . & . & . & \mathbf{N}_z \\ . & . & \mathbf{N}_z & . & . & \mathbf{N}_y \end{bmatrix}^T.$$

Here, $\mathbf{N}$ contains the interpolated functions of nine-node quadratic element, and the subscripts denote the derivatives. The Lagrangian L for an element is computed from

$$\begin{aligned} L=\frac{1}{2}\int_t\int_x\int_{-h}^{h}\int_{-b}^{b}(&\dot{\mathbf{u}}_e^T\mathbf{N}^T\boldsymbol{\rho}\mathbf{N}\dot{\mathbf{u}}_e-\overline{\mathbf{u}}_e^T\mathbf{B}_2^T\mathbf{D}\mathbf{B}_2\mathbf{u}_e \\ &-\overline{\mathbf{u}}_e^T\mathbf{B}_2^T\mathbf{D}\mathbf{B}_1\mathbf{u}_{e,x}-\overline{\mathbf{u}}_{e,x}^T\mathbf{B}_1^T\mathbf{D}\mathbf{B}_2\mathbf{u}_e \\ &-\overline{\mathbf{u}}_{e,x}^T\mathbf{B}_1^T\mathbf{D}\mathbf{B}_1\mathbf{u}_{e,x})\,dy\,dz\,dx\,dt. \end{aligned} \tag{18}$$

As before, integrating over the cross-sectional area S of the finite-width plate and using Hamilton's variational principle

$$\delta\int_{t0}^{t1}\left[\frac{1}{2}\int(\dot{\mathbf{u}}_e^T\mathbf{m}^{(e)}\dot{\mathbf{u}}_e-\overline{\mathbf{u}}_e^T\mathbf{k}_{22}^{(e)}\mathbf{u}_e-\overline{\mathbf{u}}_e^T\mathbf{k}_{21}^{(e)}\mathbf{u}_{e,x}-\overline{\mathbf{u}}_{e,x}^T\mathbf{k}_{12}^{(e)}\mathbf{u}_e-\overline{\mathbf{u}}_{e,x}^T\mathbf{k}_{11}^{(e)}\mathbf{u}_{e,x})dx\right]dt=0, \tag{19}$$

the element mass and stiffness matrices are defined as

$$\mathbf{m}^{(e)}=\int\int_S\mathbf{N}^T\boldsymbol{\rho}\,\mathbf{N}\,dy\,dz,$$

$$\mathbf{k}_{22}^{(e)}=\int\int_S\mathbf{B}_2^T\mathbf{D}\mathbf{B}_2\,dy\,dz,$$

$$\mathbf{k}_{21}^{(e)}=(\mathbf{k}_{12}^{(e)})^T=\int\int_S\mathbf{B}_2^T\mathbf{D}\mathbf{B}_1\,dy\,dz,$$

$$\mathbf{k}_{11}^{(e)}=\int\int_S\mathbf{B}_1^T\mathbf{D}\mathbf{B}_1\,dy\,dz.$$

The global mass and stiffness matrices are assembled in the usual manner. We find the governing equation of motion as

$$\mathbf{K}_1\mathbf{U}_{XX}-\mathbf{K}_2\mathbf{U}_X-\mathbf{K}_3\mathbf{U}-\mathbf{M}\ddot{\mathbf{U}}=0, \tag{20}$$

where $\mathbf{U}$ is the assembled nodal displacements $\mathbf{u}_e$, and the stiffness matrices are

$$\mathbf{K}_1=\mathbf{K}_{11},\quad \mathbf{K}_2=\mathbf{K}_{21}-\mathbf{K}_{12},\quad \mathbf{K}_3=\mathbf{K}_{22}. \tag{21}$$

Here, $\mathbf{M}$, $\mathbf{K}_1$, and $\mathbf{K}_3$ are symmetric and $\mathbf{K}_2$ is antisymmetric. As before, assuming harmonic wave propagation in the X direction, we write

$$\mathbf{U}(X,t)=\hat{\mathbf{U}}(k,\omega)e^{i(kX-\omega t)}. \tag{22}$$

Then, Eq. (22) yields the following eigenvalue problem:

$$[k^2\mathbf{K}_1+ik\mathbf{K}_2+\mathbf{K}_3]\hat{\mathbf{U}}=\omega^2[\mathbf{M}]\hat{\mathbf{U}}. \tag{23}$$

Solving the eigensystem (23) determines the eigenvalues ω and the corresponding eigenvectors for given k.

C. Wave propagation in an infinite plate

In this subsection, we describe briefly the equations governing wave propagation in an infinite plate.[12,13] Now, the plate is discretized in the thickness direction and the displacement is written as

$$\mathbf{u}(x,y,z,t)=\mathbf{N}(z)\mathbf{u}_e(x,y,t). \tag{24}$$

Here, $\mathbf{N}$ is the interpolation function across the plate thickness, and $\mathbf{u}_e$ is the nodal displacement. In linear elasticity, the strain may now be expressed as

$$\boldsymbol{\varepsilon}=(\mathbf{L}_x+\mathbf{L}_y+\mathbf{L}_z)\mathbf{u}, \tag{25}$$

where

$$\mathbf{L}_x=\begin{bmatrix} \partial_x & . & . & . & . & . \\ . & . & . & \partial_x & . & . \\ . & . & . & . & \partial_x & . \end{bmatrix},$$

$$\mathbf{L}_y=\begin{bmatrix} . & . & . & \partial_y & . & . \\ . & \partial_y & . & . & . & . \\ . & . & . & . & . & \partial_y \end{bmatrix},$$

$$\mathbf{L}_z=\begin{bmatrix} . & . & . & . & \partial_z & . \\ . & . & . & . & . & \partial_z \\ . & . & \partial_z & . & . & . \end{bmatrix},$$

are spatial derivative operators with respect to x, y, and z, respectively. Substituting Eq. (24) into Eq. (25), the strain field at any point in the plate can be expressed in the form

$$\boldsymbol{\varepsilon}=\mathbf{B}_1\mathbf{u}_{e,x}+\mathbf{B}_2\mathbf{u}_{e,y}+\mathbf{B}_3\mathbf{u}_e, \tag{26}$$

where

$$\mathbf{B}_1=\begin{bmatrix}\mathbf{N} & . & . & . & . & .\\ . & . & . & \mathbf{N} & . & .\\ . & . & . & . & \mathbf{N} & .\end{bmatrix},$$

$$\mathbf{B}_2=\begin{bmatrix}. & . & . & \mathbf{N} & . & .\\ . & \mathbf{N} & . & . & . & .\\ . & . & . & . & . & \mathbf{N}\end{bmatrix},$$

$$\mathbf{B}_3=\begin{bmatrix}. & . & . & . & \mathbf{N}_z & .\\ . & . & . & . & . & \mathbf{N}_z\\ . & . & \mathbf{N}_z & . & . & .\end{bmatrix},$$

are interpolation matrices. The Hamiltonian is given in terms of the element matrices as follows:

$$\delta L=\frac{1}{2}\,\delta\int_t\int_x\int_y(\dot{\mathbf{u}}_e^T\mathbf{m}\dot{\mathbf{u}}_e-\overline{\mathbf{u}}_{e,x}^T\mathbf{k}_{11}\mathbf{u}_{e,x}-\overline{\mathbf{u}}_{e,x}^T\mathbf{k}_{12}\mathbf{u}_{e,y}-\overline{\mathbf{u}}_{e,x}^T\mathbf{k}_{13}\mathbf{u}_e-\overline{\mathbf{u}}_{e,y}^T\mathbf{k}_{21}\mathbf{u}_{e,x}-\overline{\mathbf{u}}_{e,y}^T\mathbf{k}_{22}\mathbf{u}_{e,y}-\overline{\mathbf{u}}_{e,y}^T\mathbf{k}_{23}\mathbf{u}_e-\overline{\mathbf{u}}_e^T\mathbf{k}_{31}\mathbf{u}_{e,x}-\overline{\mathbf{u}}_e^T\mathbf{k}_{32}\mathbf{u}_{e,y}-\overline{\mathbf{u}}_e^T\mathbf{k}_{33}\mathbf{u}_e)\,dy\,dx\,dt, \quad (27)$$

where

$$\mathbf{m}=\int_{-h}^{h}\mathbf{N}^T\boldsymbol{\rho}\mathbf{N}\,dz, \qquad \mathbf{k}_{ij}=\mathbf{k}_{ji}^T=\int_{-h}^{h}\mathbf{B}_i^T\mathbf{D}\mathbf{B}_j\,dz.$$

The global mass $\mathbf{M}$ and stiffness matrices $\mathbf{K}_{ij}$ are assembled in the usual manner. The equation of motion takes the form

$$\mathbf{K}_{11}\mathbf{U}_{,XX}+\mathbf{K}_{22}\mathbf{U}_{,YY}+(\mathbf{K}_{12}+\mathbf{K}_{21})\mathbf{U}_{,XY}+(\mathbf{K}_{13}-\mathbf{K}_{31})\mathbf{U}_{,X}+(\mathbf{K}_{23}-\mathbf{K}_{32})\mathbf{U}_{,Y}+\mathbf{K}_{33}\mathbf{U}-\mathbf{M}\ddot{\mathbf{U}}=0. \quad (28)$$

For a wave propagating in a direction lying in the plane of XY, the displacement can be written as

$$\mathbf{U}(X,Y,t)=\hat{\mathbf{U}}(k_x,k_y,\omega)e^{i(k_xX+k_yY-\omega t)}. \quad (29)$$

Substituting Eq. (29) in Eq. (28), the dispersion relation is found to be

$$\mathrm{Det}[k_x^2\mathbf{K}_{11}+k_y^2\mathbf{K}_{22}+k_xk_y(\mathbf{K}_{12}+\mathbf{K}_{21})+ik_x(\mathbf{K}_{13}-\mathbf{K}_{31})+ik_y(\mathbf{K}_{23}-\mathbf{K}_{32})+\mathbf{K}_{33}-\omega^2\mathbf{M}]=0. \quad (30)$$

Here, k_x and k_y are the wave numbers along the X and Y directions. In previous studies of finite width plate, Morse,[20] Mindlin and Fox,[21] and Mindlin[22] considered special values of k_y and ω so that the traction free-boundary conditions were satisfied at the edges $y=\pm b$. Morse[20] noted that for extensional motion the shear stresses τ_{yz} and τ_{yx} were small for large B/H and neglecting those, the condition $\sigma_{yy}=0$ at $y=\pm b$ leads to $\kappa k_y=n\pi/2b$ $(n=0,1,3,\ldots)$. Here, $\kappa=C_L/C_S$. Now, for flexural motion symmetric about the z axis, the condition $\tau_{yz}=0$ at $y=\pm b$ leads to $k_y=m\pi/2b$ $(m=0,2,4,\ldots)$. It is interesting to compare the results for the infinite plate for these values of k_y with the exact solution for the bounded plate. This is discussed in the next section.

III. NUMERICAL RESULTS AND DISCUSSION

Results presented in the following include propagation in homogeneous isotropic nickel (Ni) plate of different width-to-thickness ratios ($B/H=2$ and 8) and for symmetric layered plates composed of either a core of Ni coated by superconducting YBCO layers on both sides or a YBCO layer sandwiched between two Ni layers. Results for $B/H=8$ are compared with the exact solutions[8,9] for $B/H\to\infty$. It is shown that the dispersion characteristics of guided waves in the case of $B/H=8$ differ significantly from those for an infinite plate.

In order to validate the computational strategies used in this paper, first we considered propagation in a homogeneous, isotropic rectangular plate having $B/H=2$ and Poisson's ratio, $\nu=0.3$. Comparison of the present results with those of Taweel *et al.*[11] (Fig. 2) show excellent agreement. In this case, advantage is taken of the doubly symmetric geometry. Thus, only one quarter of the cross section needs be discretized. As noted in Ref. 11, there are four independent waveforms : (a) extension (sym/sym); (b) torsion (antisym/antisym); (c) flexure about the Y axis (antisym/sym); and (d) flexure about the Z axis (sym/antisym). Both types of elements were used: six-node elements (5×5 with 66 nodes) and nine-node elements (5×5 with 121 nodes), as discussed in the previous section. Both gave results that agreed with one another and with those presented by Taweel *et al.*[11] Here, $\Omega=\omega H/C_S$ and $\mathrm{K}=kH$. Next, comparison is made with the ultrasonic data obtained by Morse[20] for extensional modes in a brass rod with $B=6.36$ mm and $H=3.18$ mm. Figure 3 shows the comparison of the phase velocity of the first two extensional modes vs frequency. Model results obtained using nine-node elements (5×5 with 121 nodes) are seen to agree well with experiments.

We now turn to the comparison of the results using six-node and nine-node elements for dispersion of extensional and flexural (about the Y axis) waves in a nickel plate having $B/H=8$. Again, because of symmetry only one quarter of the plate need be considered. Properties of nickel are given in Table I. In Fig. 4, normalized frequency is $\overline{\Omega}$, where $\overline{\Omega}=\omega H/(2\pi(1.1)\mathrm{C_S})$, and normalized phase velocity is $\overline{\mathrm{C}}$ $(=\mathrm{C}/\mathrm{C_S})$. Here, $\mathrm{C_S}$ is the shear wave velocity in nickel and the thickness of the plate is H $(=110\ \mu\mathrm{m})$. For the six-node elements, one quarter of the plate was discretized into 5×11 elements (138 nodes), whereas there were 5×5 nine-node elements (121 nodes). It is seen that there is excellent agreement in the low-frequency range considered in this paper. Also shown in this figure are the exact dispersion curves for the extensional and flexural modes of an infinite plate. These have several interesting features: (1) dispersion curve for the lowest flexural mode [Fig. 4(b)] is almost the same as in the infinite-width plate, showing that the flexural mode is not sensitive to the width B (see the discussion below); (2) the phase velocity of the lowest extensional mode in the infinite plate is higher than that in the finite-width plate. Note that as $\omega\to0$, the former tends to the plate velocity $\mathrm{C_P}$

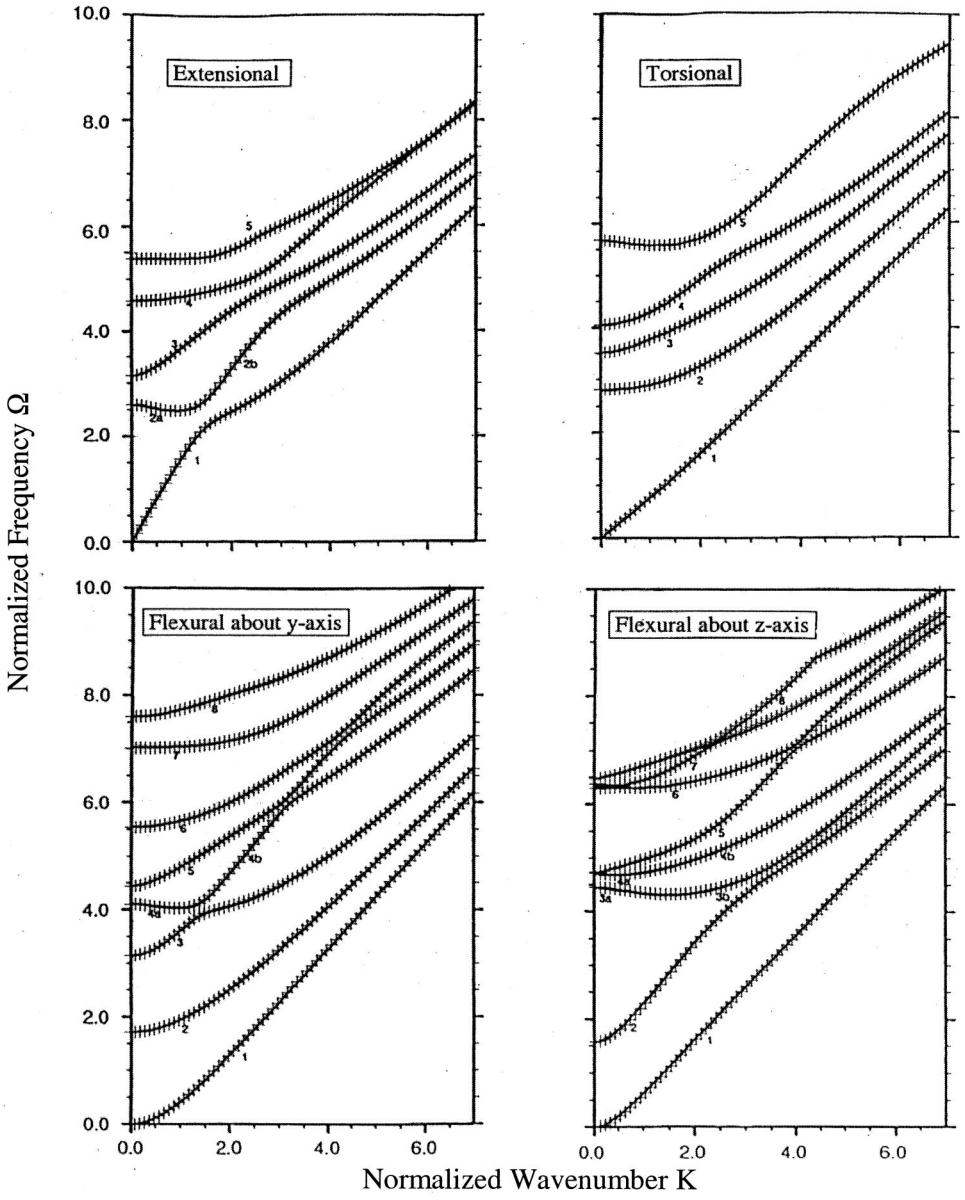


FIG. 2. Comparison of the results obtained by two methods and with those of Ref. 11. Results using nine-node element (+) and those from Ref. 11 (—) are shown. Those obtained by using six-node elements lie on the solid line and are not shown.

$=1.693C_S$, and the latter tends to the rod velocity C_R $=1.614C_S$, respectively, as shown in Fig. 5(a) for Ni plate. Note that C_P is higher than C_R because of the constrains of the finite width in a rod; (3) cutoff frequencies for the guided modes in the finite-width plate are lowered more and more as the width increases and there are more propagating modes at any frequency with increasing width (see the discussion below); (4) phase velocities of the lowest flexural and extensional modes tend to the Rayleigh wave velocity in nickel as frequency becomes large (as in the case of the infinite plate); (5) phase velocities of the low extensional modes drop rapidly as the frequency increases, tend to a plateau at the lowest extensional wave velocity in an infinite plate, and then decrease gradually to the Rayleigh wave velocity in nickel; (6) phase velocities of the low flexural modes decrease rapidly with increasing frequency tending to a plateau at the Rayleigh wave velocity (note that phase velocity of the second flexural mode drops below the Rayleigh wave velocity and then increases to that value).

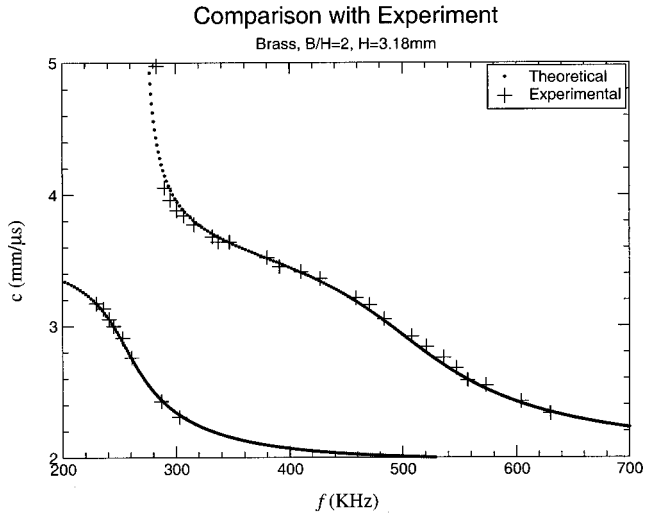
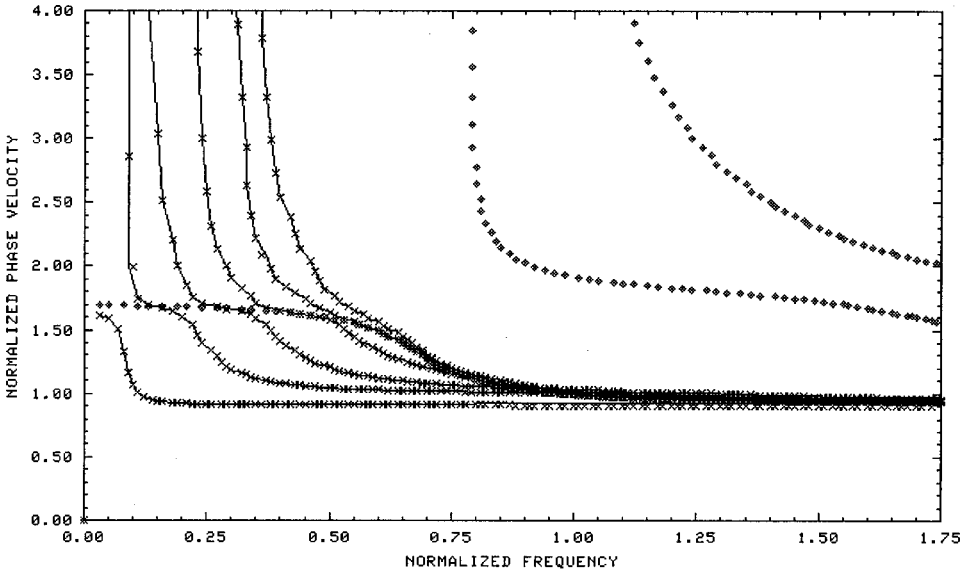


FIG. 3. Comparison of the results obtained using nine-node elements with the experiments by Morse (Ref. 20) for the first two branches of the extensional mode.

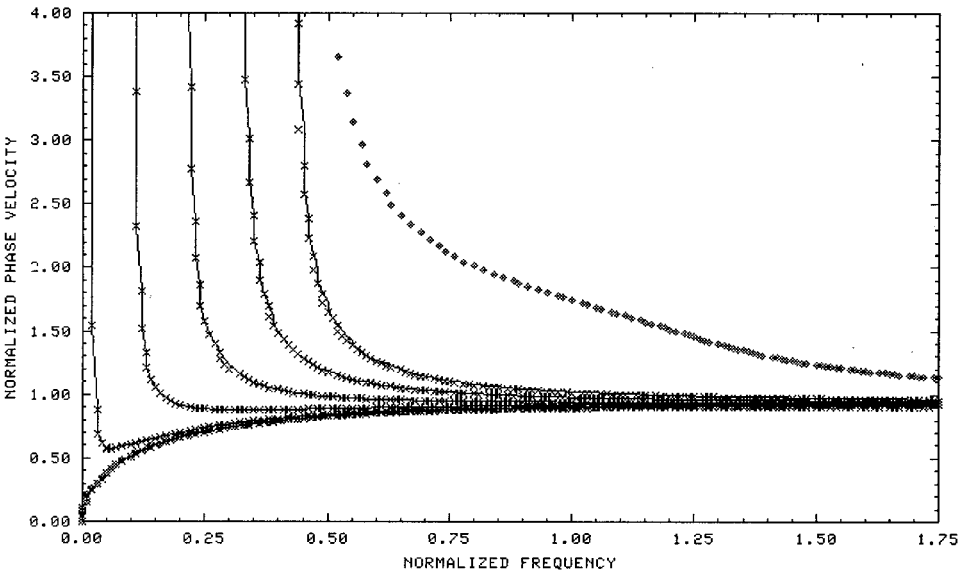
TABLE I. Material properties of Nickel and YBCO.

Material	Density (g/cm^3)	D_{11} (GPa)	D_{12} (GPa)	D_{13} (GPa)	D_{22} (GPa)	D_{23} (GPa)	D_{33} (GPa)	D_{44} (GPa)	D_{55} (GPa)	D_{66} (GPa)
Nickel	8.910	298.95	129.53	129.53	298.95	129.53	298.95	84.71	84.71	84.71
YBCO	6.330	268.00	132.00	95.00	231.00	71.00	186.00	95.00	49.00	37.00

Figure 5 shows the comparison of the dispersion of extensional and flexural modes in a plate having the width $B=8H$ with the lowest branches of the corresponding modes in an infinite plate for particular values of k_y discussed in the previous section (II C). Equation (30) relates the wave number k_x with frequency ω for different values of k_y. If $k_y=0$, this is the Lamb wave dispersion relation for propagation along the x axis. Thus, the lowest extensional and flexural branches propagate at all frequencies. If $k_y \neq 0$, then there are cutoff frequencies for propagating modes, including the lowest branches. These cutoff frequencies increase as k_y increases. Furthermore, for any k_y there is an infinite number of propagating branches. In Figs. 5(a) and (b), we have plotted the lowest extensional and flexural branches, respectively, for a particular set of k_y values for which σ_{yy} or τ_{yz} vanishes at $y=\pm b$, as appropriate. Also shown are the re-

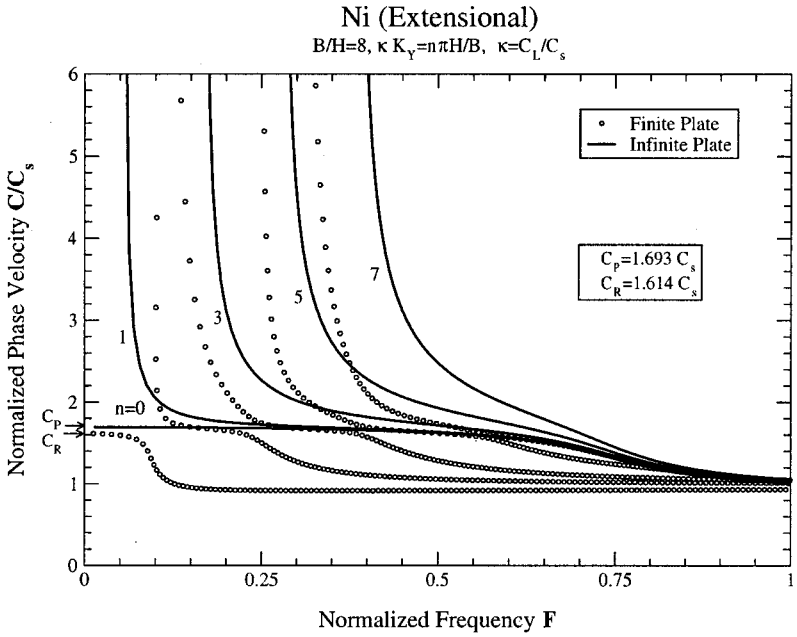


(a)

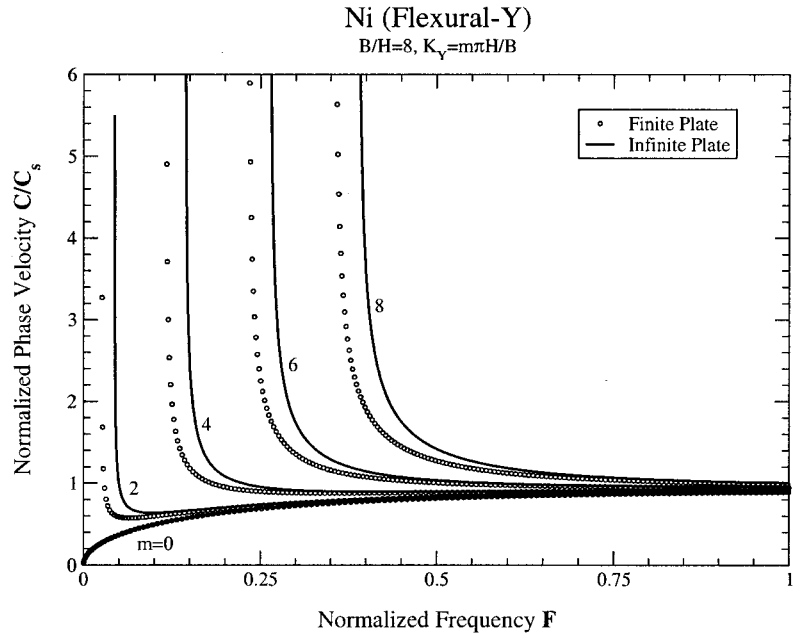


(b)

FIG. 4. Comparison of the results obtained using six-node elements (—) with those using nine-node elements ($\times$) for the extensional and flexural (flexure about Y axis) motion in a nickel plate of $B/H=8$. Also shown are the corresponding results for an infinite plate ($\blacklozenge$) having the same thickness $H=110\ \mu$m.



(a)



(b)

FIG. 5. Normalized phase velocity versus normalized frequency for (a) extensional and (b) flexural motion about the Y axis in Ni plate of finite width ($B/H=8$), and the width modes of infinite plate.

sults for a few of the propagating modes in a finite-width plate. It is seen in Fig. 5(a) that the extensional modes in an infinite plate behave quite differently than in the finite-width plate. On the other hand, Fig. 5(b) shows remarkable similarity between the flexural modes in the two cases. Thus, the width does not have strong influence on these modes.

Guided waves in a nickel plate with superconducting (YBCO) layers were studied[8,9] before. It was assumed that the plate was infinite in planar dimensions. In practice, however, these plates (tapes) are of finite width. In order to estimate the effect of finite width of the tape on the dispersion, we analyzed the guided wave modes in two different kinds of layered tapes, as in Refs. 8 and 9. In one, the tape was taken to have a core of nickel layer 100 μm in thickness coated symmetrically by 5-μm YBCO layers on top and bottom. Properties of the YBCO layer are given in Table I. The second tape was taken to be a sandwich having a thin (10-μm) YBCO layer between two 50-μm-thick nickel layers. In both cases, the thickness of the tapes was 110 μm. Although results were obtained for $B/H=2$ and 8, only those for the last case will be presented here for brevity. Cutoff frequencies are lower in the Ni/YBCO/Ni plate than in the YBCO/Ni/YBCO plate. Table II lists the cutoff frequencies of the first few modes for extensional and flexural waves in plates made of nickel, YBCO, Ni/YBCO/Ni, and YBCO/Ni/YBCO. Also shown are results for the infinite plate in these cases.

Figure 6 shows the extensional and flexural modes for Ni/YBCO/Ni plates. The plots are for the normalized phase

TABLE II. Normalized cutoff frequencies for various plates when $B/H=8$. Note: Cutoff frequencies for analogous infinitely wide tape are shown as $(\ )^+$ for SH wave, $(\ )^-$ for plane strain, and $(\ )^*$ for SH as well as plane strain waves.

Nickel tape		YBCO tape		Ni/YBCO/Ni tape		YBCO/Ni/YBCO tape	
Extension	Flexure	Extension	Flexure	Extension	Flexure	Extension	Flexure
0.000 02	0.000 05	0.000 00	0.000 07	0.000 03	0.000 06	0.000 00	0.000 05
$(0.000\ 00)^*$	$(0.000\ 00)^-$	$(0.000\ 00)^*$	$(0.000\ 00)^-$	$(0.000\ 00)^*$	$(0.000\ 00)^-$	$(0.000\ 00)^*$	$(0.000\ 00)^-$
0.603 48	0.147 09	0.654 57	0.155 67	0.605 06	0.147 31	0.605 83	0.145 69
0.714 01	0.660 73	0.896 86	0.643 56	0.727 54	0.648 53	0.727 54	0.657 46
1.428 36	1.325 81	1.793 84	1.208 09	1.455 03	1.280 07	1.455 03	1.322 74
1.790 58	2.029 63	1.946 41	1.765 74	1.796 01	1.935 69	1.797 80	2.025 82
			$(2.238\ 85)^+$				
2.144 67	2.704 13	2.691 54	2.249 27	2.182 82	2.533 28	2.182 82	2.702 75
					$(2.565\ 43)^+$		$(2.931\ 54)^+$
2.866 71	2.855 99	3.150 04	2.523 02	2.900 68	2.682 58	2.907 00	2.932 09
	$(2.855\ 99)^*$				$(2.682\ 61)^-$		$(2.932\ 10)^-$
2.893 29	2.943 89	3.591 40	2.576 31	2.911 95	2.776 09	2.911 95	3.024 28
			$(2.576\ 45)^-$				
3.600 75	3.087 40	3.925 22	2.727 95	3.644 39	2.834 89	3.644 39	3.131 61
3.756 08	3.193 26	4.019 23	2.893 03	3.736 71	3.039 44	3.785 56	3.285 31
4.216 35	3.402 92	4.459 09	3.135 31	4.112 02	3.168 90	4.317 40	3.420 52
		$(4.477\ 70)^+$					
4.305 37	3.571 60	4.480 62	3.139 30	4.226 03	3.434 13	4.361 38	3.679 60
4.354 74	3.761 48	4.496 01	3.589 86	4.383 08	3.505 76	4.383 08	3.760 68
4.750 91	4.046 57	4.638 91	3.725 82	4.621 04	3.921 85	4.776 00	4.170 38
5.013 29	4.177 77	4.780 62	3.947 16	4.826 53	3.933 22	5.131 88	4.189 37
5.119 32	4.547 86	4.997 55	4.224 18	4.922 53	4.232 88	5.146 78	4.496 90
		$(5.019\ 73)^-$					
5.131 89	4.595 88	5.099 72	4.419 90	5.041 64	4.474 74	5.200 66	4.729 84
				$(5.090\ 64)^-$			
5.257 51	5.016 66	5.152 62	4.789 30	5.119 96	4.777 70	5.299 75	5.017 07
		$(5.152\ 90)^-$					
5.334 34	5.207 73	5.218 58	5.017 20	5.131 88	4.978 71	5.450 84	5.253 93
$(5.365\ 24)^-$						$(5.508\ 44)^-$	
5.445 17	5.654 44	5.230 09	5.181 84	5.238 83	5.075 70	5.541 20	5.340 29

velocity (C/C_S) vs normalized frequency F $(=\omega\bar{H}/2\pi C_S)$, where $\bar{H}$ is total thickness (100 μm) of the nickel layer and $B/\bar{H}=8$. The effect of the interface YBCO layer is quite evident in the higher modes. In order to bring out the effect of the anisotropy of the YBCO layer, results are presented for two cases when the symmetry axis (1 axis) coincides with the X axis and when the former is rotated by 45°. In the latter case, there are two independent modes, one symmetric about the Y axis and the other antisymmetric. Comparison of the results shows that symmetric motion is somewhat affected by the orientation of the symmetry axis, but the antisymmetric motion undergoes hardly any changes. Figure 7 shows close-up views of the region where the branches shown in Fig. 6(a) cluster at around $F=0.75-0.9$. The mode coupling caused by anisotropy is evident. Figure 7(a) shows two independent families of modes, one extensional and the other bending about the z axis. Branches of the same family come very close but do not cross, but those of one family may cross ones belonging to the other. On the other hand, Fig. 7(b) shows that the two motions are now coupled and the branches do not cross. This phenomenon is strictly an effect of the anisotropy and was discussed also in Refs. 8 and 9.

Modes shapes associated with the symmetric motion about the Y axis for the Ni/YBCO/Ni plate when $B=800$ μm and $\bar{H}=100$ μm are shown in Fig. 8. There are two independent modes, extensional (L) and flexural about the Z axis (B_z), in this case. Dispersion curves are shown in Fig. 6(a). Mode shapes shown are for the first three extensional and flexural waves and for two different nondimensional wave numbers (K). Motion in only the top half of the plate is shown. It is interesting to note that at low K the displacement has very little Z component and has larger Y component than the X component. On the other hand, when K is large, the Y component is negligible except near the edges and the displacement has larger Z component than the X component. Also, the X component has a node close to the free surface. Thus, this motion can be identified with the Rayleigh wave. The figure shows clearly the effect of the finite width by the appearance of nodes along the Y axis, showing the standing waves due to the reflections from the edges.

Figure 9 shows the mode shapes associated with the antisymmetric motion about the Y axis. There are two independent waveforms: flexure about the Y axis and the torsion. Now, the motion is predominantly in the Z direction and has

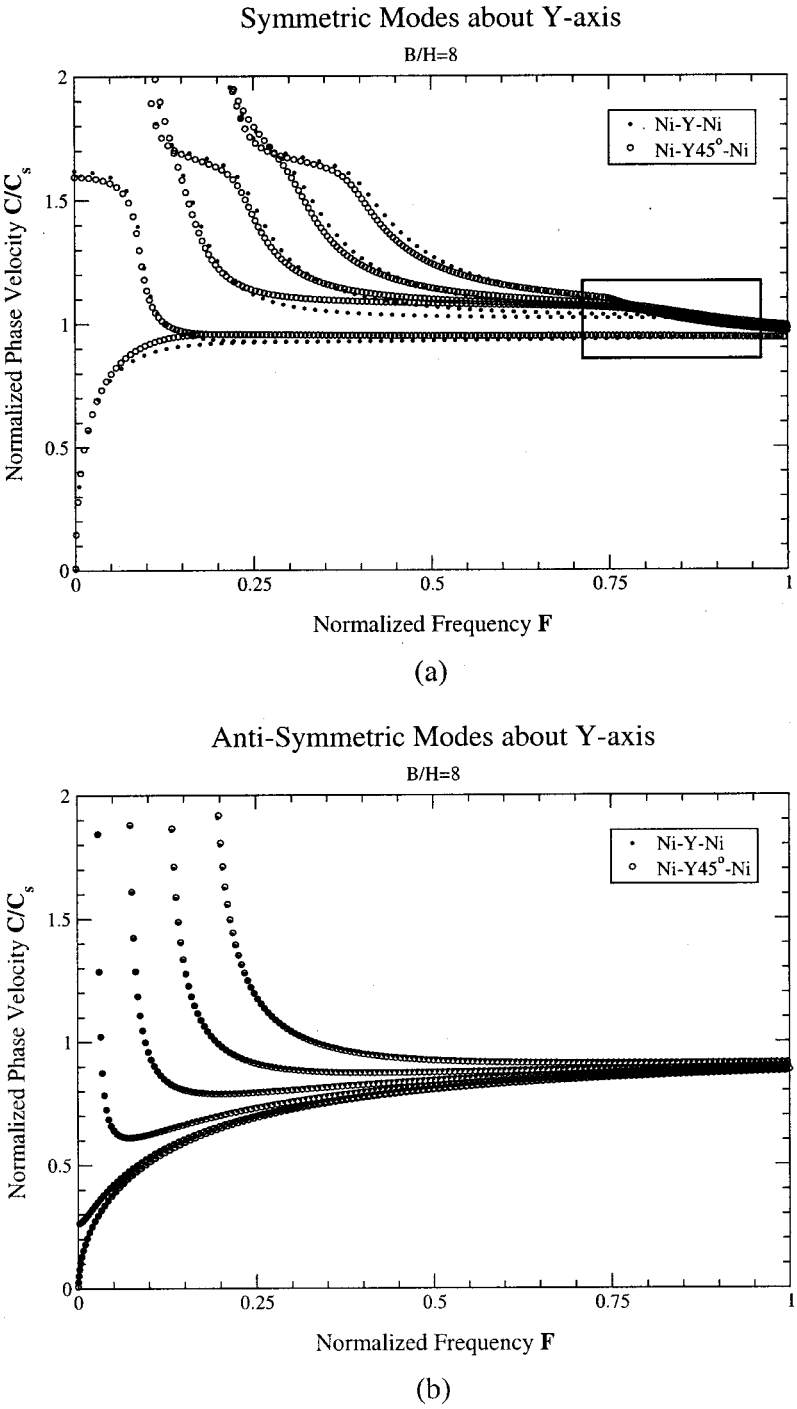
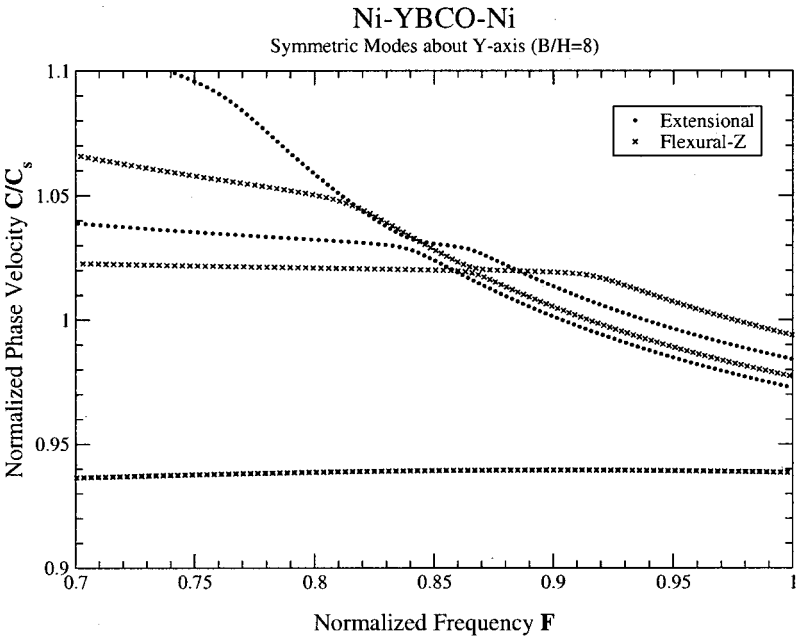


FIG. 6. Normalized phase velocity versus normalized frequency for (a) symmetric and (b) antisymmetric motion about the Y-axis in Ni/YBCO/Ni and Ni/YBCO-45°/Ni plates of finite width ($B/\overline{H}=8$).

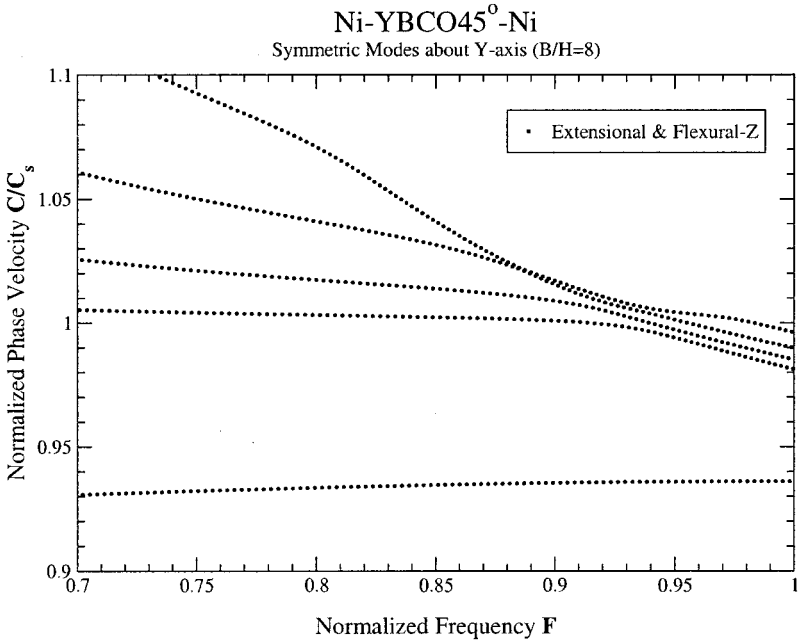
small X and Y components. At high values of K, it is found that the Y component is confined to the edges and the surface Rayleigh wave emerges.

IV. CONCLUSION

A semianalytical finite element method has been used to model ultrasonic guided waves in a plate of rectangular cross section. Two different types of planar finite elements are used to represent the cross section. In one, six-node elements that require continuity of traction and displacement components in the thickness direction and only continuity of displacement in the width direction are used. Thus, the nodes have six degrees of freedom. In the other, nine-node quadrilateral elements requiring displacement continuity at the nodes are employed. It is found that both methods give results that agree with one another and with previously reported model and (limited) experimental results. In this study, detailed investigation of the effect of changes in the width-to-thickness ratios reveals that the finite width of the plate modifies significantly the dispersion characteristics of the modes even when the width is much larger than the



(a)



(b)

FIG. 7. Close-up view of the symmetric motion about the *Y* axis in (a) Ni/YBCO/Ni; (b) Ni/YBCO-45°/Ni plates for $0.7 \leqslant F \leqslant 1.0$. Note that the extension and flexure are uncoupled in (a) and coupled in (b).

thickness. It is found that as the width increases (thickness being constant) the cutoff frequencies of the modes decrease and there are more modes for a given frequency than for an infinite plate. Modes are seen to bunch together around those of an infinite plate. The lowest flexural mode (flexure about the width axis) is not sensitive to the changes in the width and is almost the same as in the infinite plate. Investigation of the mode shapes in representative cases of interest shows that the finite width results in standing waves along the width due to reflections from the two edges. The phase velocity of the lowest extensional mode of a finite-width plate is lower than that of the infinite plate. The phase velocity of this mode drops more rapidly for a finite-width plate than for the infinite plate. The effect of thin, anisotropic superconducting layers has also been studied. This is found to be very small, but measurable, as seen by the changes in the cutoff frequencies. These are found to be dependent upon whether the layers are outside (such as coating) a thick core or are sandwiched between two thick layers. Furthermore, mode coupling due to misalignment of the material symmetry axes from the geometrical symmetry axes is evidenced as the frequency increases.

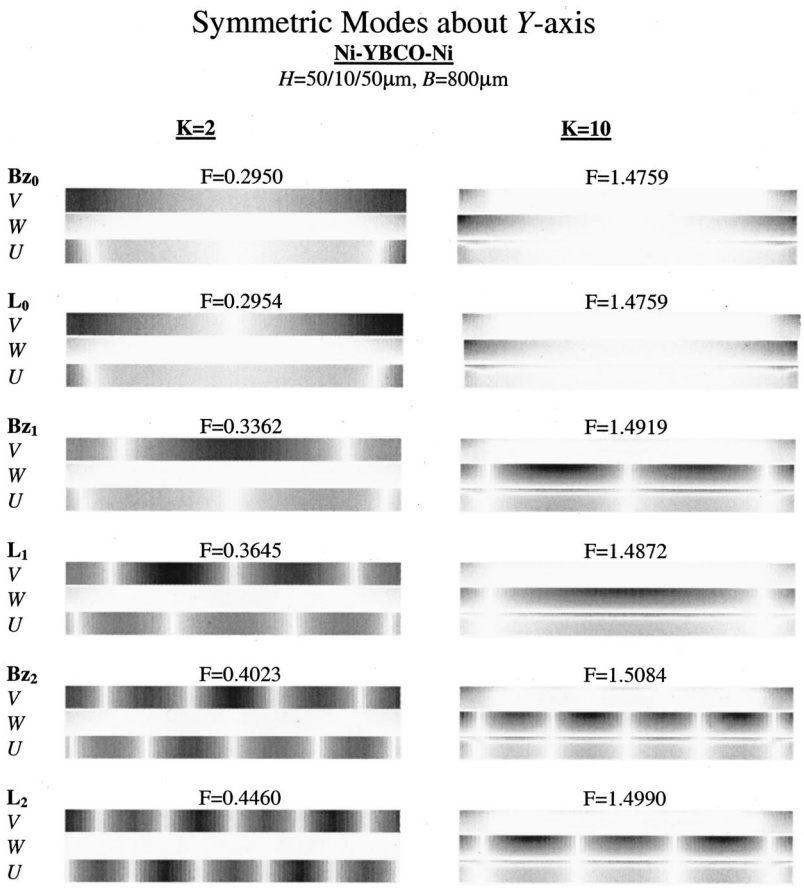


FIG. 8. Mode shapes associated the first three extensional (L) and flexural (about the *Z* axis) modes in Ni/YBCO/Ni plate for the normalized wave numbers K = 2 and K = 10.

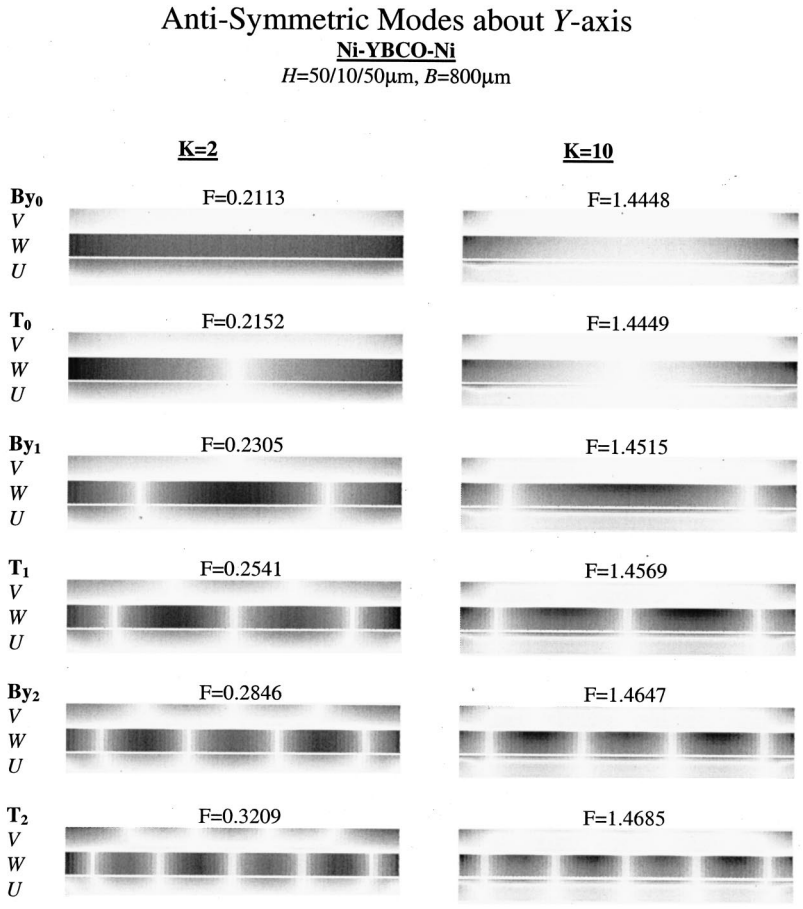


FIG. 9. Mode shapes associated with the first three flexural (about the *Y* axis) and torsional (T) modes in Ni/YBCO/Ni plate for the normalized wave numbers K = 2 and K = 10.

ACKNOWLEDGMENTS

The work of O.M.M., S.K.D., and A.J.N. was supported by a grant from the Department of Energy, Division of Materials Science and Engineering, Office of Basic Energy Sciences (Grant No. DE-FG03-97ER14738). Y.D. and A.H.S. received support from the Natural Science and Engineering Research Council of Canada. S.K.D. gratefully acknowledges a Faculty Fellowship Award from the University of Colorado at Boulder during 1999–2000 and the hospitality and support provided by the University of Queensland (Departments of Civil and Mechanical Engineering), Brisbane and the Chalmers University of Technology (Department of Mechanics), Gothenburg.

APPENDIX: ELEMENT MATRICES

Matrices appearing in Eq. (10) are defined below,

$$\mathbf{c}_1=\int_{-h}^{h}\int_{-b}^{b}\rho\mathbf{B}_1^T\mathbf{B}_1\,dy\,dz,$$

$$\mathbf{c}_2=\int_{-h}^{h}\int_{-b}^{b}\rho\mathbf{B}_1^T\mathbf{B}_2\,dy\,dz, \tag{A1}$$

$$\mathbf{m}=\int_{-h}^{h}\int_{-b}^{b}\rho\mathbf{B}_2^T\mathbf{B}_2\,dy\,dz,$$

$$\mathbf{e}_1=\int_{-h}^{h}\int_{-b}^{b}\mathbf{d}^T\mathbf{D}\mathbf{d}\,dy\,dz,$$

$$\mathbf{e}_2=\int_{-h}^{h}\int_{-b}^{b}\mathbf{d}^T\mathbf{D}\mathbf{b}\,dy\,dz,$$

$$\mathbf{e}_3=\int_{-h}^{h}\int_{-b}^{b}\mathbf{d}^T\mathbf{D}\mathbf{a}\,dy\,dz,$$

$$\mathbf{e}_4=\int_{-h}^{h}\int_{-b}^{b}\mathbf{b}^T\mathbf{D}\mathbf{b}\,dy\,dz, \tag{A2}$$

$$\mathbf{e}_5=\int_{-h}^{h}\int_{-b}^{b}\mathbf{b}^T\mathbf{D}\mathbf{a}\,dy\,dz,$$

$$\mathbf{e}_6=\int_{-h}^{h}\int_{-b}^{b}\mathbf{a}^T\mathbf{D}\mathbf{a}\,dy\,dz.$$

Matrices **d**, **b**, and **a** appearing in equation set (A2) are computed from the strain–displacement relations as

$$\mathbf{l}=[l_1\quad l_2\quad l_3\quad l_4\quad l_5\quad l_6],\quad l=\mathbf{d},\quad \mathbf{b},\ \text{and}\ \mathbf{a}, \tag{A3}$$

where

$$\mathbf{d}_i=\begin{bmatrix} . & . & -g_i f_p & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ -g_i f_p\dfrac{D_{13}}{D_{33}} & -g_i f_p\dfrac{D_{34}}{D_{33}} & . & . & . & . \\ . & . & . & . & . & . \end{bmatrix},$$
$$i=1,2,\ldots,6, \tag{A4}$$

$$\mathbf{b}_i=\begin{bmatrix} g_i f_p & . & . & g_i f_p\dfrac{D_{66}}{\Delta} & -g_i f_p\dfrac{D_{56}}{\Delta} & . \\ . & . & . & . & . & . \\ -g_i f_p\dfrac{D_{13}}{D_{33}} & -g_i f_p^*\dfrac{D_{34}}{D_{33}} & . & . & . & . \\ . & g_i f_q & (h_i-g_i^*)f_p & -g_i f_p\dfrac{D_{56}}{\Delta} & g_i f_p\dfrac{D_{55}}{\Delta} & . \\ h_i f_p\dfrac{D_{34}}{D_{33}} & h_i f_p\dfrac{D_{23}}{D_{33}} & g_i(f_q-f_p^*) & . & . & \dfrac{g_i f_p}{D_{33}} \\ -g_i^* f_p\dfrac{D_{13}}{D_{33}} & -g_i^* f_p\dfrac{D_{34}}{D_{33}} & . & . & . & . \end{bmatrix}, \tag{A5}$$

$$\mathbf{a}_i=\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & g_i^* f_q & h_i^* f_p & -g_i^* f_p \dfrac{D_{56}}{\Delta} & g_i^* f_p \dfrac{D_{55}}{\Delta} & \cdot \\ h_i f_p^* \dfrac{D_{34}}{D_{33}} & h_i f_p^* \dfrac{D_{23}}{D_{33}} & g_i f_q^* & \cdot & \cdot & \dfrac{g_i f_p^*}{D_{33}} \\ g_i^* f_q & \cdot & \cdot & g_i^* f_p \dfrac{D_{66}}{\Delta} & -g_i^* f_p \dfrac{D_{56}}{\Delta} & \cdot \\ g_i f_q^* & \cdot & \cdot & g_i f_p^* \dfrac{D_{66}}{\Delta} & -g_i f_p^* \dfrac{D_{56}}{\Delta} & \cdot \\ h_i^* f_p \dfrac{D_{34}}{D_{33}} & h_i^* f_p \dfrac{D_{23}}{D_{33}}+g_i f_q^* & g_i^* f_q+h_i f_p^* & -g_i f_p^* \dfrac{D_{56}}{\Delta} & g_i f_p^* \dfrac{D_{55}}{\Delta} & \cdot \end{bmatrix}, \tag{A6}$$

with $p=3$, $q=1$ for $i=1,2,3$ and $p=4$, $q=2$ for $i=4,5,6$. Further,

$$\begin{aligned} &f_1^*=\frac{3}{4h}(-1+\xi^2), \quad f_2^*=\frac{3}{4h}(1-\xi^2), \\ &f_3^*=\tfrac{1}{4}(-1-2\xi+3\xi^2), \quad f_4^*=\tfrac{1}{4}(-1+2\xi+3\xi^2), \\ &g_1^*=g_4^*=\frac{1}{b}\left(\eta-\frac{1}{2}\right), \\ &g_2^*=g_5^*=-\frac{2\eta}{b}, \quad g_3^*=g_6^*=\frac{1}{b}\left(\eta+\frac{1}{2}\right). \end{aligned} \tag{A7}$$

$$-2h_1^*=h_2^*=-2h_3^*=-2h_4^*=h_5^*=-2h_6^*=\frac{2}{b^2}, \tag{A8}$$

and

$$\begin{aligned} &\mathbf{E}_5=\cup\mathbf{e}_6 \\ &\mathbf{E}_4=\cup(\mathbf{e}_5^T-\mathbf{e}_5) \\ &\mathbf{E}_3=\cup(\mathbf{e}_3+\mathbf{e}_3^T-\mathbf{e}_4) \\ &\mathbf{E}_2=\cup(\mathbf{e}_2^T-\mathbf{e}_2) \\ &\mathbf{E}_1=\cup\mathbf{e}_1 \\ &\mathbf{M}=\cup\mathbf{m} \\ &\mathbf{C}_1=\cup\mathbf{c}_1 \\ &\mathbf{C}_2=\cup(\mathbf{c}_2-\mathbf{c}_2^T), \end{aligned} \tag{A9}$$

where $\cup$ stands for union.

[1] A. H. Nayfeh, *Wave Propagation in Layered Anisotropic Media with Applications to Composite Materials* (Elsevier, Amsterdam, 1995).
[2] S. K. Datta, "Wave Propagation in Composite Plates and Shells," in *Comprehensive Composite Materials*, edited by T.-W. Chou (Elsevier, Amsterdam, 2000), Vol. 1, Chap. 18.
[3] J. L. Rose, *Ultrasonic Waves in Solid Media* (Cambridge University Press, Cambridge, UK, 1999).
[4] D. E. Chimenti, "Guided Waves in Plates and Their Use in Material Characterization," Appl. Mech. Rev. **50**, 247–284 (1997).
[5] Y. SeGi, K. W. Kim, M. A. Stroscio, G. J. Iafrate, and A. Ballato, "Electron-Acoustic-Phonon Scattering Rates in Rectangular Quantum Wires," Phys. Rev. B **50**, 1733–1738 (1994).
[6] S. K. Datta, M. L. Dunn, R. Sesselman, and J. Niklasson, "Mechanical Behavior and Ultrasonic Characterization of Ductile/Brittle Layered Material System," *Proceedings of the 17th Symposium on Energy Eng. Sci.*, edited by B. Armaly (Argonne National Laboratory, Argonne, IL, 1999), 82–89.
[7] A. J. Niklasson and S. K. Datta, "On the Modeling of Guided Waves in Plates With Superconducting Layers," in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (American Institute of Physics, Melville, NY, 2000), Vol. 19.
[8] A. J. Niklasson, S. K. Datta, and M. L. Dunn, "On Approximation of Guided Waves in Plates with Thin Anisotropic Coatings by Means of Effective Boundary Conditions," J. Acoust. Soc. Am. **108**, 924–933 (2000).
[9] A. J. Niklasson, S. K. Datta, and M. L. Dunn, "On Ultrasonic Guided Waves in a Sandwich Plate with a Thin Anisotropic Middle Layer," J. Acoust. Soc. Am. **108**, 2005–2011 (2000).
[10] B. Aalami, "Waves in Prismatic Guides of Arbitrary Cross Section," J. Appl. Mech. **40**, 1067–1072 (1973).
[11] H. Taweel, S. B. Dong, and M. Kazic, "Wave Reflection from the Free End of a Cylinder with Arbitrary Cross-section," Int. J. Solids Struct. **37**, 1701–1726 (2000).
[12] S. K. Datta, A. H. Shah, R. L. Bratton, and T. Chakraborty, "Wave Propagation in Laminated Composite Plates," J. Acoust. Soc. Am. **83**, 2020–2026 (1988).
[13] W. Karunasena, A. H. Shah, and S. K. Datta, "Wave Propagation in a Multilayered Cross-ply Composite Plate," J. Appl. Mech. **58**, 1028–1032 (1991).
[14] P. J. Shull, D. E. Chimenti, and S. K. Datta, "Elastic Guided Waves and the Floquet Concept in Periodically Layered Plates," J. Acoust. Soc. Am. **95**, 99–108 (1994).
[15] S. K. Datta, A. H. Shah, and W. Karunasena, "Ultrasonic Waves and Material Defect Characterization in Composite Plates," Mech. Compos. Mater. Struct. **6**, 285–300 (1999).
[16] W. M. Visscher, A. Migliori, T. M. Bell, and R. A. Reinert, "On the Normal-Modes of Free-Vibration of Homogeneous and Anisotropic Elastic Objects," J. Acoust. Soc. Am. **90**, 2154–2164 (1991).
[17] G. J. Kynch, "The Fundamental Modes of Vibration of Uniform Beams for Medium Wavelengths," Br. J. Appl. Phys. **8**, 64–73 (1957).
[18] N. J. Nigro, "Steady-state Wave Propagation in Infinite Bars of Noncircular Cross Section," J. Acoust. Soc. Am. **40**, 1501–1508 (1966).
[19] W. B. Fraser, "Stress Wave Propagation in Rectangular Bars," Int. J. Solids Struct. **5**, 1501–1508 (1969).
[20] R. W. Morse, "Dispersion of Compressional Waves in Isotropic Rods of Rectangular Cross Section," J. Acoust. Soc. Am. **20**, 833–838 (1948).
[21] R. D. Mindlin and E. A. Fox, "Vibrations and Waves in Elastic Bars of Rectangular Cross Section," J. Appl. Mech. **27**, 152–158 (1960).
[22] R. D. Mindlin, "Waves and Vibrations in Isotropic, Elastic Plates," in *Proceedings of the First Symposium on Naval Structural Mechanics*, edited by J. N. Goodier and N. J. Hoff (Pergamon, New York, 1960).

The photoacoustic effect generated by an incompressible sphere

Gerald J. Diebold,[a] Andrew C. Beveridge, and Theron J. Hamilton
Department of Chemistry, Brown University, Providence, Rhode Island 02912

(Received 3 April 2002; revised 13 July 2002; accepted 29 July 2002)

An incompressible sphere with a vanishing thermal expansivity suspended in a fluid can generate a photoacoustic effect when the heat deposited in the sphere by a light beam diffuses into the surrounding liquid causing it to expand and launch a sound wave. The properties of the photoacoustic effect for the sphere are found using a Green's function solution to the wave equation for pressure with Neumann boundary conditions. The results of the calculation show that the acoustic wave for fast heat liberation is an outgoing compressive pulse followed by a reflected pulse whose time profile is modified as a result of frequency dependent reflection from the sphere. For slow heat release by the sphere, the photoacoustic effect is shown to be proportional to the first time derivative of the heat flux at the particle-fluid interface. © *2002 Acoustical Society of America.*
[DOI: 10.1121/1.1508788]

PACS numbers: 43.20.Rz, 43.35.Ud [LLT]

I. INTRODUCTION

The photoacoustic effect,[1–3] which refers to the production of sound through the absorption of electromagnetic radiation is frequently produced by directing the output of a pulsed laser into a weakly absorbing fluid with the acoustic pulse detected by a fast transducer. Although the heat deposition in fluids resulting from optical absorption is typically initiated by optical excitation of a dye in solution, or absorption by the fluid itself, particulate matter can act as a strong absorber of radiation acting as a source of heat that generates sound. Recent experiments with particulate suspensions irradiated with the output of high power, pulsed lasers[4–11] have shown extremely large amplitude photoacoustic waves can be produced. Similarly, transient grating experiments[11–13] have shown that diffracted light signals generated by irradiation of particulate suspensions can be strikingly different from those produced by dyes in solution. The problem of photoacoustic wave generation from particles has been treated previously,[14–17] but without consideration of the effects of reflection of sound generated in the fluid surrounding the particle by the particle itself.

In this paper we consider the problem of uniform heat deposition in a particle by absorption of optical radiation, followed by heat diffusion into a liquid which generates ultrasound. The particle is taken to be incompressible and to have a vanishing thermal expansion coefficient so that sound generation takes place solely in the fluid surrounding the particle. Section II gives the wave equation for pressure and its solution in terms of integrals over the source using the Green's function method. In Sec. III, this solution is evaluated for the problems of Gaussian and delta function deposition of heat in space outside the sphere. Important features of the time profile of the acoustic wave are explained in terms of reflection of an incoming spherical wave at a spherical interface. In Sec. IV, the photoacoustic effect is shown to be proportional to the first time derivative of the heat flux at the particle-fluid interface for slow heat release from the sphere. The Discussion section describes conditions for application of the mathematical expressions derived in the paper to generation of sound in particulate suspensions.

II. THE WAVE EQUATION AND GREEN'S FUNCTION SOLUTION

For a fluid whose heat capacity ratio is approximately unity, the coupled equations for pressure and temperature that describe the photoacoustic[18] effect decouple to give a wave equation for pressure with the source function proportional to the second time derivative of the temperature. The wave equation[19] for the photoacoustic pressure $p(\mathbf{r},t)$ can then be written for an inviscid fluid in terms of a heating function $H(\mathbf{r},t)$ as

$$\left[\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]p(\mathbf{r},t)=-\frac{\beta}{c_P}\frac{\partial H(\mathbf{r},t)}{\partial t}, \tag{1}$$

where c is the sound speed, c_P is the specific heat capacity, β is the thermal expansion coefficient, $\mathbf{r}$ is the coordinate, t is the time, and H is defined as $H=\rho c_P\partial\tau/\partial t$, where τ is the temperature rise. The approach for determination of $p(\mathbf{r},t)$ used here is to Fourier transform the wave equation for pressure to give a Helmholtz equation in the space variable, solve the Helmholtz equation, and then back transform the resulting frequency domain solution into the time domain. Here, the following convention for the Fourier transform pair $f(t)$ and $F(\omega)$ is used:

$$\begin{aligned} f(t)&=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(\omega)e^{-i\omega t}\,d\omega,\\ F(\omega)&=\int_{-\infty}^{\infty}f(t)e^{i\omega t}\,dt, \end{aligned} \tag{2}$$

where lower and upper case letters are used for time and frequency domain functions, respectively. Fourier transformation of Eq. (1) gives the inhomogeneous Helmholtz equation,

$$(\nabla^2+k^2)P(\mathbf{r},\omega)=S(\mathbf{r},\omega), \tag{3}$$

[a]Electronic mail: gerald_diebold@brown.edu

where k is the wave number given by $k=\omega/c$, where $S(\mathbf{r},\omega)$ is the Fourier transform of $-(\beta/c_P)(\partial H/\partial t)$, and ω is the angular frequency in the Fourier transform pair. The solution to the Helmholtz equation can be found by the Green's function method[20] as an integral over the radial coordinate r and over the angles θ and ϕ, denoted Ω, as

$$P(r,\Omega,\omega)=\int d\Omega' \int_0^\infty G(r,\Omega;r',\Omega')S(r',\Omega',\omega)dr'. \tag{4}$$

The derivation of the Green's function for the wave equation for pressure with a hard sphere of radius a, which has been constructed to give time derivatives of the coordinate at the interface of the sphere as zero, is given in the Appendix. Since there are two different forms of the Green's function depending on the relative magnitudes of r and r', integration over the radial coordinate requires integration from a to r, and integration from r to infinity. Thus, the frequency domain pressure is found as

$$P(r,\Omega,\mu)=\mu a^2\sum_{l,m}\left\{\frac{h_l(\mu\hat{r})}{h_l'(\mu)}Y_{l,m}(\Omega)[I_{l,m}^{(1)}(\hat{r},\mu)y_l'(\mu)-I_{l,m}^{(2)}(\hat{r},\mu)j_l'(\mu)]+I_{l,m}^{(3)}(\hat{r},\mu)Y_{l,m}(\Omega)\times\left[\frac{j_l(\mu\hat{r})y_l'(\mu)-y_l(\mu\hat{r})j_l'(\mu)}{h_l'(\mu)}\right]\right\}, \tag{5}$$

where $\hat{r}$ and μ are defined as

$$\hat{r}=r/a \quad \text{and} \quad \mu=ka=\omega a/c,$$

and where the dimensionless integrals $I_{l,m}^{(1)}$ through $I_{l,m}^{(3)}$ are defined as

$$I_{l,m}^{(1)}(\hat{r},\mu)=\int_1^{\hat{r}} j_l(\mu\xi)\xi^2\,d\xi\int Y_{l,m}^*(\Omega')S(\xi,\Omega',\omega)d\Omega',$$

$$I_{l,m}^{(2)}(\hat{r},\mu)=\int_1^{\hat{r}} y_l(\mu\xi)\xi^2\,d\xi\int Y_{l,m}^*(\Omega')S(\xi,\Omega',\omega)d\Omega',$$

$$I_{l,m}^{(3)}(\hat{r},\mu)=\int_{\hat{r}}^{\infty} h_l(\mu\xi)\xi^2\,d\xi\int Y_{l,m}^*(\Omega')S(\xi,\Omega',\omega)d\Omega',$$

where the radial integration over r' has been written in terms of a dimensionless parameter ξ where $\xi=r'/a$. Note that these three integrals are, in general, functions of $\hat{r}$, and μ, but may also be functions of other parameters defined in S. The terms containing $I_{l,m}^{(1)}$ and $I_{l,m}^{(2)}$ are used for r outside the source region, and the terms containing $I_{l,m}^{(3)}$ are used for r inside the source region.

III. GAUSSIAN SHELL SOURCE

The features of the photoacoustic effect generated in the presence of a spherical reflector can be illustrated by evaluation of Eq. (5) for the problem of instantaneous heat deposition in a spherical shell with a Gaussian spatial profile where the heating function is given by

$$H(r,t)=\frac{E}{4\pi^{3/2}r^2\sigma}\delta(t)e^{-(r-r_0)^2/\sigma^2}\mathsf{u}(r-a), \tag{6}$$

where r_0 is the distance between the center of the source and the origin, σ is the parameter that describes the spatial extent of the source, and E is the energy deposited by the light beam in the shell. Note that H in Eq. (6) is spherically symmetric and can be multiplied by $\sqrt{4\pi}Y_{0,0}(\Omega)$, which is unity, to facilitate the angular integration in Eq. (5). Differentiation of Eq. (6) with respect to time, followed by Fourier transformation and integration by parts of the resulting expression gives

$$S(r,\Omega,\omega)=\frac{i\beta E}{2\pi\sigma c_P r^2}\omega Y_{0,0}(\Omega)e^{-(r-r_0)^2/\sigma^2}\mathsf{u}(r-a), \tag{7}$$

where u is the Heaviside function. Substitution of S into Eq. (5) and integration of the resulting expression over Ω collapses the summations over l and m owing to the orthogonality of the spherical harmonics so that the frequency domain pressure contains only the $l=m=0$ terms and can be written

$$P(r,\mu)=\frac{i\beta Ec}{4\pi a^2 c_P}\mu^2\left\{\frac{h_0(\mu\hat{r})}{h_0'(\mu)}[\overline{I}_{0,0}^{(1)}(\hat{r},\mu)y_0'(\mu)-\overline{I}_{0,0}^{(2)}(\hat{r},\mu)j_0'(\mu)]+\overline{I}_{0,0}^{(3)}(\overline{r},\mu)\times\left[\frac{j_0(\mu\hat{r})y_0'(\mu)-y_0(\mu\hat{r})j_0'(\mu)}{h_0'(\mu)}\right]\right\}, \tag{8}$$

where $\hat{r}$ and μ are defined as above, and where the dimensionless integrals $\overline{I}_{0,0}^{(1)}$ through $\overline{I}_{0,0}^{(3)}$ are defined as

$$\overline{I}_{0,0}^{(1)}(\hat{r},\mu)=\frac{1}{\hat{\sigma}\sqrt{\pi}}\int_1^{\hat{r}} j_0(\mu\xi)e^{-(\xi-\xi_0)^2/\hat{\sigma}^2}\,d\xi,$$

$$\overline{I}_{0,0}^{(2)}(\hat{r},\mu)=\frac{1}{\hat{\sigma}\sqrt{\pi}}\int_1^{\hat{r}} y_0(\mu\xi)e^{-(\xi-\xi_0)^2/\hat{\sigma}^2}\,d\xi,$$

$$\overline{I}_{0,0}^{(3)}(\hat{r},\mu)=\frac{1}{\hat{\sigma}\sqrt{\pi}}\int_{\hat{r}}^{\infty} h_0(\mu\xi)e^{-(\xi-\xi_0)^2/\hat{\sigma}^2}\,d\xi,$$

where ξ_0 and $\hat{\sigma}$ are defined as $\xi_0=r_0/a$ and $\hat{\sigma}=\sigma/a$. Again, the first two integrals are used for the region outside the source, and the terms containing the third integral are used for the region between the sphere and the source.

The time domain photoacoustic pressure according to Eq. (2) is thus found to be

$$p(r,t)=\frac{i\beta Ec^2}{8\pi^2 a^3 c_P}\int_{-\infty}^{\infty} d\mu\,\mu^2\left\{\frac{h_0(\mu\hat{r})}{h_0'(\mu)}[\overline{I}_{0,0}^{(1)}(\hat{r},\mu)y_0'(\mu)-\overline{I}_{0,0}^{(2)}(\hat{r},\mu)j_0'(\mu)]+\overline{I}_{0,0}^{(3)}(\hat{r},\mu)\times\left[\frac{j_0(\mu\hat{r})y_0'(\mu)-y_0(\mu\hat{r})j_0'(\mu)}{h_0'(\mu)}\right]\right\}e^{-i\mu\hat{t}}, \tag{9}$$

where a dimensionless time $\hat{t}$ is defined as $\hat{t}=ct/a$ and where the same rule for use of the two factors in brackets as in the above frequency domain expressions obtains. The time domain wave form found by numerical Fourier transformation of Eq. (9) is shown in Fig. 1. The deposition of heat outside the sphere can be seen to cause generation of an

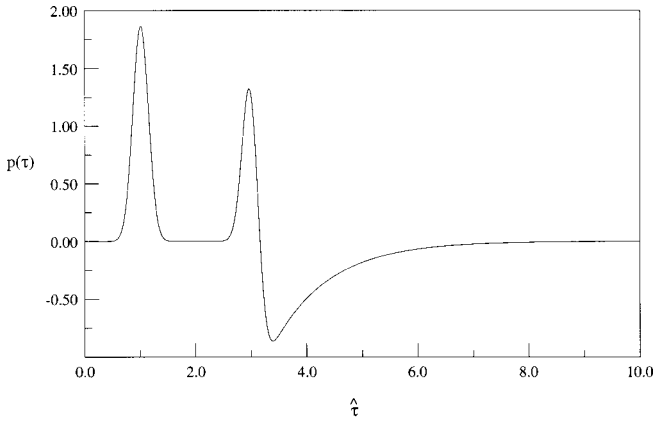


FIG. 1. Photoacoustic pressure $p(\hat{r},\hat{t})$ in arbitrary units versus dimensionless time $\hat{t}=ct/a$ for instantaneous heat deposition with a Gaussian spatial profile, calculated from Eq. (8) with $\xi_0=2$, and with $\hat{r}=3$.

initial outgoing spherical acoustic wave that has a temporal profile similar[21] to that of the Gaussian spatial profile, but with the second, reflected wave severely distorted by the reflection. In fact, the time profile of the reflected wave is necessarily modified by the reflection as can be seen by simplification of the spatial profile of the incident pulse, as is shown in the following section.

IV. INSTANTANEOUS HEAT DEPOSITION IN SPACE AND TIME

It is instructive to consider an instantaneous exciting pulse that deposits heat at a fixed distance from the origin. Since the Gaussian functions used above are normalized, they become delta functions in the limit of vanishing $\hat{\sigma}$; thus, the integrals in Eq. (8) become

$$\begin{aligned}\overline{I}^{(1)}_{0,0}(\hat{r},\mu)&=j_0(\mu\xi_0),\\ \overline{I}^{(2)}_{0,0}(\hat{r},\mu)&=y_0(\mu\xi_0),\\ \overline{I}^{(3)}_{0,0}(\hat{r},\mu)&=h_0(\mu\xi_0).\end{aligned}\tag{10}$$

Consider only the $\overline{I}^{(1)}_{0,0}$ and $\overline{I}^{(2)}_{0,0}$ terms for the waves outside the source, as would be recorded in an experiment with the detector far from the source and the reflecting sphere. The factor with brackets in Eq. (5) becomes $[j_0(\mu\xi_0)y_0'(\mu)-y_0(\mu\xi_0)j_0'(\mu)]$, which, through use of the definition of the spherical Hankel function (see the Appendix), can be shown to be identical to $-i[h_0'(\mu)j_0(\mu\xi_0)-h_0(\mu\xi_0)j_0'(\mu)]$. The photoacoustic pressure then becomes

$$p(r,t)=\frac{i\beta Ec^2}{8\pi^2 a_3 c_P}\int_{-\infty}^{\infty}d\mu\,\mu^2\left[h_0(\mu\hat{r})j_0(\mu\xi_0)-\frac{h_0(\mu\hat{r})}{h_0'(\mu)}h_0(\mu\xi_0)j_0'(\mu)\right]e^{-i\mu\hat{t}}.\tag{11}$$

Substitution of explicit expressions for the spherical Hankel and Bessel functions, $h_0(z)=e^{iz}/iz$ and $j_0(z)=\sin z/z$, and their derivatives into Eq. (11) gives, after some manipulation, the photoacoustic pressure as

$$p(r,t)=\frac{\beta Ec^2}{4\pi^2 r_0^3 c_P\hat{r}}\int_{-\infty}^{\infty}d\mu\left[e^{-i\mu\hat{\tau}_1}+\left(\frac{\mu-i}{\mu+i}\right)e^{-i\mu\hat{\tau}_2}\right],\tag{12}$$

where the dimensionless retarded times $\hat{\tau}_1$ and $\hat{\tau}_2$ have been defined as

$$\hat{\tau}_1=\frac{c}{a}\left(t-\frac{r-r_0}{c}\right)\quad\text{and}\quad\hat{\tau}_2=\frac{c}{a}\left(t-\frac{r+r_0-2a}{c}\right).$$

The photoacoustic wave can thus be seen as the sum of two waves, the first of which, given by the first term in Eq. (12), is an outgoing spherical delta function in space appearing at the field point r at a time $(r-r_0)/c$ after the heat deposition at $t=0$. The second term in Eq. (12) describes a wave that appears delayed in time from the arrival of the first pulse at r by the amount $2(r_0-a)/c$, which is the time for propagation from r_0 to the sphere and back again to r_0. Note that the second wave has nearly the form of a delta function in time and space, except for the modification of its amplitude and phase by the factor $(\mu-i)/(\mu+i)$.

The origin of the factor the modifies the reflected wave can be seen by determining the reflection coefficient of an inwardly propagating spherical wave at a spherical fluid interface. Consider an inwardly propagating wave P_I, a reflected wave P_R, and a transmitted wave P_T described by

$$\begin{aligned}P_I(r)&=\frac{1}{\hat{r}}e^{i[-k_f(r-a)-\omega t]},\\ P_R(r)&=\frac{R}{\hat{r}}e^{i[k_f(r-a)-\omega t]},\\ P_T(r)&=\frac{T}{\hat{r}}e^{i[-k_s(r-a)-\omega t]},\end{aligned}\tag{13}$$

where the subscripts k_s and k_f refer to wave numbers in the sphere and the surrounding fluid, respectively, and R and T are amplitude reflection and transmission coefficients per unit incident pressure, respectively. Requirement of contintinuty in the pressure and acceleration at the interface between the sphere and the surrounding fluid gives two equations, from which R and T are found as

$$\begin{aligned}R&=\frac{(\hat{\rho}-1)+i\tilde{\mu}(\hat{\rho}-1/\hat{c})}{-(\hat{\rho}-1)+i\tilde{\mu}(\hat{\rho}-1/\hat{c})},\\ T&=\frac{2i\tilde{\mu}\hat{\rho}}{-(\hat{\rho}-1)+i\tilde{\mu}(\hat{\rho}-1/\hat{c})},\end{aligned}\tag{14}$$

where $\tilde{\mu}$ is a dimensionless frequency parameter defined as $\tilde{\mu}=\omega a/c_f$, where the density $\hat{\rho}$ and the sound speed $\hat{c}$ ratios are defined as $\hat{\rho}=\rho_s/\rho_f$ and $\hat{c}=c_s/c_f$, respectively. In the limit of large $\hat{\rho}$, both the reflection and transmission coefficients remain complex and frequency dependent with

$$\lim_{\hat{\rho}\to\infty}R=\frac{\tilde{\mu}-i}{\tilde{\mu}+i},\tag{15}$$

which is precisely the factor multiplying the second term in Eq. (12). The complex amplitude of the frequency domain expression for the wave returning from the sphere is seen to be modified by reflection from the incompressible sphere. It is assumed that the reflection coefficient for a spherical wave

reflecting from a droplet or a solid sphere should be identical in the limit of large densities since the acceleration of the fluid at the interface must be zero in either case. The distortion in the wave reflecting from the sphere seen in Fig. 1 is thus a direct consequence of the frequency dependence of the reflection coefficient.

V. PHOTOACOUSTIC SIGNALS GENERATED BY PARTICLES

For an arbitrary heat flux emanating from a sphere $q(t)$, the photoacoustic pressure can be found by calculating the Fourier transform of $q(t)$, $Q(\omega)$, determining $S(r,\Omega,\omega)$, and carrying out the integration specified in Eq. (5) to give $P(r,\Omega,\mu)$. The time domain pressure $p(r,\Omega,t)$ is then found by Fourier transformation of $P(r,\Omega,\mu)$, which can, in most cases, be carried out numerically. However, if the heat release by the sphere is slow, that is, if the time for heat diffusion into the fluid is long compared with the transit time of sound across the heated region of the fluid, then the exact spatial profile of the heat deposition can be shown to be of no consequence in determining the time profile of the photoacoustic wave, and a much simpler method than the exact procedure for finding the photoacoustic pressure can be used.

The photoacoustic wave for slow heat release can be found by first determining the low frequency limit of the frequency domain pressure for delta function heat deposition, and then convoluting the resulting expression with the time domain heat flux from the irradiated sphere.[22] The method is based on the equivalence of frequency domain multiplication and time domain convolution,[23]

$$p(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}P(\omega)q(\omega)e^{-i\omega t}\,d\omega$$

$$=\int_{-\infty}^{\infty}p(t-\tau)Q(\tau)d\tau, \tag{16}$$

where $P(\omega)$ and $p(t)$ as well as $Q(\omega)$ and $q(t)$ are Fourier transform pairs as in Eqs. (2), where $p(t)$ here is taken as the response to delta function heat deposition, and $P(\omega)$ is appropriately normalized to unit heat flux. Since in the first of Eq. (16) it is assumed that $q(t)$ is a long pulse so that $Q(\omega)$ has only low frequency components, it follows[21] that only the low frequency components of $P(\omega)$ can contribute to the integral since the integration is over the product of $P(\omega)$ and $Q(\omega)$; thus, it is possible to take the low frequency limit of $Q(\omega)$, Fourier transform the resulting expression, and to use it in the convolution integral in the second of Eq. (16). The value of this procedure, in addition to determining the photoacoustic response to slow heat release, is to give a simple rule for the determining the functional form of the photoacoustic pressure for an arbitrary heat flux function.

Consider heat release from a sphere where the heat flux vector $q(t)$ is given by $q(t)=q_0\delta(t)$, and where the heat distribution is considered to be deposited over a region extending from a to d so that the heating function becomes

$$H(r,t)=\frac{3a^2q_0}{4\pi(d^3-a^3)}\,\delta(t)\hat{\Theta}_{a,d}(r), \tag{17}$$

where q_0 is the heat per unit area, and $\hat{\Theta}_{a,d}(r)$ is a square wave function that is unity for $a\leqslant r\leqslant d$, but zero otherwise. The source term in the wave equation for pressure can be shown to be given by

$$S(r,\Omega,\omega)=\frac{3ia^2\beta q_0}{4\pi(d^3-a^3)c_P}\,\omega\hat{\Theta}_{a,d}(r), \tag{18}$$

which can be multiplied by the factor $[\sqrt{4\pi}Y_{0,0}(\Omega)]$ as in Eq. (6) to facilitate the angular integration.

The integrals $I_{l,m}^{(1)}$ and $I_{l,m}^{(2)}$ in Eq. (5) can be evaluated for low frequencies by using explicit forms for the spherical Bessel and Neumann functions and their limiting forms for small arguments, $j_0(z)\simeq 1$, $y_0(z)\simeq -1/z$, $h_0'(z)\simeq i\exp(iz)/z^2$, and $h_0(z)=i\exp(iz)/z^2$ so that $I^{(1)}=(iq_0\beta\omega/\sqrt{4\pi}c_Pa)\,\delta_{l,0}\delta_{m,0}$. The term containing $I^{(2)}$ in Eq. (5) can be taken as zero since $j_0'(z)\simeq 0$ for small z. Substitution of the limiting forms of the spherical Hankel function and its derivative together with $I^{(1)}$ into Eq. (5) gives a frequency domain expression for the acoustic pressure, which, when Fourier transformed, yields

$$p(r,t)=-\frac{iq_0\beta c^2}{8\pi^2ac_P\hat{r}}\int_{-\infty}^{\infty}\mu e^{-i\mu(c/a)[t-(r-a)/c]}\,d\mu. \tag{19}$$

Since the quantity a/c is small compared with the time for heat evolution, Eq. (19) can be written as

$$p(r,t)=\frac{\beta q_0c^2}{\pi^2ac_P}\frac{d}{d\hat{\tau}}\,\delta(\hat{\tau}), \tag{20}$$

where the dimensionless retarded time $\hat{\tau}$ is defined as $\hat{\tau}=(c/a)[t-r/c]$. To find the time domain photoacoustic pressure for an arbitrary $q(t)$, the magnitude of the heat flux emanating from the particle (with dimensions of energy per unit area and time) q_0 is replaced by $q(t)$ in the second of Eq. (16) and the resulting expression integrated by parts to give the photoacoustic pressure for slow heat release as

$$p(r,t)=\frac{\beta a}{8\pi^2\hat{r}c_P}\frac{d}{dt}q\left(t-\frac{r}{c}\right). \tag{21}$$

It is apparent from the form of Eq. (20) that the photoacoustic pressure for any slow deposition of heat from an incompressible particle with vanishing thermal expansion coefficient is the first derivative of the heat flux transferred to the fluid.

VI. DISCUSSION

Perhaps the most important factor in determining the character of the photoacoustic wave from the imcompressible particle is the frequency dependence of the reflection coefficient for a collapsing compressional spherical wave. For high frequencies, the reflected wave returns from the sphere with an identical amplitude as the incident wave. As the wavelength of the incoming wave becomes commensurate with the radius of the sphere, its phase and amplitude are modified to produce compressions and rarefactions; finally, at very low frequencies, the incident wave returns from the sphere inverted. The behavior of the wave in the limit of low

frequency is to reflect at the origin as if no particle were present, obeying the homogeneous wave equation for pressure, as is described in Ref. 24.

In general, the photoacoustic effect from a heated incompressible sphere with a vanishing thermal expansion coefficient is given by Eq. (5) together with Eq. (16), which together provide a general method for determining the photoacoustic effect from an arbitrary heat flux from an incompressible particle. The derivation of Eq. (21), which shows that the photoacoustic wave for slow heat release from an incompressible sphere is proportional to the first derivative of the heat flux emanating from the particle, requires that the time for heat liberation be long compared with the transit time of sound across the radius of the sphere, a/c. Note that the expression for the volume in which the heat is deposited, $\frac{4}{3}\pi(d^3-a^3)$ in Eq. (17), does not appear in the low frequency expression for the acoustic pressure. It is a general principle that for slow heat deposition, the photoacoustic effect becomes proportional to various time derivatives of the light intensity depending on the dimension of the symmetry of the absorbing body;[21] the dimensions and physical properties of the absorbing body become irrelevant for determinating the time profile of the emitted wave.

The important point for the present problem is that the derivation of Eq. (21) is carried out under the assumption that the heat flux from the sphere is devoid of high frequency components. Note that for deposition of heat directly into a fluid droplet[22] with a finite thermal expansivity by a laser with a smooth time profile, it is sufficient to specify the pulse width of the optical radiation and to require it to be longer than a/c for the low frequency limit to apply.[22] However, for the present problem, it is possible that even though the overall time for heat release may be long compared to a/c, the initial release of heat through diffusion can be sufficiently fast that the conditions for the absence of high frequency components in $q(t)$ may not be fulfilled. In particular, instantaneous deposition of heat, *i.e.*, heat deposition in a time short compared with a/c, into a sphere may result in slow overall heat release, but with a rapid initial transfer of heat from the sphere to the fluid, violating the conditions for the derivation of the low frequency result. Of course, for heat deposition in the particle using a laser with a finite pulse width, the initial risetime of the optical pulse necessarily limits the frequency content of the heat flux transferred to the fluid since the heat flux from the particle is determined by a convolution of the laser time profile with the heat flux determined by delta function heat deposition in the particle. For the case of a 100 nm diameter water droplet surrounded by water, the thermal diffusion time from the sphere is 8.9 ns, and the quantitiy a/c is 33 ps. Thus, for laser pulses even as short as 1 ns, the photoacoustic effect will be proportional to the first derivative of the heat flux from the droplet since a/c is so small.

It is well known in photoacoustics that deposition of heat in a body with a short optical pulse gives a photoacoustic wave that contains information about physical properties of the body, or, in cases where linear thermoelastic sound generation is not operative, information about the mechanism of sound generation itself. Examples of this are the determination of the acoustic properties and diameter of a uniformly irradiated sphere,[22,24–26] the determination of the position of an irradiated body in space,[27–29] and determination of the sound wave generation mechanism for heat deposited at a point in space.[30] The results given here show that, in principle, the above generalization is operative for a pulse generated outside an incompressible sphere: for the initial outgoing wave, the spatial distribution of heat is mapped into the time profile of the acoustic wave, while the change in time profile of reflected wave necessarily contains information about the acoustic properties of the reflecting sphere.

ACKNOWLEDGMENT

The authors are grateful for support of this research by the US Department of Energy under Grant No. ER-13235.

APPENDIX: THE GREEN'S FUNCTION FOR A HARD SPHERE

For the problem of a hard sphere[31] of radius a surrounded by a fluid, the time derivatives of the coordinate at $r=a$ should be zero so that the fluid motion vanishes at the surface of the sphere. It is sufficient to take the acceleration $\dot{\mathbf{u}}$ to be zero. Since the acceleration is related to the gradient of the pressure and the density ρ through the relation $\dot{\mathbf{u}}=-\boldsymbol{\nabla} p/\rho$, the Green's function $G(r,\Omega;r',\Omega')$ for a problem with spherical symmetry must be constructed so that $\partial G(r,\Omega;r',\Omega')/\partial r'=0$. The Green's function for the Helmholtz equation must obey the condition that it vanish at infinite distance from the sphere and satisfy

$$(\boldsymbol{\nabla}^2+k^2)G(r,\Omega;r',\Omega')=\frac{\delta(r-r')}{r^2}\,\delta(\bar{\Omega}), \tag{A1}$$

where $\bar{\Omega}$ is the angle between Ω and Ω'.

A function that satisfies these conditions can be constructed from the solutions to the homogeneous Helmholtz equation, namely, the spherical Bessel, Neumann, and Hankel functions, $j_l(kr)$, $y_l(kr)$, and $h_l(kr)$, respectively, where the last of these is the Hankel function of the first kind, defined as $h_l(kr)=j_l(kr)+iy_l(kr)$. For the incompressible sphere, it follows that a Green's function of the form

$$\begin{aligned} G_{r>r'}(r,\Omega;r',\Omega')&=\sum_{l,m} A_l h_l(kr) Y_{l,m}(\bar{\Omega}),\\ G_{r<r'}(r,\Omega;r',\Omega')&=\sum_{l,m} B_l[j_l(kr)y_l'(ka)\\ &\quad -y_l(kr)j_l'(ka)]Y_{l,m}(\bar{\Omega}) \end{aligned} \tag{A2}$$

satisfies both Eq. (A1) and the boundary conditions for the incompressible sphere.

Integration of Eq. (A1) from $r=r'-\epsilon$ to $r'+\epsilon$ for vanishing ϵ, gives an expression that can be satisfied for G continuous at r' but with $\partial G/\partial r$ discontinuous at $r=r'$. The conditions that $G(r,\Omega;r',\Omega')$ must obey are thus given by[20]

$$G|_{r'-\epsilon} - G|_{r'+\epsilon} = 0,$$
$$\left.\frac{\partial G}{\partial r}\right|_{r'+\epsilon} - \left.\frac{\partial G}{\partial r}\right|_{r'-\epsilon} = \frac{1}{r^2}\,\delta(\bar{\Omega}) \tag{A3}$$

which determine the coefficients $A_{l,m}$ and $B_{l,m}$ in the expansion of G in Eq. (A2). Substitution of the expression for the Green's function, Eq. (A2), into Eqs. (A3), and multiplication of the resulting expressions by $Y^*_{l',m'}(\bar{\Omega})$ gives two equations, which, when integrated over $d\Omega$, collapses the summation over l and m to give

$$A_{l,m}h_l(kr') - B_{l,m}[j_l(kr')y_l'(ka) - y_l(kr')j_l'(ka)] = 0,$$
$$A_{l,m}h_l(kr') - B_{l,m}[j_l'(kr')y_l'(ka) - y_l'(kr')j_l'(ka)] = \frac{1}{kr'^2}\,Y^*_{l,m}(0), \tag{A4}$$

where the orthogonality relation for spherical harmonics, $\int Y^*_{l,m}(\Omega)\; Y_{l',m'}(\Omega)\; d\Omega = \delta_{l,l'}\delta_{m,m'}$ has been used. The solution to these equations gives $A_{l,m}$ and $B_{l,m}$ as

$$B_{l,m} = k\left[\frac{h_l(kr')}{h_l'(ka)}\right]Y^*_{l,m}(0),$$
$$A_{l,m} = k\left[\frac{j_l(kr')y_l'(ka) - y_l(kr')j_l'(ka)}{h_l'(ka)}\right]Y^*_{l,m}(0).$$

Note that $Y^*_{l,0}(0,0)$ is real and can be expressed as

$$Y^*_{l,0}(0,0) = \sqrt{\frac{2l+1}{4\pi}}\,\delta_{m,0}$$

so that on substitution of $A_{l,m}$ and $B_{l,m}$ from Eq. (A5) into Eq. (A2) the summation over m collapses, leaving only the summation over l. Note that Ω refers to the angle between Ω and Ω'. It is convenient to write the Green's function in terms of Ω and Ω' which refer to the coordinates of the field and source, respectively, referred to a fixed coordinate system, which can be accomplished by first writing $Y_{l,0}$ in terms of the Legendre polynomial P_l through use of the definition of the spherical harmonics, which for $m=0$ gives

$$Y_{l,0}(\bar{\Omega}) = \sqrt{\frac{2l+1}{4\pi}}\,P_l(\cos\bar{\Omega}),$$

followed by application of the spherical harmonic addition theorem

$$P_l(\cos\theta') = \frac{4\pi}{2l+1}\sum_{m=-l}^{l} Y(\theta,\phi)Y^*(\theta',\phi'),$$

which gives the Green's function as

$$G_{r>r'}(r,\Omega;r',\Omega') = k\sum_{l,m}\frac{h_l(kr)}{h_l'(ka)}[j_l(kr')y_l'(ka) - y_l(kr')j_l'(ka)]Y^*_{l,m}(\Omega')Y_{l,m}(\Omega),$$
$$G_{r<r'}(r,\Omega;r',\Omega') = k\sum_{l,m}\frac{h_l(kr')}{h_l'(ka)}[j_l(kr)y_l'(ka) - y_l(kr)j_l'(ka)]Y^*_{l,m}(\Omega')Y_{l,m}(\Omega), \tag{A5}$$

where the summation over l and m extends from 0 to ∞, and the angles Ω and Ω' are independently referred to a fixed coordinate system.

Note that since the form of the Green's function changes depending on whether r is greater or smaller than r', the integration over r' specified in Eq. (4) must be carried out in two steps as

$$P(r,\Omega,\omega) = \int d\Omega'\int_a^r G(r,\Omega;r',\Omega')S(r',\Omega',\omega)dr' + \int d\Omega'\int_r^\infty G(r,\Omega;r',\Omega')S(r',\Omega',\omega)dr'. \tag{A6}$$

It should be noted that the Green's function derived here does not have the property $G(-\omega)=G^*(\omega)$ which permits the time domain pressure to be written as

$$p(r,\Omega,t) = \frac{1}{\pi}\,\mathrm{Re}\int_0^\infty P(r,\Omega,\omega)e^{-i\omega t}\,d\omega,$$

which is commonly taken as the form for the inverse Fourier transform in most numerical transformation algorithms. Instead, the Green's function derived here, or any frequency domain pressure found from Eq. (5), must be evaluated in two steps as

$$p(r,\Omega,t) = \frac{1}{2\pi}\int_0^\infty P(r,\Omega,\omega)e^{-i\omega t}\,d\omega + \frac{1}{2\pi}\int_0^\infty P(r,\Omega,-\omega)e^{i\omega t}\,d\omega,$$

the second integral being the common form of the Fourier transform (but with the integration over ω, and the first integral, the common form for the inverse Fourier transform).

[1] C. K. N. Patel and A. C. Tam, "Pulsed optoacoustic spectroscopy of matter," Rev. Mod. Phys. **53**, 517 (1981).
[2] V. E. Gusev and A. A. Karabutov, *Laser Optoacoustics* (AIP, New York, 1993), Chap. 2.
[3] F. V. Bunkin, Al. A. Kolomensky, and V. G. Mikhalevich, *Lasers in Acoustics* (Harwood Academic, Reading, MA, 1991).
[4] S. V. Egerev, A. E. Pashin, and Ya. O. Simanovskii "Sound radiation from cavitation axisymmetric sources induced with intense laser pulses," *Advances in Nonlinear Acoustics* (World Scientific, Singapore, 1993), p. 436.
[5] S. S. Alimpiev, Ya. O. Simanovskii, S. V. Egerev, and A. E. Pashin, "Optoacoustic detection of microparticles in liquids at laser fluences below the optical breakdown threshold," Laser Chem. **16**, 63 (1995).
[6] T. Kitamori, M. Fujii, T. Sawada, and Y. Gohshi, "Linear response theory in photoacoustic spectroscopy and characterization of turbid particles," Jpn. J. Appl. Phys., Part 1 **24**, 210 (1984).
[7] A. Yoshinaga, Y. Hsieh, T. Sawada, and Y. Gohshi, "Laser induced photoacoustic spectrometry with a single particle sample," Anal. Sci. **5**, 147 (1989).
[8] C. P. Lin and M. W. Kelly, "Cavitation and acoustic emission around laser-heated microparticles," Appl. Phys. Lett. **72**, 2800 (1998).
[9] C. P. Lin and M. W. Kelly, "Ultrafast time-resolved imaging of stress transient and cavitation from short pulsed laser irradiated melanin particles," Proc. SPIE **2391A**, 294 (1995).
[10] A. C. Beveridge, T. E. McGrath, and G. J. Diebold, "Photoacoustic shock generation in carbon suspensions," Appl. Phys. Lett. **75**, 4204 (1999).
[11] H. X. Chen and G. J. Diebold, "Chemical generation of acoustic waves: A "giant" photoacoustic effect," Science **270**, 963 (1995).
[12] T. E. McGrath, A. C. Beveridge, and G. J. Diebold, "Laser induced "regeneration" of colloidal particles: The effects of thermal inertia on the

chemical reactivity of laser heated particles,'' Angew. Chem. Int. Ed. Engl. **38**, 3353 (1999).
[13] H. X. Chen and G. J. Diebold, ''Production of the photoacoustic effect and transient gratings by molecular volume changes,'' J. Chem. Phys. **104**, 6730 (1996).
[14] L. M. Lyamshev, ''Optical-acoustic probing of heterogeneous condensed media,'' Sov. Phys. Dokl. **24**, 463 (1979).
[15] S. V. Egerev and A. E. Pashin, ''Optoacoustic diagnostics of liquids with ultrafine inhomogeneities,'' Acoust. Phys. **39**, 43 (1993).
[16] O. M. Zozula and O. V. Puchenkov, ''On theory of opto-acoustic effect in liquid disperse systems,'' Acoust. Phys. **39**, 46 (1993).
[17] Y. N. Cao, G. J. Diebold, and M. B. Zimmt, ''Transient grating studies of ultrasonic attenuation in reverse micellar solutions,'' Chem. Phys. Lett. **276**, 388 (1997).
[18] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1986).
[19] In P. J. Westervelt and R. S. Larson, ''Laser excited broadside array,'' J. Acoust. Soc. Am. **54**, 121 (1973), heat conduction is ignored, which rigorously gives a wave equation for pressure with H the energy per unit volume and time deposited by the light beam. Also, the term describing the temperature change arising from pressure fluctuations remains in the differential equation for the temperature. In the present paper, the heat capacity ratio is taken as approximately unity giving a finite thermal expansivity and a wave equation for p as an approximation. In this case, H is constructed to give $H(\mathbf{r},t)=\rho c_P \partial\tau(\mathbf{r},t)/\partial t$, for whatever process that changes τ. See Ref. 18, Eq. 6.4.22.
[20] J. Mathews and R. L. Walker, *Mathematical Methods of Physics* (Benjamin Cummings, Menlo Park, CA, 1970).
[21] The theorem for three-dimensional waves given in G. J. Diebold, T. Sun, and M. I. Khan, ''Photoacoustic monopole radiation in one, two, and three dimensions,'' Phys. Rev. Lett. **67**, 3384 (1991) should be valid for the Gaussian heat source in the absence of a reflecting surface.
[22] G. J. Diebold and P. J. Westervelt, ''The photoacoustic effect generated by a spherical droplet in a fluid,'' J. Acoust. Soc. Am. **84**, 2245 (1988).
[23] C. C. Goodyear, *Signals and Information* (Wiley-Interscience, New York, 1971). See Eq. (42); the sign convention for the Fourier transform pair in this text the opposite of that used here.
[24] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed. (Pergamon, Oxford, 1986), see Sec. 70.
[25] G. J. Diebold, M. I. Khan, and S. M. Park, ''Photoacoustic 'signatures' of particulate matter: Optical production of acoustic monopole radiation,'' Science **250**, 101 (1990).
[26] L. M. Lyamshev and K. A. Naugol'nykh, ''Sound generation by thermal sources,'' Sov. Phys. Acoust. **22**, 354 (1976).
[27] A. A. Oraevsky, S. L. Jacques, and F. K. Tittel, ''Measurement of tissue optical properties by time resolved detection of laser induced transient stress,'' Appl. Opt. **36**, 402 (1997).
[28] R. A. Kruger and P. Liu, ''Photoacoustic ultrasound: Pulse production and detection of 0.5% liposyn,'' Med. Phys. **21**, 1179 (1994).
[29] A. A. Karabutov, E. V. Savateeva, N. B. Podymova, and A. A. Oraevsky, ''Backward mode detection of laser induced wideband ultrasonic transients with optoacoustic transducers,'' J. Appl. Phys. **87**, 2003 (2000).
[30] I. G. Calasso, W. Craig, and G. J. Diebold, ''Photoacoustic point source,'' Phys. Rev. Lett. **86**, 3350 (2001).
[31] Reference 18 gives a Green's function for an incompressible sphere, without derivation, but using an uncommon convention for the spherical harmonics.

Rectified heat transfer into translating and pulsating vapor bubbles

Y. Hao and A. Prosperetti[a)]
Department of Mechanical Engineering, The Johns Hopkins University, Baltimore, Maryland 21218

(Received 20 February 2002; revised 3 July 2002; accepted 1 August 2002)

It is well known that, when a stationary vapor bubble is subject to a sufficiently intense acoustic field, it will grow by rectified heat transfer even in a subcooled liquid. The object of this paper is to study how translation, and the ensuing convective effects, influence this process. It is shown that, depending on the initial temperature distribution and other factors, convection can cause a destabilization of the bubble or its faster growth. Significant effects occur in parameter ranges readily encountered in practice. The phenomena described can therefore be exploited for bubble management, e.g., by increasing the condensation rate or promoting faster bubble growth and coalescence. In a saturated or a superheated liquid, heat rectification and convection reinforce each other and the bubble growth is accelerated by a translatory motion. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1508789]

PACS numbers: 43.25.Yw [AJS]

I. INTRODUCTION

A vapor bubble subject to a sound field in a subcooled liquid exhibits the phenomenon of rectified heat transfer, by which it can either grow, or collapse at a slower rate than without sound (see, e.g., Wang, 1974; Khabeev, 1976; Akulichev *et al.*, 1979; Patel *et al.*, 1985; Gumerov, 2000; Hao and Prosperetti, 1999; Hao *et al.*, 2001; Prosperetti and Hao, 2001, 2002). Similarly, in a superheated liquid, sound causes a faster growth than would occur in a constant pressure environment. The process is quite similar to the perhaps better known rectified diffusion of mass (see, e.g., Fyrillas and Szeri, 1994).

The physical origin of this behavior is due to the opposite effect of two competing mechanisms, which justifies the adjective ''rectified'' given to this phenomenon. When the bubble is compressed, some vapor condenses, the surface temperature rises, and heat is conducted away into the adjacent liquid. When the bubble expands during the following half cycle, evaporation causes a temperature drop of the bubble surface, with a consequent heat flux from the liquid. The net result of these opposing processes exhibits a bias toward an energy gain by the bubble because (i) the spherical geometry forces a thicker boundary layer—and hence a smaller heat flux—during compression, and (ii) the surface area available for the phase change is bigger during expansion than during compression.[1]

In view of the reliance of these effects on heat transfer between the bubble and the liquid, it may be expected that they would be strongly influenced by a translational motion of the bubble. This expectation is borne out by the results of this paper, which presents a numerical investigation of the process under the simplifying assumption of a spherical bubble. Depending on conditions, in a subcooled liquid, we find that convection can either make the bubble more labile or increase its growth rate. The bubble condenses if it is abruptly exposed to cold liquid and the velocity is sufficiently large. A bubble which survives this initial stage, on the other hand, may grow faster than without translation. The explanation for this paradoxical finding is that, as explained before, the amount of heat transferred is greater during the expansion phase than the compression phase. During the expansion phase the bubble is cooler than its surroundings and, therefore, convection helps to increase its vapor content.

With this study we continue the investigation of the behavior of vapor bubbles under the joint action of acoustic and flow fields started in Hao and Prosperetti (1999) and continued in subsequent papers (Hao and Prosperetti, 2000; Hao *et al.*, 2001; Prosperetti and Hao, 2001). This series of studies responds to the practical need of vapor bubble management under microgravity conditions, both to increase the critical heat flux in boiling and to control the void fraction of flowing two-phase cryogenic mixtures (Oka *et al.*, 1992; Ervin *et al.*, 1992; Sitter *et al.*, 1998a, b).

II. MATHEMATICAL FORMULATION AND NUMERICAL METHOD

The mathematical model is based on our earlier work (Hao and Prosperetti, 1999, 2000); here we only present a summary of the formulation.

The bubble is assumed to remain spherical, with a radius $R(t)$ determined by a form of the Rayleigh–Plesset equation which accounts for weak liquid compressibility effects:

$$\left(1-\frac{\dot R}{c}\right)R\ddot R+\frac{3}{2}\left(1-\frac{\dot R}{3c}\right)\dot R^2 = \frac{1}{\rho}\left(1+\frac{\dot R}{c}+\frac{R}{c}\frac{d}{dt}\right)[p_b-P(t)]. \tag{1}$$

Here dots denote time differentiation, ρ and c are the liquid density and speed of sound, p_b is the pressure on the liquid

a)Author to whom correspondence should be addressed; electronic mail: prosperetti@jhu.edu. Also with Department of Applied Physics, Twente Institute of Mechanics, and Burgerscentrum, University of Twente, AE 7500 Enschede, The Netherlands.

side of the interface, and $P(t)$ is the ambient pressure which we will take in the form

$$P(t)=P_\infty \pm P_A \sin \omega t, \tag{2}$$

in which P_∞ is the static pressure, P_A is the acoustic pressure amplitude, and ω is the angular frequency of the sound field; the upper sign is chosen when the initial phase of the sound is compressive, and the lower one when it is expansive. It will be seen that the phase of the sound field is very important in the initial stages of the process. We assume the vapor pressure in the bubble p_V to be spatially uniform and in saturated conditions with respect to the bubble surface temperature T_S; thus

$$p_V=p_{\mathrm{sat}}(T_S)=p_b+\frac{2\sigma}{R}+4\mu\frac{\dot{R}}{R}, \tag{3}$$

in which σ and μ are the surface tension and viscosity coefficients; here and in the following we append the subscript V to quantities pertaining to the vapor, and the subscript sat indicates quantities evaluated along the saturation line.

The energy equation in the liquid is

$$\frac{\partial T}{\partial t}+\mathbf{u}\cdot\boldsymbol{\nabla} T=D\nabla^2 T, \tag{4}$$

where T is the liquid temperature, D is the liquid thermal diffusivity, and $\mathbf{u}$ is the liquid velocity field in the bubble rest frame. This equation is solved subject to the condition of undisturbed temperature far from the bubble, $T\to T_\infty$ for $r\to\infty$ while, at the bubble surface, conservation of energy dictates that (Hao and Prosperetti, 1999)

$$4\pi R^2 k\left.\frac{\partial T}{\partial r}\right|_{r=R(t)}=L\frac{d}{dt}\left(\frac{4}{3}\pi R^3\rho_V\right)+\frac{4}{3}\pi R^3\rho_V c_s\frac{dT_S}{dt}. \tag{5}$$

Here k and L are the liquid thermal conductivity and latent heat, respectively, while $c_s=c_{pV}-L/T_S$ is the specific heat of the vapor along the saturation line; c_{pV} is the vapor specific heat at constant pressure. For most of the calculations reported below, the initial temperature field equals T_∞ everywhere, including the bubble surface.

We assume the bubble to move rectilinearly with velocity U with respect to the liquid at rest at infinity. Keeping in mind the temperature dependence of viscosity, we find that typical Reynolds numbers in the cases we consider are of the order of several hundreds, which justifies the approximation of potential flow except possibly in the last stages of the total condensation of a bubble. Therefore we assume that, in the rest frame of the bubble,

$$\mathbf{u}=\boldsymbol{\nabla}\left[-\frac{R^2\dot{R}}{r}-Ur\left(1+\frac{R^3}{2r^3}\right)\cos\theta\right], \tag{6}$$

where θ is the polar angle measured from the front stagnation point and r is the distance from the bubble center. In this paper we will mostly take the velocity of the bubble to be a constant.

For the solution of the energy equation (4) we use the same method of Hao and Prosperetti (2000): the temperature field is expanded in a series of Legendre polynomials

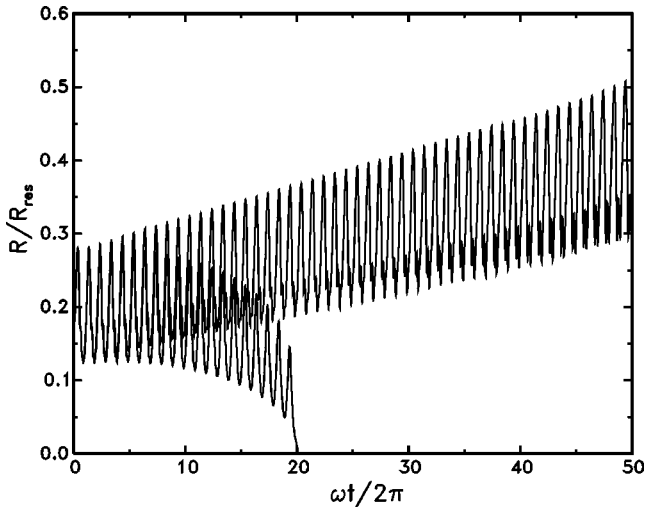


FIG. 1. Nondimensional radius versus time of bubbles translating with constant velocity in water at 95 °C and 101.3 kPa static pressure in a 1 kHz, 30.39 kPa sound field. The upper line is for a velocity of 0.2 m/s, and the lower line for 0.3 m/s. The radius is normalized by the resonant radius of a stationary bubble equal to 2.367 mm; the initial radius is 0.5 mm. The initial phase of the sound field promotes expansion.

$$T=T_\infty+\sum_{N=0}^{\infty} S_N(r,t)P_N(\cos\theta), \tag{7}$$

and the coupled equations for the various S_N that arise upon substitution into (4) are solved by expanding each S_N in a series of Chebyshev polynomials; the resulting equations are solved by collocation. The reader is referred to the reference for details.

In the numerical implementation of this scheme we typically use 8 terms in the Legendre expansion (7) and 16 terms in the Chebyshev expansion for the S_N's with 16 collocation points. These values have been chosen on the basis of standard convergence tests. The resulting ordinary differential equation systems are solved by the backward differentiation algorithm of the LSODE package which automatically selects the time step.

In the limit cases of no sound or no translational velocities, we have verified that the code developed for this work gave the same results as those used in our earlier papers dealing with these limit cases (Hao and Prosperetti, 1999, 2000). Those codes had been validated against analytical solutions and independent numerical solutions.

III. NUMERICAL RESULTS

The role of convection in the different stages of the process under study is quite different and, accordingly, we treat separately phenomena taking place during the first few acoustic cycles and over a longer time period.

A. Short-time behavior

Figure 1 compares the evolution of two bubbles, initially both with a radius $R(0)=0.5$ mm, subject to a sound field with acoustic amplitude $P_A=30.4$ kPa ($P_A/P_\infty=0.3$) and frequency $\omega/2\pi=1$ kHz; the liquid is water at $T_\infty=95$ °C

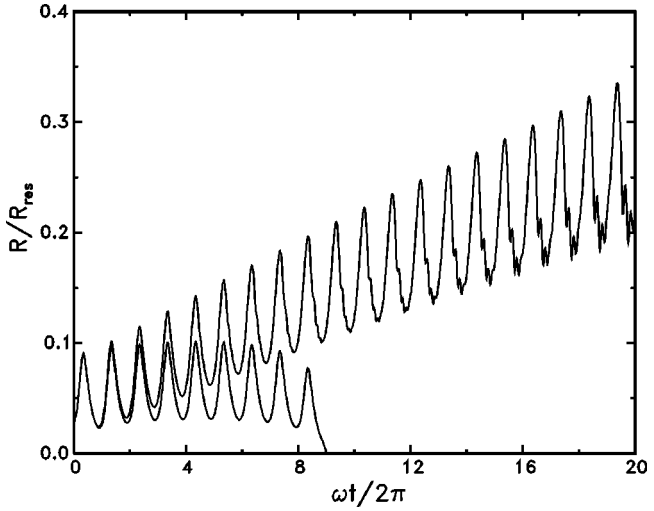


FIG. 2. As in Fig. 1 with $P_A=40.52$ kPa and $\omega/2\pi=0.2$ kHz; upper line $U=0.2$ m/s, lower line 0.24 m/s; here $R_{\text{res}}=13.47$ mm.

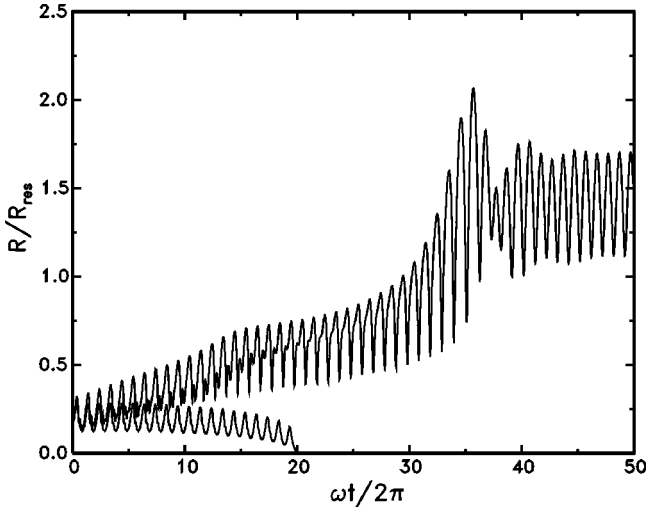


FIG. 3. Nondimensional radius versus time of bubbles translating at 0.3 m/s in water at 95 °C and 101.3 kPa static pressure in a 1 kHz, sound field. The upper line is for an acoustic pressure amplitude $P_A=35.46$ kPa, the lower one for $P_A=30.39$ kPa. The radius is normalized by the linear resonant radius, $R_{\text{res}}=2.367$ mm, and the initial radius is 0.5 mm. The initial phase of the sound field promotes expansion.

and $P_\infty=101.3$ kPa, so that the subcooling $\Delta T=T_{\text{sat}}(P_\infty)-T_\infty$ (in which T_{sat} is the saturation temperature at P_∞) is 5 °C. The Jakob number, defined by

$$\text{Ja}=\rho c_{pL}\frac{T_{\text{sat}}(P_\infty)-T_\infty}{L\rho_{V,\text{sat}}(P_\infty)}, \tag{8}$$

in which c_{pL} is the liquid specific heat and $\rho_{V,\text{sat}}(P_\infty)$ is the saturation vapor density at P_∞, is 15.2 approximately. The upper curve is for a bubble translating with a constant velocity of 0.2 m/s, while the lower curve is for a translation velocity of 0.3 m/s;[2] the radius is normalized by R_{res}, the linear resonant value for a stationary bubble (Hao and Prosperetti, 1999) which here is 2.367 mm. The initial Péclet number, defined by

$$\text{Pe}=\frac{2R(0)U}{D}, \tag{9}$$

has values 1184 and 1777 in the two cases. The first bubble grows by rectified diffusion of heat while the other one executes oscillations around a decreasing mean radius until it eventually condenses completely. For comparison, it may be noted that, without sound, condensation is completed at $t=3.0$ ms ($\omega t/2\pi=3.0$) for $U=0.2$ m/s and at $t=2.5$ ms ($\omega t/2\pi=2.5$) for $U=0.3$ m/s.

A similar behavior is shown in Fig. 2, where $P_A=40.52$ kPa ($P_A/P_\infty=0.4$) and $\omega/2\pi=0.2$ kHz; the other parameters have the same value as before and $R_{\text{res}}=13.47$ mm. At this lower frequency the duration of the compression phase, during which the bubble loses mass, is longer, and a velocity of 0.24 m/s (lower line, Pe=1,420) is sufficient to cause a complete condensation after a few cycles while for $U=0.2$ m/s the bubble grows. The behavior shown in these figures is typical of the phenomena encountered during the initial stages of the process object of this study. However, a bubble that survives the first few cycles may actually grow faster when translating. We will address this point shortly. For the time being, we limit our analysis to the initial stages as in Figs. 1 and 2.

Similar results are found in dependence of other parameters as well. The effect of different acoustic pressures with a fixed velocity, $U=0.3$ m/s (Pe=1,777) are shown in Figs. 3 and 4. At 1 kHz (Fig. 3), with acoustic pressure amplitudes $P_A=35.46$ kPa (upper curve), rectified heat transfer is sufficiently strong to prevent the bubble from collapsing, but if P_A is reduced to 30.39 kPa, the bubble ultimately condenses. At 0.2 kHz (Fig. 4), the bubble is more labile and one needs a higher pressure amplitude, $P_A=45.57$ kPa (upper curve), to stabilize it. The effect of frequency is explored directly in Fig. 5, for the same initial radius and subcooling, $P_A=40.52$ kPa, $U=0.24$ m/s, and acoustic frequencies of 1 kHz (upper curve) and 0.2 kHz. In all these cases the minus sign (initial expansion) is taken in (2).

The general behavior is similar also for smaller bubbles; Figs. 6 and 7 are analogous to Figs. 1 and 4 for $R(0)=0.1$ mm. In Fig. 6 (1 kHz) the acoustic pressure amplitude

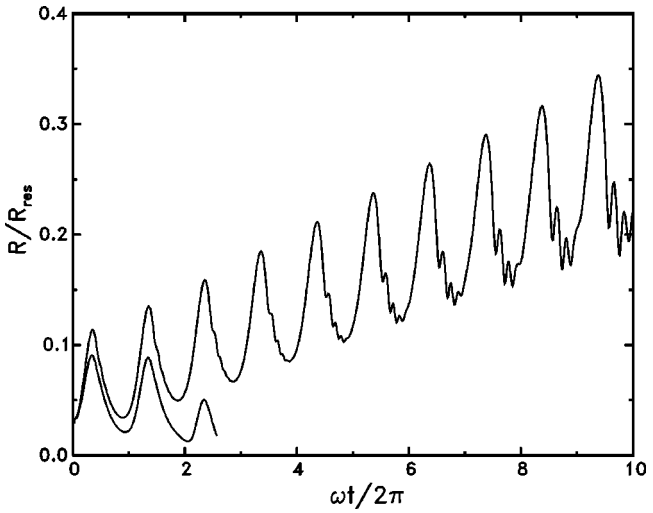


FIG. 4. As in Fig. 3 with $\omega/2\pi=0.2$ kHz ($R_{\text{res}}=13.47$ mm); upper curve $P_A=45.57$ kPa; lower curve $P_A=40.52$ kPa.

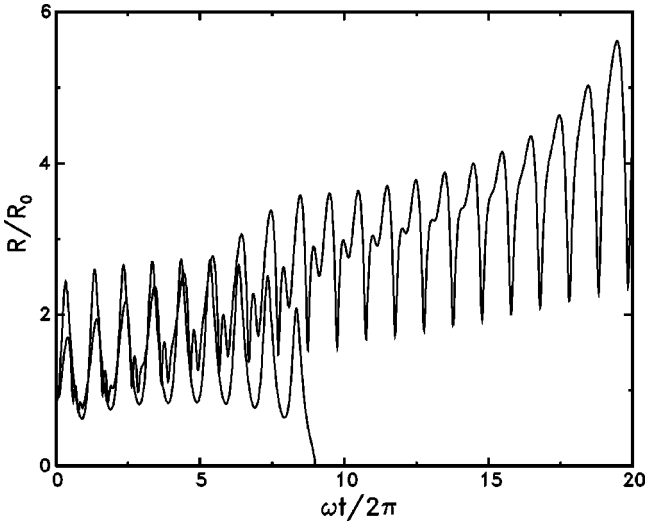


FIG. 5. Normalized radius versus time of bubbles translating at 0.24 m/s in water at 95 °C and 101.3 kPa static pressure in a 40.52 kPa sound field. The upper line is for $\omega/2\pi=1$ kHz ($R_{\rm res}=2.367$ mm), the lower line for $\omega/2\pi=0.2$ kHz ($R_{\rm res}=13.47$ mm). Since the resonant radius depends on frequency, unlike the other figures, here the radius is normalized by the initial value, $R(0)=0.5$ mm. The initial phase of the sound field promotes expansion.

is 40.52 kPa and, therefore, the bubble can withstand a bigger velocity (0.35 m/s, upper line) without condensing. When the velocity is raised to 0.42 m/s, however, rectified growth is impossible. An example at a lower frequency, $\omega/2\pi=0.2$ kHz, for two pressure amplitudes (50.65 kPa, upper curve, and 44.57 kPa) is shown in Fig. 7: upon comparison with the bigger bubble of Fig. 4, it is seen that a stronger forcing is needed to stabilize the bubble. Both Figs. 6 and 7 display the prominent presence of harmonics of the driving sound. These are a nonlinear effect and set in when the growing radius makes the bubble resonant with harmonics of the sound frequency (Hao and Prosperetti, 1999).

These results are affected by the initial temperature distribution around the bubble. As an example, in Fig. 8 we consider the collapsing bubble of Fig. 1 for $U=0.3$ m/s in an

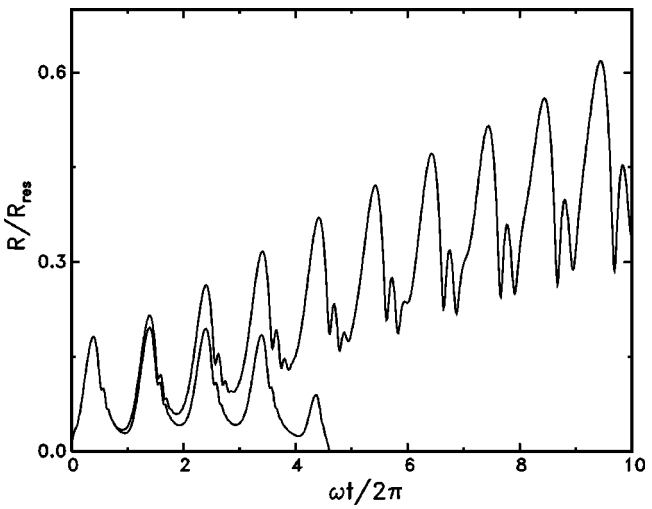


FIG. 6. As in Fig. 1 for an initial radius $R(0)=0.1$ mm. The acoustic pressure amplitude is 40.52 kPa; upper line, $U=0.35$ m/s, lower line 0.42 m/s.

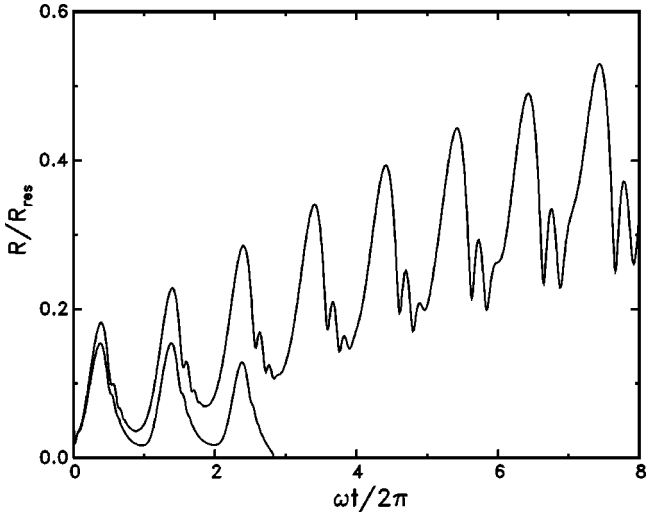


FIG. 7. As in Fig. 4 for an initial radius $R(0)=0.1$ mm; $U=0.3$ m/s; upper line $P_A=50.65$ kPa, lower line 44.57 kPa.

initial temperature field given, for $R\leqslant r\leqslant R+\delta$, by

$$T(r,\theta,0)=T_\infty+[T_{\rm sat}(P_\infty)-T_\infty]\left(1-\frac{r-R(0)}{\delta}\right)^2, \quad (10)$$

with $\delta/R(0)=0.5$ (solid line). The dashed line is the same result plotted in Fig. 1, with $T(r,\theta,0)=T_\infty$. With the initial temperature distribution (10) the bubble grows. As $\delta/R(0)$ is reduced, the growth process is found to slow down.

These results illustrate the sensitivity of the boundary between stable and unstable bubbles to the precise initial temperature distribution in the bubble neighborhood.[3] In this regard, the situation is similar to that encountered in Hao and Prosperetti (2000) for bubbles in the absence of sound. In view of the marked dependence of the phenomenon on such details and of the abundance of parameters and physical

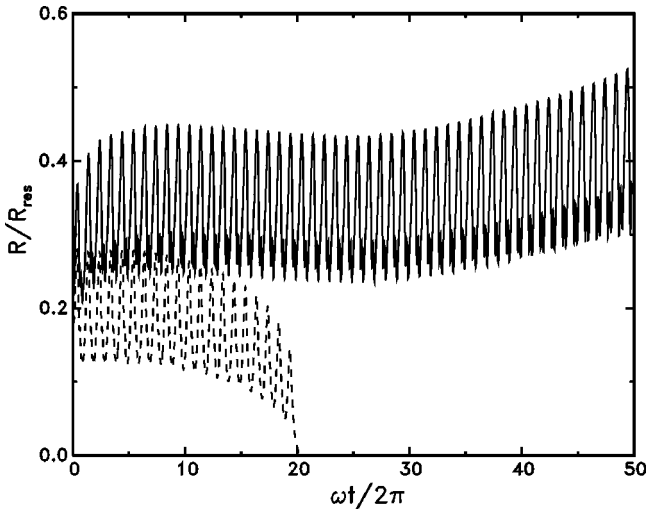


FIG. 8. Nondimensional radius versus time of bubbles translating with constant velocity $U=0.3$ m/s in water at 95 °C and 101.3 kPa static pressure in a 1 kHz, 30.39 kPa ($P_A/P_\infty=0.3$) sound field for two initial temperature distributions. For the dashed line $T(r,\theta,0)=T_\infty$, while, for the solid line, the initial temperature distribution is given by (10) with $\delta/R(0)=0.5$. The radius is normalized by the resonant radius of a stationary bubble equal to 2.367 mm; the initial radius is 0.5 mm. The initial phase of the sound field promotes expansion.

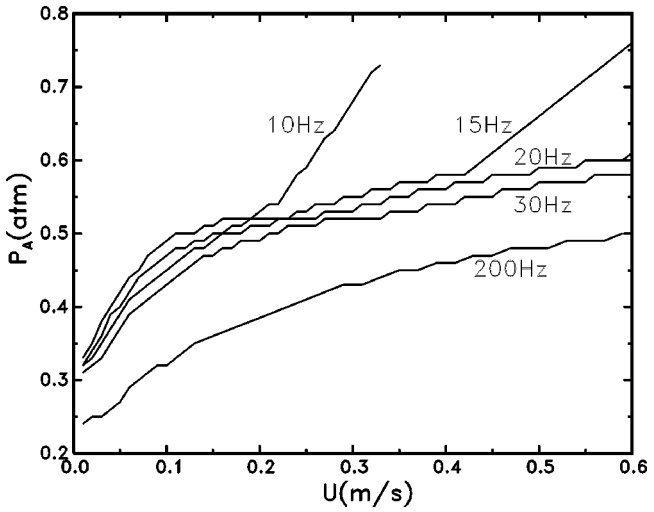


FIG. 9. Approximate location of the stability boundary for a vapor bubble translating with velocity U in a sound field with pressure amplitude P_A for different values of the sound frequency; the initial bubble radius is 0.5 mm. The liquid is water at $T_\infty = 95\,^\circ\mathrm{C}$ and $P_\infty = 101.3$ kPa and, at $t=0$, $T=T_\infty$, everywhere. The bubble grows by rectified diffusion for acoustic amplitudes above the lines. The initial phase of the sound is expansive [minus sign in Eq. (2)]. The curves are labeled by the sound frequency in Hz.

properties, it is somewhat futile to attempt a complete quantitative characterization of the stability limit. Thus, we only present a rather crude characterization of parameter space in Figs. 9 [$R(0)=0.5$ mm] and 10 [$R(0)=0.1$ mm] for an initial temperature equal to $T_\infty = 95\,^\circ\mathrm{C}$ everywhere and $P_\infty = 101.3$ kPa. The figures show approximate stability lines in the (P_A, U) plane for different sound frequencies: for acoustic amplitudes above these lines the bubble grows by rectified diffusion of heat, while it eventually collapses for smaller P_A's.

These results have been obtained with an initially expansive phase of the sound field [minus sign in Eq. (2)], as the bubble behavior is strongly affected by this phase. A typical near-threshold behavior with an expansive initial phase [minus sign in Eq. (2)] is shown in Fig. 11, where $R(0)$

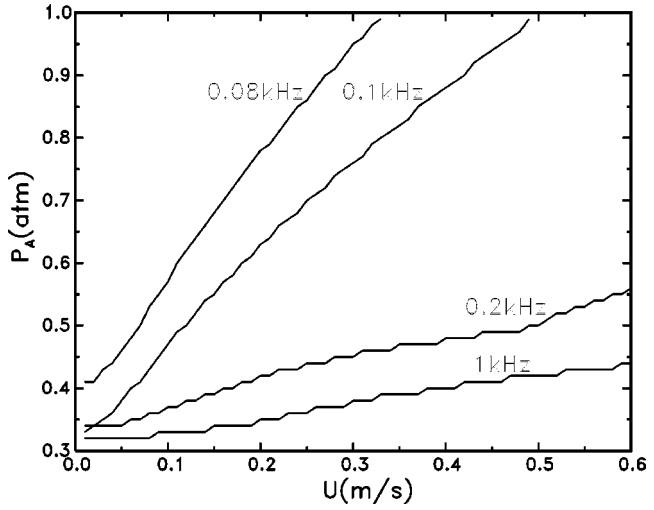


FIG. 10. As in Fig. 9 for $R(0)=0.1$ mm. The curves are labeled by the sound frequency in kHz.

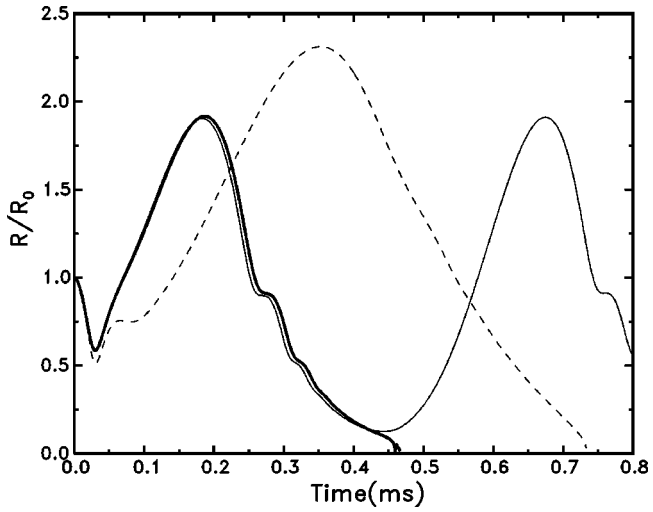


FIG. 11. Examples of the initial bubble behavior with an expansive sound phase [minus sign in Eq. (2)] with $R(0)=0.1$ mm, $P_A=30.39$ kPa, $U=0.2$ m/s and $T_{\mathrm{sat}}-T_\infty=5\,^\circ\mathrm{C}$; 1 kHz (dashed line), 2 kHz (thick solid line), 2.05 kHz. The threshold for growth lies between 2 and 2.05 kHz.

$=0.1$ mm, $P_A=30.39$ kPa, $U=0.2$ m/s and $T_{\mathrm{sat}}-T_\infty=5\,^\circ\mathrm{C}$. Both for 1 kHz (dashed line) and 2 kHz (thick solid line) the bubble condenses completely after a first expansion but, for 2.05 kHz, there is sufficient rectified heat transfer to stabilize the bubble and actually promote growth (the latter not shown in the figure). Figure 12 is for an initial compression phase [plus sign in Eq. (2)]: in the absence of sound (dash-and-dot line) or at 100 Hz (short dashes) a relatively slow condensation takes place; at the higher frequencies of 1 kHz (long dashes) and 3.8 kHz (thick solid line) condensation is more rapid, while at 3.9 kHz a strong growth is encountered.

It is interesting to briefly examine the consequences of the assumption of spatially uniform vapor conditions inside the bubble. In order to study this issue precisely, it would be necessary to account for the vapor flow induced by the translation of the bubble, which is a nontrivial task. Thus, we

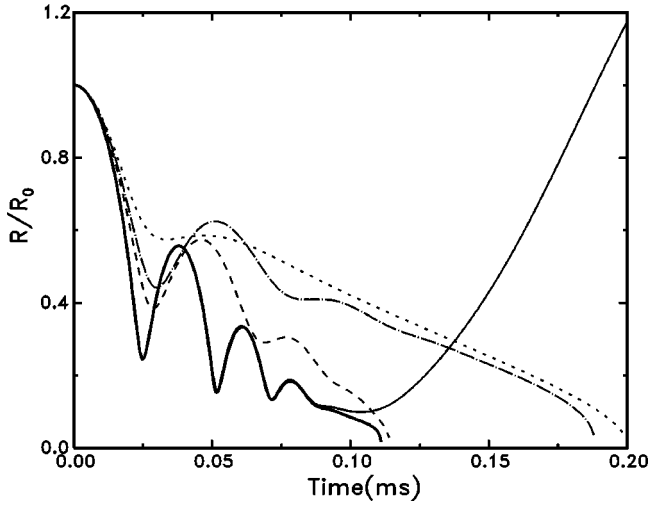


FIG. 12. Examples of the initial bubble behavior with a compressive sound phase [plus sign in Eq. (2)] with $R(0)=0.1$ mm, $P_A=30.39$ kPa, $U=0.2$ m/s and $T_{\mathrm{sat}}-T_\infty=5\,^\circ\mathrm{C}$; no sound (dash-and-dot line), 100 Hz (short dashes), 1 kHz (long dashes), 3.8 kHz (thick solid line), 3.9 kHz.

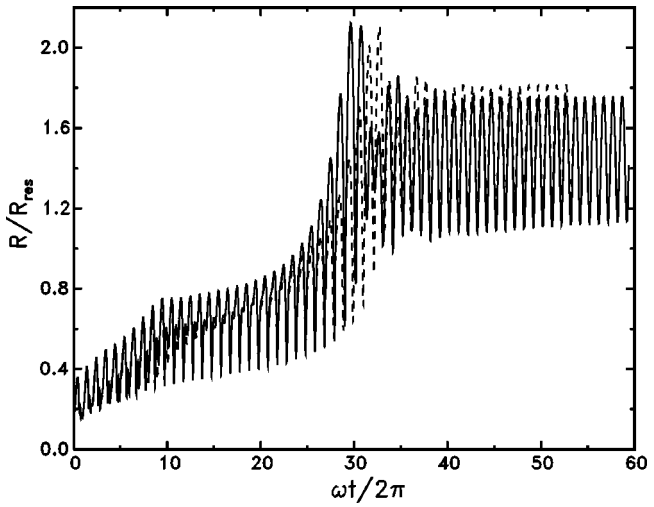


FIG. 13. Comparison between the growth rate of a stationary bubble modeled with a uniform (solid line) and nonuniform interior. The conditions are $R(0)=0.5$ mm, $P_A=40.9$ kPa $(P_A/P_\infty=0.4)$, $\omega/2\pi=1$ kHz, $T_\infty=95\,^\circ$C and $P_\infty=101.3$ kPa.

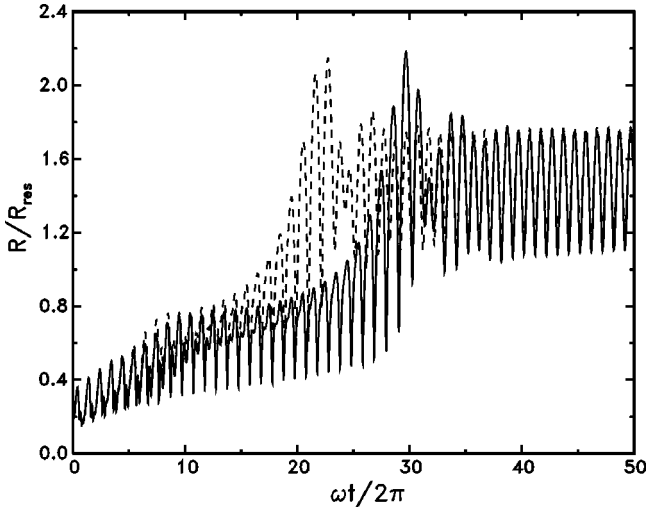


FIG. 14. Comparison between the growth rate of a stationary (solid line) and translating bubble. The conditions are $R(0)=0.5$ mm, $U=0.3$ m/s, $P_A=40.9$ kPa $(P_A/P_\infty=0.4)$, $\omega/2\pi=1$ kHz, $T_\infty=95\,^\circ$C and $P_\infty=101.3$ kPa.

content ourselves with considering the difference between homogeneous and nonhomogeneous conditions for a stationary bubble. Figure 13 shows a typical result. The conditions are: $R(0)=0.5$ mm, $P_A=40.9$ kPa $(P_A/P_\infty=0.4)$, $\omega/2\pi=1$ kHz, $T_\infty=95\,^\circ$C and $P_\infty=101.3$ kPa. The solid line is for spatially uniform vapor conditions, while the dashed line accounts for vapor nonuniformity according to our earlier model for a stationary bubble (Hao and Prosperetti, 1999). It is seen here that spatial nonuniformity gives rise to a somewhat slower growth. This result is not surprising: there is a tendency to develop stronger vapor temperature gradients near the interface during condensation than evaporation, because the departing vapor tends to ''wash them out'' [for explicit results on the temperature distribution inside the bubble, see Prosperetti and Hao (2002)]. Now, during condensation, the bubble surface temperature rises and, therefore, if the vapor temperature is allowed to be spatially nonuniform, there will be conduction of heat into the vapor, which favors condensation and hence heat loss. In any case, relatively large differences are only found near the resonance radius, where the oscillation amplitude is strongest. The vapor circulation neglected in this result would have the effect of mixing the vapor and may be expected, therefore, to decrease the consequences of spatial nonuniformity.

B. Long-time behavior

It would be wrong to conclude from the results shown so far that translation in a subcooled liquid always hinders bubble growth. An example is shown in Fig. 14, where the conditions are as in the previous figure. The solid line is for a stationary bubble, while the dashed line is for a bubble translating at 0.3 m/s. Convection increases the heat loss during compression as well as the heat gain during expansion: this result therefore implies that the latter effect is greater than the former. That this must be so can be understood from the approximate expression derived by Ruckenstein (1959) for the Nusselt number for flow past a sphere of constant diameter in the limit of large Péclet numbers (see also Legendre *et al.*, 1998):

$$\mathrm{Nu}=\frac{2Rh}{k}=2\sqrt{\frac{Pe}{\pi}}, \tag{11}$$

in which h is the heat transfer coefficient. The instantaneous heat transfer is proportional to hR^2, i.e., for constant U and temperature difference, to $R^{-1/2}R^2=R^{3/2}$, i.e., greater during expansion.

In all the cases considered so far the bubble velocity was held fixed. In reality, a translating and pulsating bubble will have a variable velocity due to changes in added mass, drag, and possibly buoyancy. Details of the velocity dependence on bubble volume depend in a subtle way on Reynolds and Weber numbers. A simple model, which we used in an earlier paper (Hao and Prosperetti, 2000), consists in assuming that the bubble remains spherical and the impulse is conserved:

$$\tfrac{2}{3}\pi\rho R^3U=\tfrac{2}{3}\pi\rho R(0)^3U(0). \tag{12}$$

Figure 15 compares a constant velocity case, $U=0.3$ m/s (dashed line), with a situation in which $U(0)=0.3$ m/s, but U at later times is found from Eq. (12). The pressure amplitude is $P_A=40.9$ kPa $(P_A/P_\infty=0.4)$, $\omega/2\pi=1$ kHz, $T_\infty=95\,^\circ$C, $P_\infty=101.3$ kPa, and $R(0)=0.5$ mm. The figure shows that allowing the velocity to vary slows down the growth of the bubble. This result can be understood by noting that, according to (12), the translational velocity increases when the bubble shrinks, which increases heat loss, while it decreases when the bubble expands, which is the phase of the acoustic cycle during which heat is gained.

In a saturated or superheated liquid, the growth of a bubble is always enhanced by motion relative to the liquid. An example is shown in Fig. 16, where the water is saturated at 101.3 kPa, the acoustic pressure amplitude is 30.39 kPa, the frequency is 1 kHz, and the initial bubble radius is 0.5 mm. The solid line is for a bubble translating at 0.3 m/s,

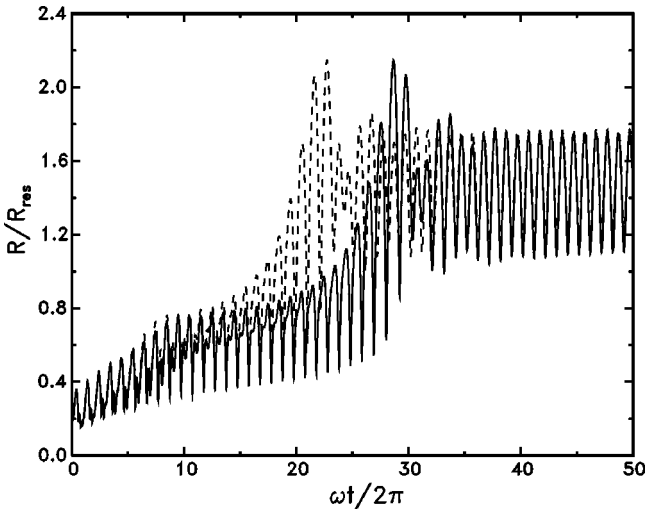


FIG. 15. Comparison between the growth rate of a bubble translating with fixed velocity $U=0.3$ m/s (dashed line), and a variable velocity calculated from (12). The conditions are $R(0)=0.5$ mm, $P_A=40.9$ kPa ($P_A/P_\infty=0.4$), $\omega/2\pi=1$ kHz, $T_\infty=95$ °C and $P_\infty=101.3$ kPa.

while the dashed line is for a stationary bubble; the radius is normalized by the resonant radius of a stationary bubble equal to 2.71 mm.

IV. DISCUSSION

The numerical results of the previous section, as well as the theory of rectified heat transfer with no translation (see, e.g., Gumerov, 2000), illustrate the complexity of the phenomenon under study. It is not possible to give a simple yet detailed physical interpretation of the results presented in the previous section: we shall content ourselves with some rather general arguments.

The most dangerous phase of the process for the survival of the bubble is the initial one and, accordingly, we limit ourselves to this stage. It may be noted, however, that, if the bubble velocity is large enough to move by about one diameter or more during one cycle of the sound, the surrounding temperature field will approximately always be close to the initial undisturbed distribution and, therefore, the bubble behavior during the initial stage may be expected to characterize also its long-term behavior. A bubble translating with a velocity U will move a distance equal to its diameter after a number N of acoustic cycles given by

$$N=\frac{\omega}{2\pi}\frac{2R}{U}. \tag{13}$$

For $R=0.5$ mm, $\omega/2\pi=1$ kHz, $N=1$ when U has the rather large value of 1 m/s. These conditions, therefore, are probably not frequently encountered.

In the first place, it is obvious that no rectified heat transfer is possible unless the liquid becomes temporarily superheated during the expansion phase of the sound. With the aid of the Clausius–Clapeyron relation, this condition gives

$$P_A \geqslant L\rho_V \frac{T_{\mathrm{sat}}-T_\infty}{T_{\mathrm{sat}}}. \tag{14}$$

For water at $T_\infty=95$ °C and $P_\infty=101.3$ kPa, we find $P_A \geqslant 18.2$ kPa. When this relation is barely satisfied, there is only a small fraction of the sound cycle during which the liquid is momentarily superheated. Clearly, if this interval of time is too short, the energy gain will not be sufficient to compensate the loss during the remainder of the cycle. Hence, one may expect that the estimate (14) will actually be lower than the true threshold; this conclusion is in agreement with the behavior shown in the previous figures.

Another condition for stability is that the bubble should not condense completely during the compression half-cycle of the sound, before the expansion phase renders the liquid temporarily superheated. To estimate this low-frequency limit, we may use a simplified form of the heat balance at the bubble surface:

$$q=\rho_V L\dot{R}, \tag{15}$$

where q is the liquid-side heat flux. This relation is based on the assumption that the vapor-side heat flux is negligible and

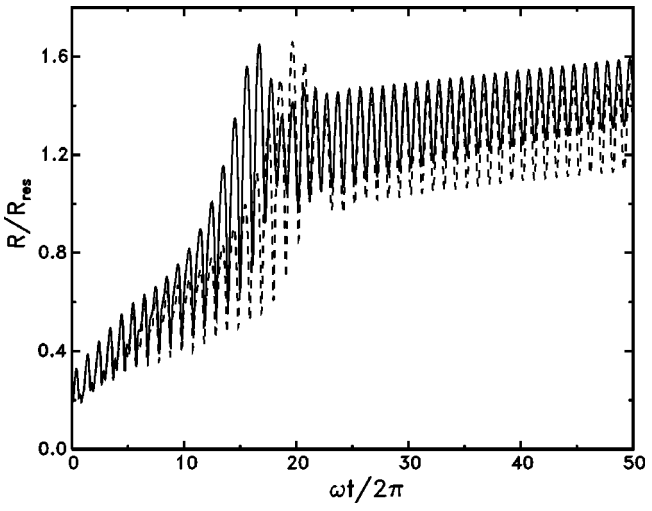


FIG. 16. Nondimensional radius versus time of a bubble translating at 0.3 m/s in water at 100 °C and 101.3 kPa in a 1 kHz, 30.39 kPa sound field (solid line); the dashed line is for a stationary bubble. The radius is normalized by the resonant radius of a stationary bubble equal to 2.71 mm; the initial radius is 0.5 mm.

TABLE I. Bubble collapse time for stationary vapor bubbles in water at 95 °C and 101.3 kPa as calculated numerically and as estimated from Eq. (17).

R_0 (μm)	Collapse time (ms) (Numerical)	Collapse time (ms) [Eq. (17)]	Ratio
50	0.0788	0.052	1.51
60	0.09	0.075	1.2
70	0.112	0.102	1.1
80	0.141	0.133	1.06
90	0.178	0.168	1.06
100	0.22	0.21	1.05
200	1.06	0.83	1.28
300	2.69	1.87	1.44
400	5.14	3.32	1.55
500	8.5	5.18	1.64
600	12.8	7.47	1.72
700	17.8	10.2	1.75
800	23.4	13.3	1.76

TABLE II. Comparison of the approximate expression (18) for the frequency above which a bubble would survive the first compression phase of the sound field, evaluated both in terms of Ja [Eq. (8)] and Ja_e [Eq. (19)], with numerical results obtained with an initially compressive phase of the sound field [plus sign in Eq. (2)]. The liquid is water at 95 °C and 101.3 kPa and the bubbles are stationary.

R_0 (mm)	P_A (kPa)	Numerical (First cycle)	Numerical (Growth)	Eq. (18) with Ja	Eq. (18) with Ja_e
0.5	20.26	100±10 Hz	3.7±0.1 kHz	98.8 Hz	288 Hz
0.5	22.29	101±1 Hz	125±5 Hz	98.8 Hz	306 Hz
0.5	30.39	130±10 Hz	130±10 Hz	98.8 Hz	374 Hz
0.1	20.26	3.2±0.1 kHz	17.5±0.5 kHz	2.47 kHz	7.2 kHz
0.1	22.29	3.3±0.1 kHz	4.2±0.1 kHz	2.47 kHz	7.8 kHz
0.1	30.39	3.6±0.1 kHz	3.6±0.1 kHz	2.47 kHz	9.3 kHz

that all the heat conducted into the liquid accounts for the latent heat released by the saturated vapor condensing at the bubble surface. If, following Florschuetz and Chao (1965), we use for q the estimate $q \simeq k(T_{\text{sat}}-T_\infty)/\sqrt{\pi D t}$, we find

$$\frac{R(t)}{R(0)} = 1 - 2\text{Ja}\sqrt{\frac{Dt}{\pi R^2(0)}}, \tag{16}$$

from which, upon setting $R(t_c)=0$, one deduces the following estimate for the characteristic bubble collapse time t_c:

$$t_c = \frac{\pi}{4\text{Ja}^2}\frac{R^2(0)}{D}. \tag{17}$$

As shown in Table I, while this estimate of the collapse time is of the right order of magnitude, depending on specific conditions, numerically it may be off by as much as a factor of 2. With this caveat in mind, if we set $t_c \geqslant \pi/\omega$ for stability, we find

$$\frac{\omega}{2\pi} \geqslant \frac{2D}{\pi R^2(0)}\text{Ja}^2. \tag{18}$$

As the static pressure in the liquid changes due to the presence of the sound field, the instantaneous Jakob number also changes; in particular, at the pressure maximum, it has an ''effective'' value given by

$$\text{Ja}_e = \rho c_{pL}\frac{T_{\text{sat}}(P_\infty + P_A) - T_\infty}{L\rho_{V,\text{sat}}(P_\infty + P_A)}, \tag{19}$$

where the variation of L, ρ, and c_{pL} has been neglected for simplicity.

A comparison of (18), evaluated both in terms of Ja and Ja_e, with some numerical results is given in Table II. It should be noted that these results have been obtained with an initially compressive phase of the sound field [i.e., the plus sign in Eq. (2)] because otherwise the correct value of the radius to use would be that at the end of the first expansion, which it is difficult to estimate. It is seen that, in general, the frequency threshold lies somewhere between the two predictions of Eq. (18). This estimate refers to the survival of the bubble during the first sound cycle but, in the table, we also show the numerically determined threshold for a sustained growth of the bubble. The substantial difference between the two numerical thresholds that appears in some cases arises when the acoustic pressure does not exceed by much the very low estimate of Eq. (14).

The previous estimate is applicable at low translational velocities, when the heat loss from the bubble is dominated by conduction. When convection dominates, we may estimate q from the expression (11) of the Nusselt number. Upon setting in (15)

$$q = \text{Nu}\,k\frac{T_{\text{sat}}-T_\infty}{2R} = 2k(T_{\text{sat}}-T_\infty)\sqrt{\frac{U}{2\pi DR(t)}}, \tag{20}$$

and integrating, we find

$$\frac{R(t)}{R(0)} = \left(1 - 3\text{Ja}\sqrt{\frac{DU}{2\pi R^3(0)}}t\right)^{2/3}, \tag{21}$$

from which

$$t_c' = \frac{1}{3\text{Ja}}\sqrt{\frac{2\pi R^3(0)}{DU}} = \frac{R^2(0)}{3\text{Ja}D}\sqrt{\frac{2\pi}{\text{Pe}(0)}}. \tag{22}$$

The ratio of the two collapse times is

$$\frac{t_c}{t_c'} = \frac{3}{8\text{Ja}}\sqrt{\pi\text{Pe}(0)}, \tag{23}$$

TABLE III. The numerically calculated collapse time of bubbles translating with constant velocity in water at 95 °C and 101.3 kPa in the absence of a sound field compared with the estimate (22). Note that for smaller bubbles at low velocity the Péclet number is too small to justify use of (22).

	Collapse time (ms)					
	Numerical			Eq. (22)		
R_0 (mm)	0.2 m/s	0.4 m/s	0.6 m/s	0.2 m/s	0.4 m/s	0.6 m/s
0.05	0.070	0.063	0.055	0.11	0.076	0.062
0.075	0.10	0.09	0.08	0.20	0.14	0.11
0.10	0.18	0.15	0.13	0.30	0.21	0.17
0.25	1.0	0.73	0.60	1.2	0.85	0.69
0.50	3.0	2.2	1.8	3.4	2.4	2.0
0.75	5.6	4.0	3.2	6.2	4.4	3.6

TABLE IV. Comparison of the estimate (25) for the frequency above which a translating bubble would survive the first compression phase of the sound field, with numerical results obtained with an initially compressive phase of the sound field [plus sign in Eq. (2)]. For each bubble radius, the numerical result in the first line is the threshold for the bubble to survive the first compression induced by the sound field; that in the second line is the threshold for bubble growth by rectified heat transfer. The liquid is water at 95 °C and 101.3 kPa and the acoustic pressure amplitude is 30.39 kPa.

	Threshold (kHz)					
	Numerical (First cycle/growth)			Estimate from Eq. (25), with Ja_e/Ja		
R_0 (mm)	0.2 m/s	0.4 m/s	0.6 m/s	0.2 m/s	0.4 m/s	0.6 m/s
0.5	0.275±0.025/ 0.700±0.050	0.425±0.025/ 1.45±0.050	0.525±0.025/ 1.85±0.050	0.290/ 0.150	0.410/ 0.212	0.502/ 0.260
0.1	3.85±0.05/ 3.85±0.05	4.45±0.05/ 5.35±0.05	5.45±0.05/ 8.5±0.05	3.24/ 1.68	4.58/ 2.37	5.60/ 2.90

and we may therefore conclude that convection dominates for Pe(0) greater than the critical value

$$\mathrm{Pe}_c = \frac{64}{9\pi}\mathrm{Ja}^2. \tag{24}$$

The collapse time estimated from (22) is compared with some numerical results in Table III. The Péclet number for $R(0)=0.05$ mm and $U=0.2$ m/s is 119, well below the estimate (24) of Pe_c which, for this subcooling, is 523. This case is therefore conduction dominated and, not surprisingly, the estimate based on (22) is rather poor. As Pe increases, however, the table shows that there is a good agreement with the numerical results.

The same argument as before applied to (22) now gives

$$\frac{\omega}{2\pi} \geqslant \frac{3D\mathrm{Ja}}{4R^2(0)}\sqrt{\frac{\mathrm{Pe}(0)}{\pi}}. \tag{25}$$

This estimate is compared with some numerical results in Table IV whereas, in Table III, separate thresholds are shown for bubble survival during the first compressive half-cycle of the sound and for bubble growth by rectified heat transfer. The first threshold is estimated rather well by (25). Upon comparing the results of Tables II and IV, it is seen that, for a stationary bubble, the stability threshold is better predicted by the Jacob number (8) while the effective Jacob number (19) seems to be more relevant for a translating bubble. This reflects the greater instability of the latter situation.

The stability boundaries in the (P_A, U) plane shown in Figs. 9 and 10 must correspond to conditions such that the rectified influx of heat due to the sound-induced oscillations balances the heat lost through convection. When conditions are such that the bubble lasts many cycles, the heat exchange with the liquid may approximately be considered as governed by a Nusselt number of the form (11) averaged over one cycle. In these conditions since, to lowest order, rectified heat flux into the bubble is a second order effect and, therefore, proportional to P_A^2 (see, e.g., Wang, 1974; Gumerov, 2000), one may expect from (11) a dependence of the stability boundary on velocity of the form $P_A^2 \propto \sqrt{U}$, i.e., $P_A \propto U^{1/4}$. Figures 17 and 18 show the same results as Figs. 9 and 10 replotted in this way. At very low translational velocity, conduction dominates, the threshold becomes independent of U and, therefore, $P_A/U^{1/4}$ diverges. To explain the high-velocity behavior of some of the curves, we observe that proportionality of rectified heat transfer to P_A^2 can only be assumed when the bubble lasts long enough to take time averages. If the bubble only lasts a few cycles, this estimate is inapplicable and, indeed, we see in Figs. 17 and 18 that, for some frequencies, the proportionality of P_A to $U^{1/4}$ is violated at higher velocities. Nevertheless, we see that there are velocity and frequency ranges where the prediction $P_A \propto U^{1/4}$ is substantiated by the numerical results.

It is evident from many of the examples shown in the previous figures that the rectified diffusion of heat into the bubble becomes much smaller past the resonance radius. It is therefore conceivable that a bubble for which rectified diffusion is just about strong enough to make it grow below resonance would start to condense past resonance then grow again once it has shrunk below the resonance radius, and so on. We have searched numerically for this behavior but have concluded that, if it exists, it requires such a fine tuning of conditions as to be of little practical interest; therefore, we have not pursued the matter further.

V. CONCLUSIONS

In view of the dependence of heat transfer on convection, it is not surprising to find a strong effect of translatory

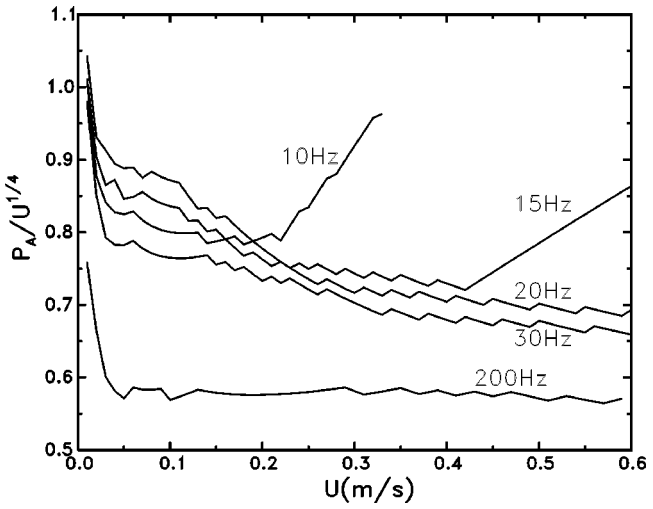


FIG. 17. The data of Fig. 9 [$R(0)=0.5$ mm] replotted as $P_A/U^{1/4}$ vs U according to the argument given at the end of Sec. IV. The horizontal or near-horizontal portion of the curves corresponds to situations in which convection dominates the heat transfer from the bubble. The curves are labeled by the sound frequency in Hz.

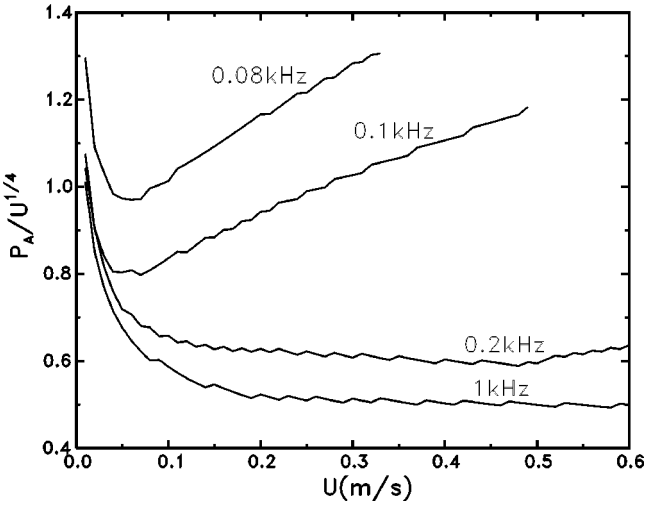


FIG. 18. The data of Fig. 10 [$R(0)=0.1$ mm] replotted as $P_A/U^{1/4}$ vs U according to the argument given at the end of Sec. IV. The horizontal or near-horizontal portion of the curves corresponds to situations in which convection dominates the heat transfer from the bubble. The curves are labeled by the sound frequency in kHz.

motion on the process of rectified heat transfer into a vapor bubble. The details of the process are complex and depend on the initial liquid temperature distribution around the bubble, in addition to the bubble radius, velocity, and sound field characteristics. This circumstance prevents one from making simple statements of general validity. If the liquid is initially subcooled and at a uniform temperature, convection may destabilize the bubble. For example, in water at 95 °C and at normal gravity, a 1-mm-diam bubble in a 1-kHz, 30.39-kPa sound field would grow if moving at 0.2 m/s, while it would collapse at 0.3 m/s. In a stronger sound field, 35.46 kPa, the bubble would be stabilized even at this higher velocity, which is of the order of the buoyant rise velocity of such a bubble in normal gravity. Qualitatively, an interesting finding of this study is that the boundary between stable and unstable bubbles occurs in parameter ranges readily encountered in practice.

These effects may be useful for bubble management in several conditions. For example, in situations where large bubbles may be easier to control than smaller ones, one may want to promote growth, which facilitates coalescence under the action of pressure-radiation forces (Hao *et al.*, 2001). Our results show that this is possible even in the presence of subcooled convection and give an indication as to the necessary acoustic parameters.

ACKNOWLEDGMENTS

The authors are grateful to Dr. N. Gumerov for some useful exchanges in the course of this work and to the reviewers for some useful suggestions. This study has been supported by NASA under Grant No. NAG3-1924.

[1]Dr. Gumerov (private communication) points out that an additional mechanism might be the different curvature of the saturation and adiabatic curves: the temperature oscillations have a shift due to the nonlinearity of these curves which causes an effective liquid-vapor temperature difference to arise.

[2]A 0.5-mm-radius bubble in water has a rise velocity of the order of 0.3 m/s under normal buoyancy; larger or smaller velocities are possible due to acceleration of the container, different gravitational fields, the presence of shock waves, and others.

[3]We loosely refer to a bubble prevented from collapsing as being ''stabilized'' by the action of the sound field. The true equilibrium state corresponds to a condition in which mass loss by condensation balances gain by rectified diffusion; it is clear, however, that this equilibrium is unstable.

Akulichev, V. A., Alekseev, V. N., and Yushin, V. P. (**1979**). ''Growth of vapor bubbles in an ultrasonic field,'' Sov. Phys. Acoust. **25**, 453–457.

Ervin, J. S., Merte, H., Jr., Keller, R. B., and Kirk, K. (**1992**). ''Transient pool boiling in microgravity,'' Int. J. Heat Mass Transf. **35**, 659–674.

Florschuetz, L. W., and Chao, B. T. (**1965**). ''On the mechanics of vapor bubble collapse,'' J. Heat Transfer **87**, 209–220.

Fyrillas, M. M., and Szeri, A. J. (**1994**). ''Dissolution or growth of soluble spherical oscillating bubbles,'' J. Fluid Mech. **277**, 381–407.

Gumerov, N. A. (**2000**). ''Dynamics of vapor bubbles with non-equilibrium phase transitions in isotropic acoustic fields,'' Phys. Fluids **12**, 71–88.

Hao, Y., and Prosperetti, A. (**1999**). ''The dynamics of vapor bubbles in acoustic pressure fields,'' Phys. Fluids **11**, 2008–2019.

Hao, Y., and Prosperetti, A. (**2000**). ''The collapse of vapor bubbles in a spatially non-uniform flow,'' Int. J. Heat Mass Transf. **43**, 3539–3550.

Hao, Y., Oğuz, H. N., and Prosperetti, A. (**2001**). ''Pressure-radiation forces on pulsating vapor bubbles,'' Phys. Fluids **13**, 1167–1177.

Khabeev, N. S. (**1976**). ''Heat transfer and phase-transition effects in the oscillation of vapor bubbles,'' Sov. Phys. Acoust. **21**, 501–505.

Legendre, D., Borée, J., and Magnaudet, J. (**1998**). ''Thermal and dynamic evolution of a spherical bubble moving steadily in a superheated or subcooled liquid,'' Phys. Fluids **10**, 1256–1272.

Oka, T., Abe, Y., Tanaka, K., Mori, Y. H., and Nagashima, A. (**1992**). ''Observational study of pool boiling under microgravity,'' JSME Int. J.: Series B **35**, 280–286.

Patel, G. M., Nicholas, R. E., and Finch, R. D. (**1985**). ''Rectified heat transfer in vapor bubbles,'' J. Acoust. Soc. Am. **78**, 2122–2131.

Prosperetti, A., and Hao, Y. (**2001**). ''Vapor bubbles in flow and acoustic fields,'' in *Free-Surface Flows*, edited by A. C. King and Y. Shikhmurzaev (Kluwer, Dordrecht), pp. 249–256.

Prosperetti, A., and Hao, Y. (**2002**). ''Vapor bubbles in flow and acoustic fields,'' in *Annals of the New York Academy of Sciences*, vol. 974, *Microgravity Transport Processes in Fluid, Thermal, Biological, and Materials, Sciences*, edited by S. S. Sadhal, in press.

Ruckenstein, E. (**1959**). ''On heat transfer between vapour bubbles in motion and the boiling liquid from which they are generated,'' Chem. Eng. Sci. **10**, 22–30.

Sitter, J. S., Snyder, T. J., Chung, J. N., and Marston, P. L. (**1998a**). ''Acoustic field interaction with a boiling system under terrestrial gravity and microgravity,'' J. Acoust. Soc. Am. **104**, 2561–2569.

Sitter, J. S., Snyder, T. J., Chung, J. N., and Marston, P. L. (**1998a**). ''Terrestrial and microgravity pool boiling heat transfer from a wire in an acoustic field,'' Int. J. Heat Mass Transf. **41**, 2143–2155.

Wang, T. (**1974**). ''Rectified heat transfer,'' J. Acoust. Soc. Am. **56**, 1131–1143.

Validity of the sonar equation and Babinet's principle for scattering in a stratified medium

Purnima Ratilal, Yisan Lai, and Nicholas C. Makris[a)]
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 27 July 2001; revised 8 April 2002; accepted 12 June 2002)

The sonar equation rests on the assumption that received sound pressure level after scattering can be written in decibels as a sum of four terms: source level, transmission loss from the source to the target, target strength, and transmission loss from the target to the receiver. This assumption is generally not valid for scattering in a shallow water waveguide and can lead to large errors and inconsistencies in estimating a target's scattering properties as well as its limiting range of detection. By application of coherent waveguide scattering theory, the sonar equation is found to become approximately valid in a shallow water waveguide when the object's complex scatter function is roughly constant over the equivalent horizontal grazing angles $\pm\Delta\psi$ spanned by the dominant waveguide modes. This is approximately true (1) for all objects of spatial extent L and wavelength λ when $2\Delta\psi<\lambda/2L$ and (2) for spheres and certain other rounded objects in nonforward scatter azimuths, even when (1) does not hold. The sonar equation may be made valid by lowering the active frequency of operation in a waveguide. This is often desirable because it greatly simplifies the analysis necessary for target classification and localization. Similarly, conditions are given for when Babinet's principle becomes approximately valid in a shallow water waveguide. [DOI: 10.1121/1.1499136]

PACS numbers: 43.30.Ft, 43.30.Gv, 43.30.Vh [DLB]

I. INTRODUCTION

The sonar equation is the most widely used analytical tool in applications of active sonar.[1–5] It is typically employed to estimate a target's scattering properties and limiting range of detection.[1] The sonar equation, in its active form for an omnidirectional source, rests on the assumption that received sound pressure level in decibels can be written as a sum of four terms: source level, transmission loss from the source to the target, target strength, and transmission loss from the target to the receiver. This assumption has two related implications: (1) that propagation and scattering effects are completely factorable from each other, and (2) that a linear combination of the incoherent quantities, target strength, transmission loss, and source level completely specifies the sound pressure level at the receiver.

When this assumption is axiomatically adopted in an ocean waveguide, fundamental inconsistencies can occur when experimental data is examined. For example, the target strength of an object, which should be invariant, can vary significantly with range when experimentally estimated in a range-independent environment where the direct return arrives together with multiple returns from the waveguide boundaries. This has been noted in the classic text, *Physics of Sound in the Sea*.[2] Regardless of this inconsistency, target strength is still used in that text to describe the scattering properties of an object because a more fundamental approach had not presented itself, just as it is still used today by many practitioners of ocean acoustics.

Our goal in the present paper is to (1) demonstrate that the assumption that the sonar equation rests on and its implications are not generally valid for scattering in a shallow water waveguide, and (2) provide conditions necessary for the sonar equation to become valid in a shallow water waveguide.[6] In the process we show that the invariant scattering properties of an object in a waveguide cannot generally be described by target strength, an incoherent quantity, but rather require a coherent representation that arises from the fundamental waveguide scattering theory[7,8] upon which our analysis is based.

First developed in World War II,[1,2] the sonar equation, analogous to the radar equation, *is only valid when propagation and scattering dependencies are approximately separable.* For example, given an omnidirectional point source and receiver in free space, propagation and scattering are separable when the source and receiver are in the farfield of the target, where the incident as well as the scattered wave may be approximated as planar. The sonar equation has a long history of legitimate usage in deep water,[1,2] where these free-space conditions are effectively achieved in many practical scenarios due to the significant time separation that often occurs between direct and surface reflected arrivals and the adiabatic nature of refraction in the ocean.

In continental shelf environments, referred to as shallow water waveguides in ocean acoustics, multiple reflections from the surface and bottom typically overlap and *coherently* interfere with each other and the direct arrival. No unique incident and scattered angle exists. To understand the implications for scattering, it is convenient to decompose the incident field at a target in a waveguide from a farfield point

a)Electronic mail: makris@mit.edu

source, into modal plane waves. Each incident plane wave arrives with a specific elevation angle from the azimuth of the farfield source and then scatters *coherently* from the target into outgoing plane waves in all elevation and azimuthal angles. The target affects both the amplitude and phase of each scattered plane wave. *This phase change cannot be described by an incoherent quantity such as target strength*. At a farfield receiver, the scattered field is the *coherent* sum of all scattered plane waves from all incident plane waves. Just as waveguide propagation models must account for the *coherent* interference of multiple arrivals from a source to receiver to accurately determine transmission loss, waveguide scattering models must account for the *coherent* interference of all scattered waves from every wave incident on the target. Propagation and scattering are in this way *coherently* convolved for objects submerged in a waveguide.

To establish when the sonar equation can be applied in a shallow water waveguide, we calculate the scattered field from a variety of target types in various shallow water waveguides using a physics-based waveguide scattering model[7,8] that takes into account the coupling between propagation and scattering. We then compare the results to those predicted by the sonar equation. The waveguide scattering model, based on Green's Theorem, expresses the scattered field in terms of normal modes, convenient for long-range propagation, and the plane wave scatter function of the object. The only assumptions needed for the waveguide scattering model to be valid are that the propagation medium is horizontally stratified and range independent, multiple scattering between the object and waveguide boundaries is negligible, the object lies within a constant sound speed layer, and the range from the object to the source or receiver is large enough that the scattered field can be expressed as a linear function of the object's plane wave scatter function.

Under these conditions, the plane wave scatter function of the object, which depends on absolute object orientation and direction of both the incident and scattered plane wave, is the invariant quantity that describes the scattering properties of an object in a waveguide. The scatter function is a coherent quantity. The object's incoherent target strength is simply $20\log$ of the scatter-function-magnitude-to-wave number ratio. Target strength then only contains the amplitude but not the essential phase information necessary to describe the scattering process in a waveguide.

We show analytically that if the scatter function of the object is approximately constant over the equivalent angles spanned by the waveguide modes for the given bistatic geometry, the scatter function of the object, which effectively couples the modes of the incident and scattered field, can be factored with little error. This leads to an approximation for the sound pressure level of the scattered field that is the sonar equation. Many rounded objects, such as spheres and certain spheroids, exhibit this behavior in nonforward scatter. Flat homogeneous objects, such as plates and disks, are the most highly directional convex targets. These have nonuniform scatter functions with strong main lobes in the forward and backscatter directions of diffraction-limited angular width λ/L, for $\lambda/L \ll 1$, where L is the object's length and λ the wavelength. Other targets that are nonconvex or inhomogeneous, for example, may have a narrower scatter function main lobe due to interference from distinct parts of the target. In the limit, the scatter function of an object consisting solely of two point scatterers separated by L has the narrowest main lobe of angular width $\lambda/2L$. When the angular width of the object's scatter function main lobe is much smaller than that spanned by the propagating modes $\pm\Delta\psi$ about the horizontal, which is often limited by the critical angle ψ_c of the seabed beyond a few water column depths in range, the modes of the waveguide are scattered nonuniformly. This leads to strong coupling between propagation and scattering in both the forward and backscatter azimuths. The sonar-equation approximation is found to be in error, often by tens of decibels, when applied to such highly directional targets in shallow water waveguides.

These findings explain the physical basis for the discrepancy noted in Ref. 9 between the sonar equation and a waveguide scattering model. Some special cases were previously noted. For example, Ingenito pointed out that propagation and scattering become factorable in a waveguide that supports only a single mode.[7] Makris noted that this factorization is possible for compact objects, i.e., those with $ka = \pi L/\lambda \ll 1$.[9] We note that the sonar equation is always valid for compact pressure-release objects since the scattered field is effectively omnidirectional, but is more approximate for rigid compact objects since their scattered fields always maintain some directionality as ka decreases.

As a general conclusion, we find that *the sonar equation is valid when the target's scatter function is roughly constant over the equivalent horizontal grazing angles $\pm\Delta\psi$ spanned by the dominant waveguide modes*. This is approximately true (1) for all objects when $2\Delta\psi < \lambda/2L$ and (2) for spheres and certain other rounded objects in nonforward scatter azimuths even when (1) does not hold. For homogeneous convex objects condition (1) is the less stringent $2\Delta\psi < \lambda/L$. A quantitative definition of $\Delta\psi$ is provided in Sec. III B. It should be noted that $\Delta\psi$ is range dependent in realistic ocean waveguides and is not necessarily equal to the critical angle ψ_c. This is true even in Pekeris waveguides at small ranges where the leaky modes with elevation angles larger than ψ_c make substantial contribution to the scattered field and at long ranges where modal stripping reduces the dominant modes to elevation angles much smaller than ψ_c.

This conclusion is significant because, in an active scenario, the sonar operator has the ability to lower the frequency of transmission until the target's scatter function becomes approximately constant over $\pm\Delta\psi$. The sonar equation then becomes a valid approximation when $f < c/(4L\,\Delta\psi)$. Operating in this frequency regime is desirable because when the sonar equation is valid, only a single parameter is necessary to characterize the scattering properties of the target for that measurement. This greatly simplifies target classification by making the classic approach of estimating a single-parameter target strength valid in a shallow water waveguide. It also simplifies other problems such as estimating target depth in a waveguide. When the sonar equation is not valid, the problem of classifying the target becomes much more complicated. Up to $2(2N)^2$ parameters would be necessary to characterize the scattering properties

of the target, for a waveguide that supports N modes, because the amplitude and phase of the object's scatter function would have to be determined for each incident and scattered pairing of each mode's equivalent up- and downgoing plane wave elevation angles.

Babinet's principle maintains that the forward scattered fields of impenetrable objects that have identical projected areas with respect to a given incident plane wave in free space are asymptotically equal for large ka.[10,12,13] This also holds true for some penetrable objects.[10] For an object submerged in a waveguide, the incident and scattered fields are often characterized by a wide angular spectrum of plane waves. Despite this difference between waveguide and free space scattering, simulations in several typical shallow water waveguides with a variety of targets types show that Babinet's principle can hold approximately in a waveguide *in the forward-scatter azimuth* if the equivalent propagation angles of the modes are sufficiently close to horizontal, as if often the case after long-range propagation in lossy media. By Babinet's principle, objects that are large compared to the wavelength cast the same free space shadow as flat objects with the same projected area. Since flat objects of high ka are the most directional, the sonar equation approximation breaks down rapidly in a shallow water waveguide as ka increases beyond unity for scattering in the *forward* azimuth for all object shapes, including spheres. Extreme caution should then be used in applying the sonar equation in forward scatter.

In Sec. II, we describe the waveguide scattering model that takes into account the coupling between propagation and scattering in a shallow water waveguide. A detailed derivation of the sonar equation from Green's Theorem is provided in Sec. III to show when propagation and scattering become decoupled and when incoherent target strength is sufficient to describe the scattering properties of an object in a waveguide. Illustrative examples are presented in Sec. IV. Babinet's principle and issues involved with applying it in a waveguide are discussed in Sec. V.

II. WAVE-THEORETIC MODEL FOR 3-D SCATTERING FROM AN OBJECT OF ARBITRARY SHAPE IN A STRATIFIED MEDIUM

We adopt a wave-theoretic normal mode model based on Green's Theorem for the field scattered by an object in a stratified medium, following Refs. 7 and 8. In the formulation, the origin of the coordinate system is placed at the object centroid. The source coordinates are defined by $\mathbf{r}_0 = (x_0, y_0, z_0)$ and the receiver coordinates by $\mathbf{r} = (x, y, z)$, where the positive z axis points downward and normal to the interface between horizontal strata. Spatial cylindrical (ρ, ϕ, z) and spherical systems (r, θ, ϕ) are defined by $x = r \sin\theta \cos\phi$, $y = r \sin\theta \sin\phi$, $z = r \cos\theta$, and $\rho = x^2 + y^2$. The horizontal grazing angle is $\psi = \pi/2 - \theta$. The horizontal and vertical wave number components for the nth mode are, respectively, $\xi_n = k \sin\alpha_n$ and $\gamma_n = k \cos\alpha_n$, where α_n is the elevation angle of the mode measured from the z axis. Here $0 \leqslant \alpha_n \leqslant \pi/2$ so that the down- and upgoing plane wave components of each mode will then have elevation angles α_n and $\pi - \alpha_n$, respectively. The corresponding vertical wave number of the down and upgoing components of the nth mode are γ_n and $-\gamma_n$, respectively, where $\mathrm{Re}\{\gamma_n\} \geqslant 0$. Moreover, $k^2 = \xi_n^2 + \gamma_n^2$, and the wave number magnitude k equals the angular frequency ω divided by the sound speed c in the object layer. For economy, the notation of Ref. 11 is used here and in the remainder of this article. Figure 2 of Ref. 11 shows the geometry of spatial and wave number coordinates.

The spectral component of the scattered field from the object at the origin for a source at $\mathbf{r}_0$ and a receiver at $\mathbf{r}$ is

$$\begin{aligned}\Phi_s(\mathbf{r}|\mathbf{r}_0) = \mathcal{A} \sum_{m=1}^{M_{\max}} \sum_{n=1}^{M_{\max}} \frac{(4\pi)^2}{k} \\ \times [A_m(\mathbf{r}) A_n(\mathbf{r}_0) S(\pi - \alpha_m, \beta; \alpha_n, \beta_i) \\ - B_m(\mathbf{r}) A_n(\mathbf{r}_0) S(\alpha_m, \beta; \alpha_n, \beta_i) \\ - A_m(\mathbf{r}) B_n(\mathbf{r}_0) S(\pi - \alpha_m, \beta; \pi - \alpha_n, \beta_i) \\ + B_m(\mathbf{r}) B_n(\mathbf{r}_0) S(\alpha_m, \beta; \pi - \alpha_n, \beta_i)], \end{aligned} \tag{1}$$

where

$$\begin{aligned} A_m(\mathbf{r}) &= \frac{i}{d(0)} (8\pi\xi_m\rho)^{-1/2} u_m(z) N_m^{(1)} e^{i(\xi_m\rho + \gamma_m D - \pi/4)}, \\ B_m(\mathbf{r}) &= \frac{i}{d(0)} (8\pi\xi_m\rho)^{-1/2} u_m(z) N_m^{(2)} e^{i(\xi_m\rho - \gamma_m D - \pi/4)}, \\ A_n(\mathbf{r}_0) &= \frac{i}{d(z_0)} (8\pi\xi_n\rho_0)^{-1/2} u_n(z_0) N_n^{(1)} e^{i(\xi_n\rho_0 + \gamma_n D - \pi/4)}, \\ B_n(\mathbf{r}_0) &= \frac{i}{d(z_0)} (8\pi\xi_n\rho_0)^{-1/2} u_n(z_0) N_n^{(2)} e^{i(\xi_n\rho_0 - \gamma_n D - \pi/4)}, \end{aligned} \tag{2}$$

are the incident and scattered down- and upgoing plane wave amplitudes in the layer of the object, $\mathcal{A}$ is the source amplitude, D is the depth of the object center from the sea surface, $d(z)$ is the density at depth z, $u_n(z)$ are the mode functions, $S(\alpha, \beta, \alpha_i, \beta_i)$ is the object's plane wave scatter function, and $M_{\max}$ is the mode number at which the series can be truncated and still accurately represent the field. The definition of the plane wave scatter function here follows that defined in Ref. 8, where the incident plane wave on the object is described in terms of the direction it goes to so that for forward scatter in free space $\alpha = \alpha_i$, $\beta = \beta_i$. The product of $e^{-i2\pi ft}$ and the right hand side of Eq. (1) yields the time-harmonic scattered field. The mode functions are normalized according to

$$\delta_{nm} = \int_{-D}^{\infty} \frac{u_m(z) u_n^*(z)}{d(z)} dz, \tag{3}$$

and can be expressed in the layer of the object as

$$u_n(z) = N_n^{(1)} e^{i\gamma_n(z+D)} - N_n^{(2)} e^{-i\gamma_n(z+D)}, \tag{4}$$

where $N_n^{(1)}$ and $N_n^{(2)}$ are normalization constants.

A more general expression than Eqs. (1)–(4), for the scattered field from an arbitrarily shaped object in a stratified medium, is given in Refs. 9 and 11 in terms of wave number integrals. A number of assumptions have to be satisfied for the above formulation to be valid, as noted in Ref. 8. In particular, multiple scattering between the object and wave-

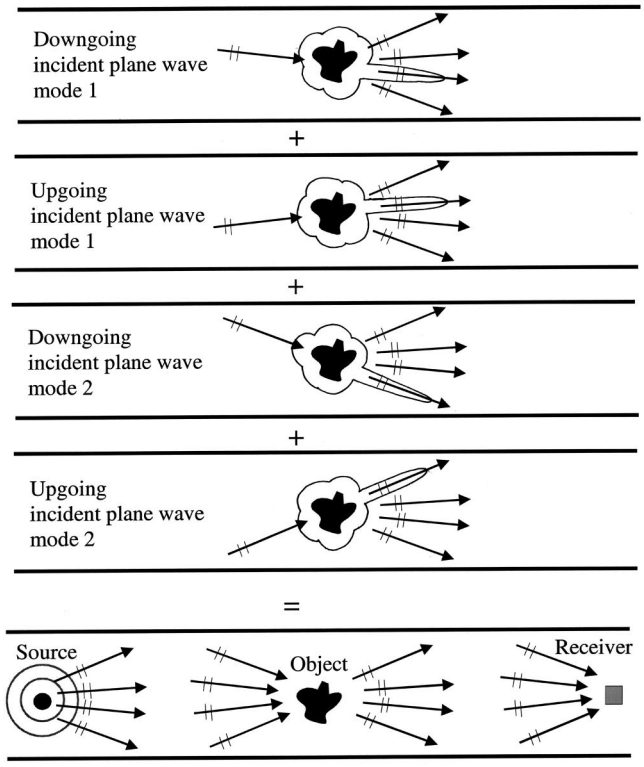


FIG. 1. (a) Scattering in a waveguide with only two modes. Each mode is composed of a downgoing and an up-going plane wave. Each incoming plane wave is scattered by the object into various outgoing plane waves. The scattered fields from each incident plane wave are coherently superposed to form the total scattered field at the receiver. (b) Scattering in free space. Scattering by the object depends only on the direction of the source and receiver relative to the target in free space in the farfield.

guide boundaries is negligible, the object lies within a layer of constant sound speed, and the range from the object to source or receiver must be large enough that the scattered field can be approximated as a linear function of the object's plane wave scatter function and by modal summations.

In Eq. (1), the field radiated by the source is decomposed into modes incident on the object. Each incoming mode is composed of a pair of plane waves: one downgoing with amplitude $A_n(\mathbf{r}_0)$ and one up-going with amplitude $B_n(\mathbf{r}_0)$ with incident elevation angles, α_n and $\pi-\alpha_n$, respectively. Each scattered mode is also composed of a pair of plane waves with amplitude $B_m(\mathbf{r})$ and elevation angle α_m for the downgoing plane wave and amplitude $A_m(\mathbf{r})$ and elevation angle $\pi-\alpha_m$ for the upgoing plane wave.[15] Each of the four terms in Eq. (1) represents the *coherent* scattering of one of the two incoming plane wave components of the *n*th mode, into one of the two outgoing plane wave components of the *m*th mode. The farfield physics of the interaction is determined by the scatter function that depends on the elevation angles of both the incident and scattered plane waves. The scatter function is a coherent quantity that affects both the amplitude and phase of the scattered plane waves. The scattered fields from each incident plane wave are coherently superposed to form the total scattered field at the receiver. The scattering process couples the modes so that propagation and scattering are coherently convolved for objects submerged in a shallow water waveguide. A conceptual diagram of scattering in a waveguide is given in Fig. 1(a) for a waveguide that excites only two modes.

We would like to point out some errata in Ref. 8. In Eqs. (14)–(17) and (37) of Ref. 8, the functions B_m and $B_{m'}^*$ should be replaced by A_m and $A_{m'}^*$, respectively, in Eqs. (41) and (43)–(45), the functions B_l and B_l^* should be replaced by A_l and A_l^* respectively, and in Eqs. (14) and (15), the angles α_m and $\alpha_{m'}$ should be replaced with $\pi-\alpha_m$ and $\pi-\alpha_{m'}$ respectively. In Eqs. (42)–(45), the vertical wavenumber γ_l should be replaced with $-\gamma_l$ and the vertical wavenumber $-\gamma_m$ should be replaced by $+\gamma_m$.[15]

III. THE SONAR EQUATION

A. Sonar equation model for the scattered field from an object in free space

Here we derive the sonar equation from first principles using Green's theorem and some steps from Appendix A of Ref. 8. With the object centroid at the origin of the coordinate system, let the coordinates of a point on the surface of the object be defined by $\mathbf{r}_t=(x_t,y_t,z_t)$. Using the Helmholtz–Kirchoff integral equation,[12] the harmonic field scattered by an object can be expressed as

$$\Phi_s(\mathbf{r})=-\oiint_{A_t}\left\{[\Phi_i(\mathbf{r}_t)+\Phi_s(\mathbf{r}_t)]\,\frac{\partial G(\mathbf{r}|\mathbf{r}_t)}{\partial n_t}-G(\mathbf{r}|\mathbf{r}_t)\frac{\partial}{\partial n_t}[\Phi_i(\mathbf{r}_t)+\Phi_s(\mathbf{r}_t)]\right\}dA_t, \tag{5}$$

where $G(\mathbf{r}|\mathbf{r}_t)$ is the medium's Green's function and $\Phi_i(\mathbf{r}_t)$ is the incident field, each satisfying the Helmholtz equation driven by a source at angular frequency $\omega=2\pi f$. The area integral encloses the scatterer and the surface normal points into the enclosed volume.

In free space, for a point source at location $\mathbf{r}_0$, the field incident on the object is approximately planar for $r_0\gg r_t$ and can be expressed as

$$\Phi_i(\mathbf{r}_t)=\mathcal{A}\frac{e^{ik|\mathbf{r}_t-\mathbf{r}_0|}}{|\mathbf{r}_t-\mathbf{r}_0|}\approx\mathcal{A}\frac{e^{ikr_0}}{r_0}e^{ikr_t\eta(\alpha_i,\beta_i,\theta_t,\phi_t)}, \tag{6}$$

where (α_i,β_i) is the direction the plane wave from the source travels to, and the cosine between the directions (θ_t,ϕ_t) and (α_i,β_i) is

$$\eta(\alpha_i,\beta_i,\theta_t,\phi_t)=\cos\,\alpha_i\cos\theta_t+\sin\,\alpha_i\sin\theta_t\,\cos\,(\beta_i-\phi_t). \tag{7}$$

The Green's function in free space is

$$G(\mathbf{r}|\mathbf{r}_t)=\frac{1}{4\pi}\frac{e^{ik|\mathbf{r}-\mathbf{r}_t|}}{|\mathbf{r}-\mathbf{r}_t|}\approx\frac{1}{4\pi}\frac{e^{ikr}}{r}e^{-ikr_t\eta(\alpha,\beta,\theta_t,\phi_t)}, \quad (8)$$

where the last approximation is for a receiver so far from the object that $r\gg r_t$. By application of Green's theorem, the scattered field at this distant receiver then can be written as

$$\Phi_s(\mathbf{r})=-\mathcal{A}\frac{e^{ikr_0}}{r_0}\frac{e^{ikr}}{4\pi r}\oiint_{A_t}\left\{\left(e^{ikr_t\eta(\alpha_i,\beta_i,\theta_t,\phi_t)}+\frac{r_0}{\mathcal{A}}e^{-ikr_0}\Phi_s(\mathbf{r}_t)\right)\frac{\partial}{\partial n_t}e^{-ikr_t\eta(\alpha,\beta,\theta_t,\phi_t)}\right.$$
$$\left.-e^{-ikr_t\eta(\alpha,\beta,\theta_t,\phi_t)}\frac{\partial}{\partial n_t}\left(e^{ikr_t\eta(\alpha_i,\beta_i,\theta_t,\phi_t)}+\frac{r_0}{\mathcal{A}}e^{-ikr_0}\Phi_s(\mathbf{r}_t)\right)\right\}dA_t. \quad (9)$$

By definition of the plane wave scatter function $S(\alpha,\beta,\alpha_i,\beta_i)$, however, Eq. (9) can also be written as

$$\Phi_s(\mathbf{r})=\mathcal{A}\frac{e^{ikr_0}}{r_0}\frac{e^{ikr}}{kr}S(\alpha,\beta,\alpha_i,\beta_i), \quad (10)$$

in an object-centered coordinate system, which leads to the equality

$$S(\alpha,\beta,\alpha_i,\beta_i)=-\frac{k}{4\pi}\oiint_{A_t}\left\{\left(e^{ikr_t\eta(\alpha_i,\beta_i,\theta_t,\phi_t)}+\frac{r_0}{\mathcal{A}}e^{-ikr_0}\Phi_s(\mathbf{r}_t)\right)\frac{\partial}{\partial n_t}e^{-ikr_t\eta(\alpha,\beta,\theta_t,\phi_t)}\right.$$
$$\left.-e^{-ikr_t\eta(\alpha,\beta,\theta_t,\phi_t)}\frac{\partial}{\partial n_t}\left(e^{ikr_t\eta(\alpha_i,\beta_i,\theta_t,\phi_t)}+\frac{r_0}{\mathcal{A}}e^{-ikr_0}\Phi_s(\mathbf{r}_t)\right)\right\}dA_t, \quad (11)$$

that relates Eq. (9) directly to the Green's theorem when $r\gg r_t$.

Using the free space Green's function, Eq. (8), we can write Eq. (10) as

$$\Phi_s(\mathbf{r})=\mathcal{A}(4\pi)^2G(\mathbf{0}|\mathbf{r}_0)G(\mathbf{r}|\mathbf{0})\frac{S(\alpha,\beta,\alpha_i,\beta_i)}{k}. \quad (12)$$

From Eq. (12), we see that the Green's functions that describe the propagation of waves from source to object $G(\mathbf{0}|\mathbf{r}_0)$ and from object to receiver $G(\mathbf{r}|\mathbf{0})$ are decoupled from the scattering function $S(\alpha,\beta,\alpha_i,\beta_i)$ of the object that governs the scattering process. Only the directions of the source and the receiver relative to the object matter for farfield scattering in free space where propagation and scattering effects become factorable from each other. The approximation given in Eq. (12) is always valid in the farfield, where r, $r_0>L^2/\lambda$, and may be valid at much closer ranges for certain targets, such as spheres.[9] The incident wave effectively arrives at the target as a plane wave propagating from the direction of the source and the scattered wave at the point receiver behaves as a plane wave propagating from the target centroid. The scatter function of the target determines how a plane wave from the source is scattered in the direction of the receiver.

It is important to notice that Eq. (12) is in the frequency domain for a time harmonic source. If the source was broadband with spectrum $Q(f)$, the received field would be the inverse Fourier transform of the product of $Q(f)$ and the right-hand side of Eq. (12). For a broadband source signal, it is impossible to separate scattering from propagation even in free space. For a narrow band source, $\mathcal{A}\approx Q(f)df$.

Equation (10) can be recast as a sonar equation by taking 10 log of the squared magnitude of both sides,

$$10\log\left(\frac{|\Phi_s(\mathbf{r})|^2}{P_{\text{ref}}^2}\right)=\text{SL}-\text{TL}(\mathbf{0}|\mathbf{r}_0)+\text{TS}-\text{TL}(\mathbf{r}|\mathbf{0}), \quad (13)$$

where $P_{\text{ref}}=1$ μPa, $r_{\text{ref}}=1$ m, and

$$\text{SL}=20\log\left|\frac{\mathcal{A}}{P_{\text{ref}}r_{\text{ref}}}\right| \text{ dB } re \text{ 1 } \mu\text{Pa at 1 m}, \quad (14)$$

$$\text{TL}(\mathbf{0}|\mathbf{r}_0)=20\log\frac{r_0}{r_{\text{ref}}} \text{ dB } re \text{ 1 m}, \quad (15)$$

$$\text{TS}=20\log\left|\frac{S(\alpha,\beta,\alpha_i,\beta_i)}{kr_{\text{ref}}}\right| \text{ dB } re \text{ 1 m}, \quad (16)$$

$$\text{TL}(\mathbf{0}|\mathbf{r})=20\log\frac{r}{r_{\text{ref}}} \text{ dB } re \text{ 1 m}. \quad (17)$$

Following the sonar equation, the radiated sound has a source level of SL, which is the sound pressure level measured at 1 m from the source. This is reduced by the transmission loss $\text{TL}(\mathbf{r}_0|\mathbf{0})$, from source to target centroid. The level is augmented by the target strength TS, and further diminished by transmission loss $\text{TL}(\mathbf{0}|\mathbf{r})$ from target centroid to receiver. The level of the scattered field in decibels, Eq. (13), is a linear combination of these incoherent quantities. The incoherent target strength is obtained from the magnitude of the free space scatter function following Eq. (16). It contains only the amplitude but not the phase information of the coherent scatter function and depends only on the direc-

tion of the source and receiver, relative to the target. A conceptual diagram of free space scattering is given in Fig. 1(b).

B. Application of the sonar equation in a waveguide

It is common practice when using the sonar equation in a waveguide to replace the transmission loss in free space with that in the waveguide.[2] This can be done analytically by replacing the free-space Green's function with the waveguide Green's function in Eq. (12). Using a modal formulation, the Green's function in the waveguide between a point at the origin **0** and a field point at ***r*** can be expressed as a sum of normal modes,

$$G(\mathbf{r}|\mathbf{0}) = \frac{i}{d(0)}(8\pi)^{-1/2}e^{-i\pi/4}\sum_{m}^{M_{\max}} u_m(z)u_m(0)\frac{e^{i\xi_m\rho}}{\sqrt{\xi_m\rho}}. \tag{18}$$

Using Eqs. (2) and (4), we can express the Green's function in the waveguide, Eq. (18), as

$$G(\mathbf{r}|\mathbf{0}) = \sum_{m}^{M_{\max}} [A_m(\mathbf{r}) - B_m(\mathbf{r})]. \tag{19}$$

By reciprocity,

$$G(\mathbf{0}|\mathbf{r}_0) = G(\mathbf{r}_0|\mathbf{0}) = \sum_{n}^{M_{\max}} [A_n(\mathbf{r}_0) - B_n(\mathbf{r}_0)]. \tag{20}$$

Substituting the waveguide Green's functions, Eqs. (19) and (20) into Eq. (12), we obtain the sonar equation approximation for the scattered field from an object in a waveguide:

$$\Phi_s(\mathbf{r}|\mathbf{r}_0) = \mathcal{A}(4\pi)^2\left(\sum_{n}^{M_{\max}} [A_n(\mathbf{r}_0) - B_n(\mathbf{r}_0)]\right) \times\left(\sum_{m}^{M_{\max}} [A_m(\mathbf{r}) - B_m(\mathbf{r})]\right)\frac{S(\alpha,\beta,\alpha_i,\beta_i)}{k}. \tag{21}$$

The wave-theoretic model for object scattering in a waveguide, Eq. (1), differs significantly from the sonar equation model in Eq. (21). In the waveguide scattering model, the scattered field depends on the direction of each incoming and outgoing modal plane wave. Each incoming plane wave is *coherently* scattered to each outgoing plane wave by the object depending on the scatter function, which can vary with the azimuth and elevation angles of the incoming and outgoing plane waves. In the sonar equation model, since propagation and scattering are assumed to decouple, the scattered field depends only on the direction of the source and receiver relative to the object and not the direction of the individual modal plane waves.

The sonar equation (21) is a special case of the general coherent scattering formulation for a waveguide of Eq. (1), and so is only valid under restrictive conditions. If the scatter function remains constant over the horizontal grazing angle span of the waveguide modes $\pm\Delta\psi$ for the given measurement scenario, the scatter function factors from the modal sums of the waveguide scattering model, Eq. (1), that then reduces to the sonar equation, Eq. (21). Propagation and scattering are then separable, and the sonar equation becomes valid in a waveguide where α and α_i are approximately $\pi/2$ in Eq. (21). Target strength, along with the other incoherent terms of the sonar equation, SL and TL, then become sufficient to determine the scattered field level in decibels. We approximate the horizontal grazing angle span of the waveguide modes by

$$\pm\Delta\psi = \pm\left(\frac{\pi}{2} - \alpha_{M_{\max}}\right), \tag{22}$$

where

$$\alpha_{M_{\max}} = \tan^{-1}\frac{\xi_{M_{\max}}}{\gamma_{M_{\max}}}. \tag{23}$$

Here $M_{\max}$ and $\Delta\psi$ are range-dependent, even in realistic range-independent waveguides, and tend to decrease with range due to attenuation from absorption and scattering in the ocean, following the process known as ''mode stripping.'' This is significant because the sonar equation approximation improves as $\Delta\psi$ decreases for fixed λ/L.

IV. ILLUSTRATIVE EXAMPLES IN SHALLOW WATER

We now use examples to illustrate the fact that the sonar equation is valid when the scatter function is roughly constant over the equivalent horizontal grazing angles spanned by the dominant waveguide modes. We show that the sonar equation is generally a good approximation (1) for all objects when $2\Delta\psi < \lambda/2L$, for homogeneous convex objects when $2\Delta\psi < \lambda/L$, and (2) for spheres and certain other rounded objects in nonforward scatter azimuths, even when (1) does not hold. We proceed by analyzing active sonar examples for a variety of target types and shallow water waveguides with both the sonar equation and the waveguide scattering model.

In all the illustrative examples, a water column of 100 m depth is used to simulate a typical continental shelf environment. The sound speed in the water column is isovelocity at 1500 m/s with a constant density of 1 g/cm^3 and attenuation of 6.0×10^{-5} dB/λ. The seabed is either perfectly reflecting or comprised of sand or silt half-spaces. The density, sound speed, and attenuation are taken to be 1.9 g/cm^3, 1700 m/s, and 0.8 dB/λ for sand, 1.4 g/cm^3, 1520 m/s, and 0.3 dB/λ for silt. The receiver is either colocated with the source, in which case we calculate the backscattered field where $\beta_i = 0$, $\beta = \pi$, or in the forward azimuth of the object where $\beta_i = \beta = 0$ and we calculate the forward scattered field. The scattered fields from a disk, sphere, spheroid, and composite target are computed as a function of increasing range between source–receiver and object. In all the examples, the range increases along the *x* axis and depth along the *z* axis.

A. Effect of bottom type on the validity of the sonar equation

We use examples to illustrate how the validity of the sonar equation depends on the bottom type through the grazing angle span of the waveguide modes $\Delta\psi$. In a Pekeris

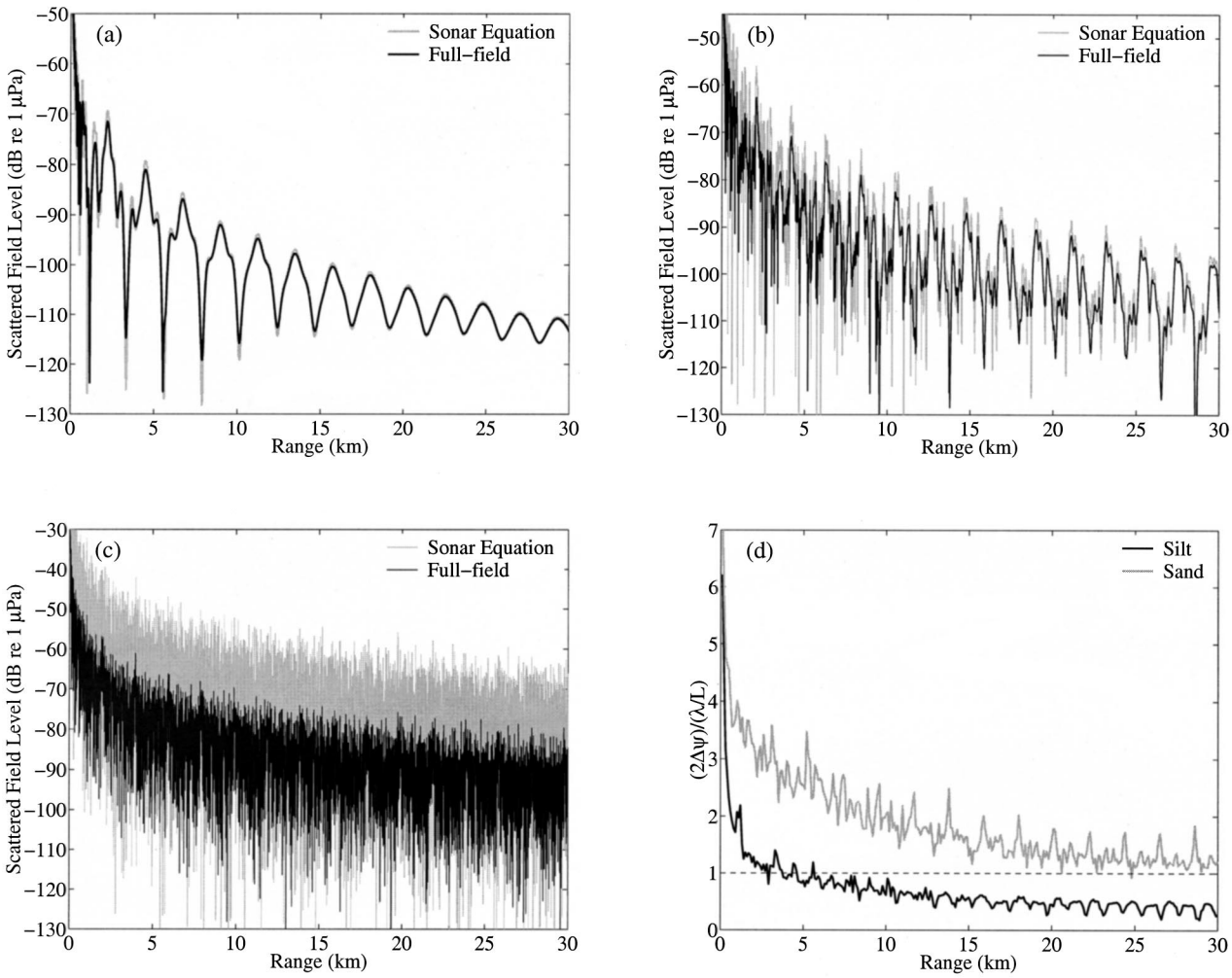


FIG. 2. The backscattered field from an upright 10 m radius rigid circular disk at 300 Hz in Pekeris waveguides with (a) silt, (b) sand, and (c) perfectly reflecting bottoms, respectively, calculated using the waveguide scattering model, Eq. (1), and compared to the sonar equation, Eq. (21). The water depth is 100 m with the source–receiver and object at 50 m depth in the middle of the water column. The range increases along the *x* axis and depth along the *z* axis. The circular disk is aligned with its plane normal to the *x* axis. The results are plotted in decibels, i.e., $20\log|\Phi_s|$, as a function of increasing range between the object and monostatic source–receiver. The source level is 0 dB *re* 1 μPa at 1 m. *ka* is 12.6 for this example. (d) The ratio $2\Delta\psi/(\lambda/L)$ for the examples given in (a) and (b). For the perfectly reflecting waveguide $2\Delta\psi/(\lambda/L)$ is 12.6. The sonar equation provides a good approximation to the scattered field in the waveguide when $2\Delta\psi/(\lambda/L)<1$.

waveguide, $\Delta\psi$ is bounded by the critical grazing angle of the bottom ψ_c beyond a few waveguide depths in range where the leaky modes no longer contribute significantly. A bottom with a large sound speed contrast relative to the water column has a correspondingly large ψ_c, $M_{\max}$ and $\Delta\psi$. So for fixed source frequency and object size, the sonar equation approximation is expected to improve as this sound speed contrast decreases. We show this by examining Pekeris waveguides with silt, sand, and perfectly reflecting bottoms that, respectively, exhibit an increase in sound speed contrast.

The backscattered field from a homogeneous convex object, an upright 10 m radius rigid circular disk at 300 Hz, is plotted in Figs. 2(a)–(c) for the three bottom types. For this example, the product $ka=\pi L/\lambda$ is 12.6, where *a* is the radius of the disk, and *ka* is the ratio of the object circumference to the wavelength. By applying Green's theorem, the scatter function for the rigid circular disk is found to be

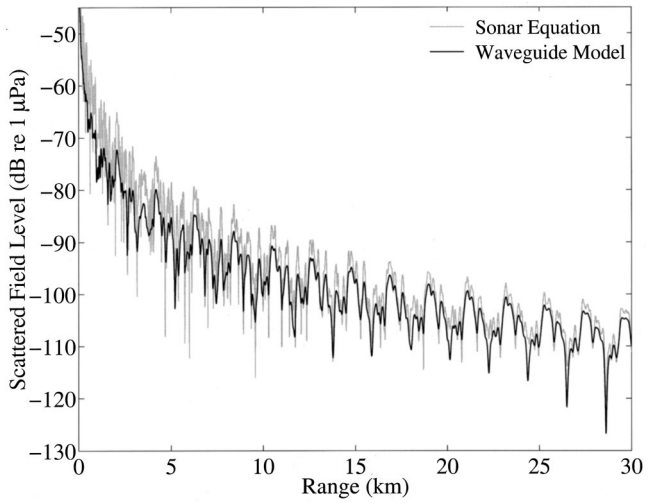


FIG. 3. Similar to Fig. 2(b), except the scattered field in the back azimuth is averaged over depth throughout the water column from 0 to 100 m.

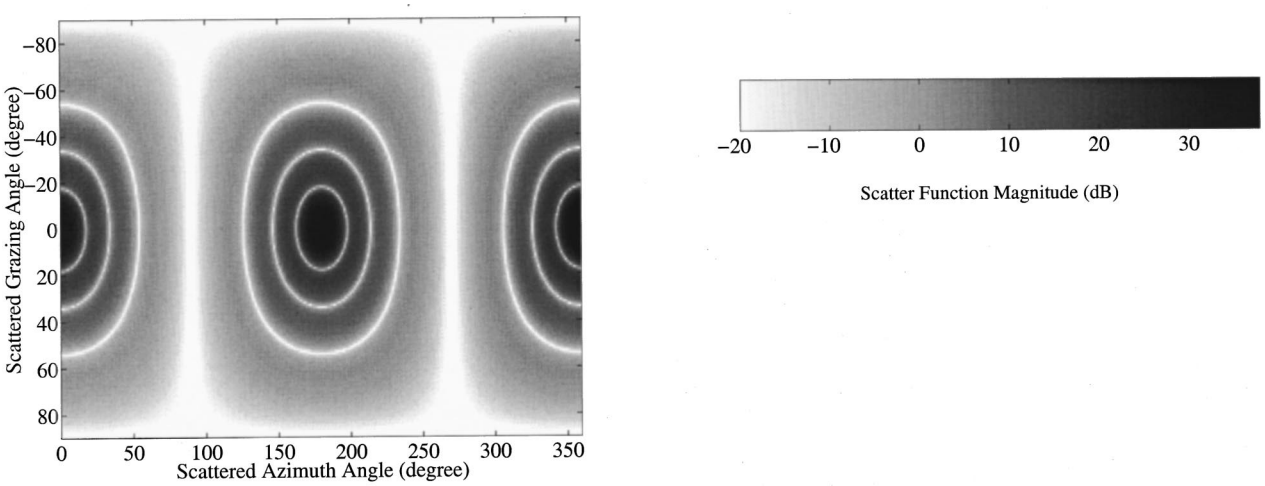


FIG. 4. The magnitude of the plane wave scatter function $20\log|S(\alpha,\beta,\alpha_i=90°,\beta_i=0°)|$ for an upright rigid circular disk of $ka=12.6$ is plotted as a function of horizontal grazing $(90°-\alpha)$ and azimuth β angles of scattered plane waves for an incident plane wave traveling in the direction $(90°-\alpha_i)=0°,\beta_i=0°$. The scatter function is antisymmetric about the plane of the disk, as can be seen from Eq. (24).

$$S(\alpha,\beta,\alpha_i,\beta_i)=-\frac{i}{2}\left(\frac{kL}{2}\right)^2 \sin\alpha\cos\beta\,\mathrm{circ}\left(\frac{kL}{2}\sqrt{(\sin\alpha_i\sin\beta_i-\sin\alpha\sin\beta)^2+(\cos\alpha_i-\cos\alpha)^2}\right), \tag{24}$$

where $L=2a$ is the diameter of the disk and $\mathrm{circ}(x)=2J_1(x)/x$. An alternative but equivalent derivation and form for the scatter function can be found in Ref. 14.

As expected, the sonar equation matches the waveguide scattering model well in the Pekeris silt waveguide of Fig. 2(a) beyond roughly 5 km range. In the Pekeris sand waveguide, Fig. 2(b), the difference between the sonar equation and the waveguide scattering model is as much as 10 dB within a few kilometers range. The match, however, improves as the range increases. In the perfectly reflecting waveguide, Fig. 2(c), the sonar equation overestimates the scattered field level by as much as 20 dB. This error is more than half the object's maximum target strength of 36 dB *re* 1 m, which occurs monostatically at broadside. The scattered field shows greater fluctuation with range for the sonar equation approximation than for the waveguide scattering model. This is true in all three waveguides.

In practice, measurements are often averaged over space, time, or frequency to reduce the fluctuations that arise from waveguide interference. Less fluctuation with range is observed in both of the depth-averaged scattered fields of Fig. 3 than in Fig. 2(b) for the sand bottom. The discrepancy between the sonar equation and waveguide scattering model remains, however, with differences as large as 10 dB still occurring within a few kilometers range.

The reason that the sonar equation is found to be a good approximation for the silt beyond roughly 5 km but not the sand or perfectly reflecting waveguides is that the condition $2\Delta\psi<\lambda/L$ only holds for the silt waveguide. This can be seen by examining Fig. 2(d), where the ratio $2\Delta\psi/(\lambda/L)$ is plotted. Here $\Delta\psi$ is computed from Eqs. (22) and (23) at a given receiver depth by determining the minimum value for $M_{\max}$ at which the modal sum of Eq. (1) differs less than 1 dB from that of an infinite sum. The $\Delta\psi$ used in the figures is the average value over the receiver depth throughout the water column. For ranges beyond roughly 5 km in the silt waveguide, the ratio is less than unity and the condition for the sonar equation to be valid holds. Figure 2(d) also shows that the condition is generally not satisfied in the sand waveguide for the ranges shown. The condition is never satisfied in the

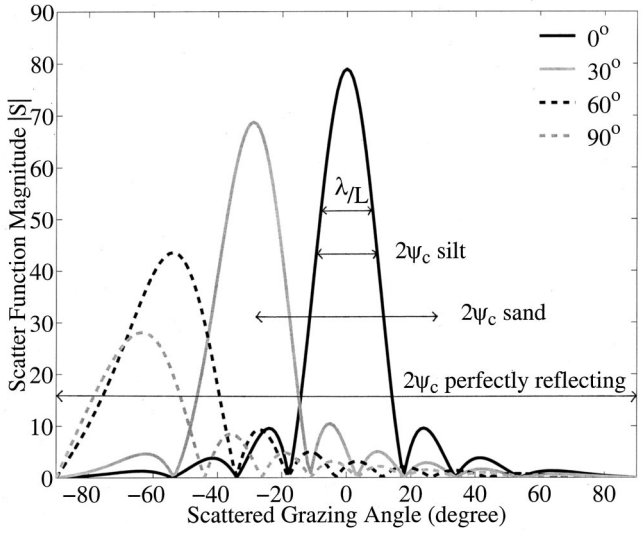


FIG. 5. The magnitude of the plane wave scatter function $|S(\alpha,\beta,\alpha_i,\beta_i=0°)|$ for an upright rigid circular disk of $ka=12.6$ is plotted as a function of the horizontal grazing angle $(90°-\alpha)$ of scattered plane waves in the backscatter azimuth $\beta=180°$ for several incident plane waves with horizontal grazing angles $(90°-\alpha_i)=0°$, 30°, 60°, and 90°. The solid curve in this figure for broadside incidence $[(90°-\alpha_i)=0°]$ is a slice through Fig. 4 at the backscatter azimuth $\beta=180°$ of the scattered plane waves. The width λ/L of the scatter function main lobe for broadside incidence is 14.3° or 0.25 rad. Also shown is the critical grazing angle ψ_c of the seabed in the Pekeris slit waveguide of 9.3°, Pekeris sand waveguide of 28°, and the perfectly reflecting waveguide of 90°.

perfectly reflecting waveguide because the ratio $2\Delta\psi/(\lambda/L)$ is always 12.6. As expected, the performance of the sonar equation improves with range in realistic ocean waveguides because $M_{\max}$ and $\Delta\psi$ decrease due to modal stripping, as is evident in Figs. 2(a), (b), and (d). Note that $\Delta\psi$ can be determined from Fig. 2(d) since $\lambda/L=0.25$ rad or 14.3°, as can be seen from Figs. 4 and 5. Here the scatter function magnitude is plotted for the upright disk, where the main lobe has a minimum width of λ/L for a plane wave incident at broadside, where $(\pi/2-\alpha_i)=0$.

To visualize why the sonar equation is not valid when the condition $2\Delta\psi<(\lambda/L)$ does not hold, it is instructive to plot the width of the bottom critical angle for each bottom type across the scatter function main lobe, as is done in Fig. 5. When the condition does not hold, the object scatters the dominant incident modes with widely varying amplitudes and the scatter function cannot be approximated as a constant over $\pm\Delta\psi$. In this case, both the magnitude and phase variations of the scatter function are important in describing the scattering process. The sonar equation *overestimates* the level of the scattered field because it depends only on two directions: those of the source and receiver relative to the object. The two relevant directions for the upright disk are $(\pi/2-\alpha_i)=0$ and $\beta_i=0$ for the incident and $(\pi/2-\alpha)=0$ and $\beta=\pi$ for the scattered field. These correspond to global maxima in both the scatter function and target strength, as can be seen in Fig. 5, which is inappropriately assigned to all incoming and outgoing directions by the sonar equation.

B. Effect of object size and frequency on the validity of the sonar equation

In this section, we investigate sonar equation performance as a function of object size and frequency for a rigid circular disk in various waveguides. At the high *ka* of 62.8 shown in Figs. 6(a), (c), and (e), the object is large compared to the wavelength and the sonar equation significantly overestimates the scattered field level. This is to be expected from Fig. 6(g), where the condition $2\Delta\psi<\lambda/L$ is not satisfied in any of the waveguides. At the lower *ka* of 1.3, Figs. 6(b), (d), and (f), the sonar equation provides a good approximation in all except the perfectly reflecting waveguide,

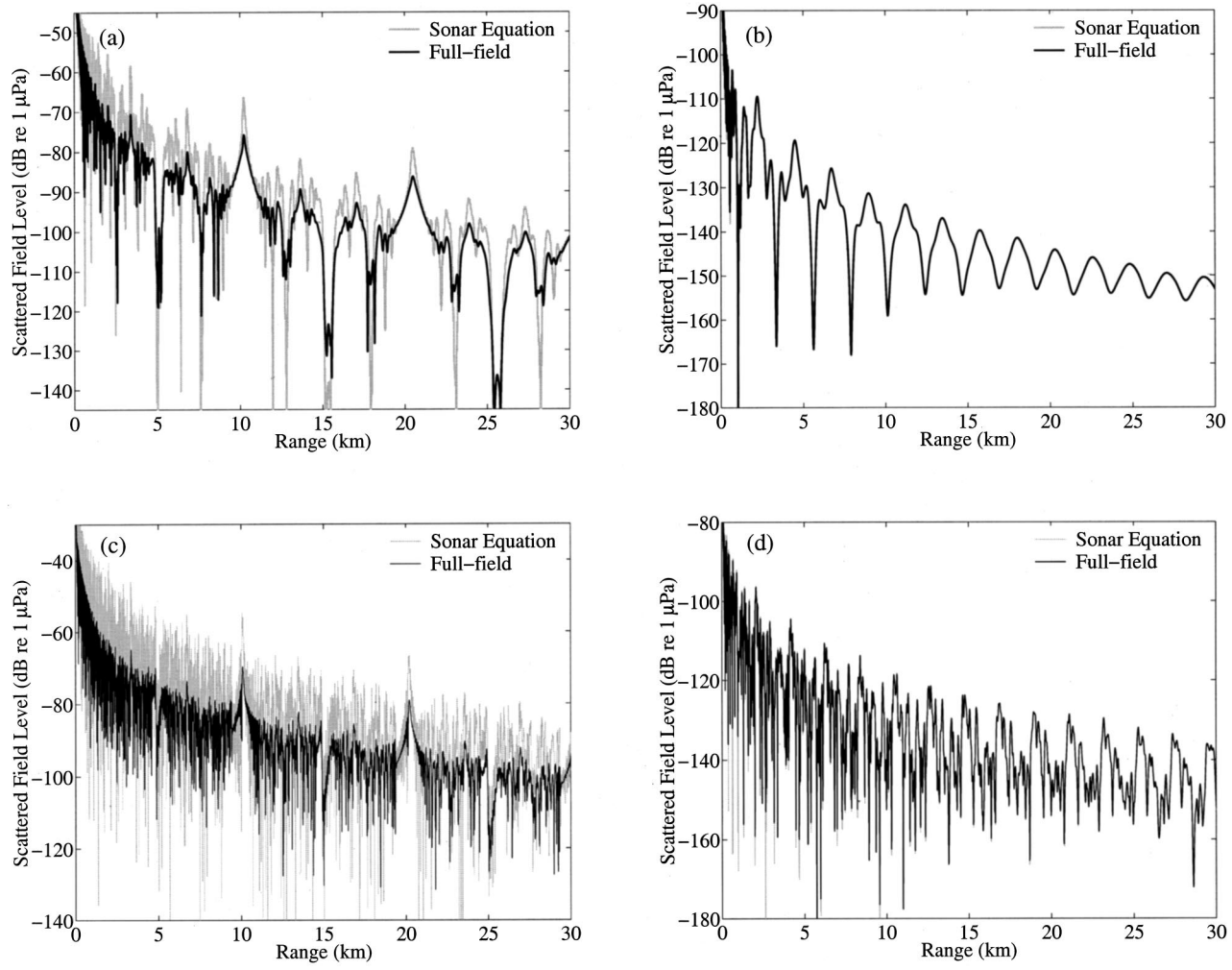


FIG. 6. Similar to Figs. 2(a)–(c), but for an upright rigid circular disk of (a) high and (b) low *ka* in the Pekeris silt waveguide, (c) high and (d) low *ka* in the Pekeris sand waveguide, and (e) high and (f) low *ka* in the perfectly reflecting waveguide. The high *ka* case corresponds to a disk of 10 m radius at 1500 Hz with $ka=62.8$ while the low *ka* case corresponds to a disk of 1 m radius at 300 Hz with $ka=1.3$. (g) and (h) are similar to Fig. 2(d) but for *ka* of 62.8 and 1.3, respectively for the cases shown in (a)–(d) in the Pekeris silt and sand waveguides. The ratio $2\Delta\psi/(\lambda/L)=62.8$ for (e) with *ka* of 62.8 and $2\Delta\psi/(\lambda/L)=1.3$ for (f) with *ka* of 1.3 in the perfectly reflecting waveguide.

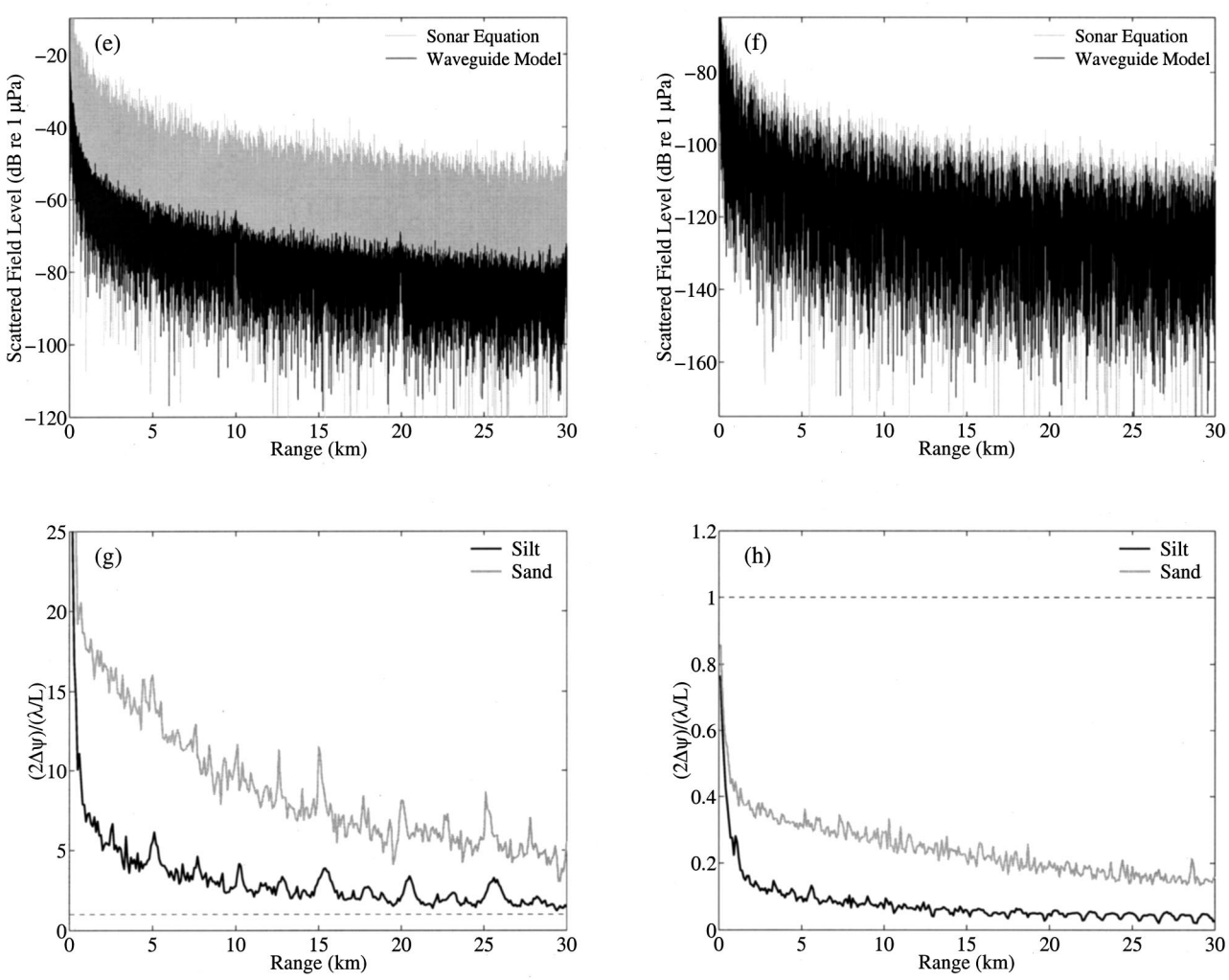


FIG. 6. (*Continued.*)

which is consistent with the results of Fig. 6(h), where the condition is satisfied for the sand and silt waveguides for all ranges shown. This shows that the sonar equation can be made valid for a given object and measurement geometry by lowering the frequency of operation. The scatter functions for the high and low *ka* cases are plotted in Figs. 7(a) and (b), respectively. For the low *ka* case in the perfectly reflecting waveguide, the width of the scatter function lobe is only

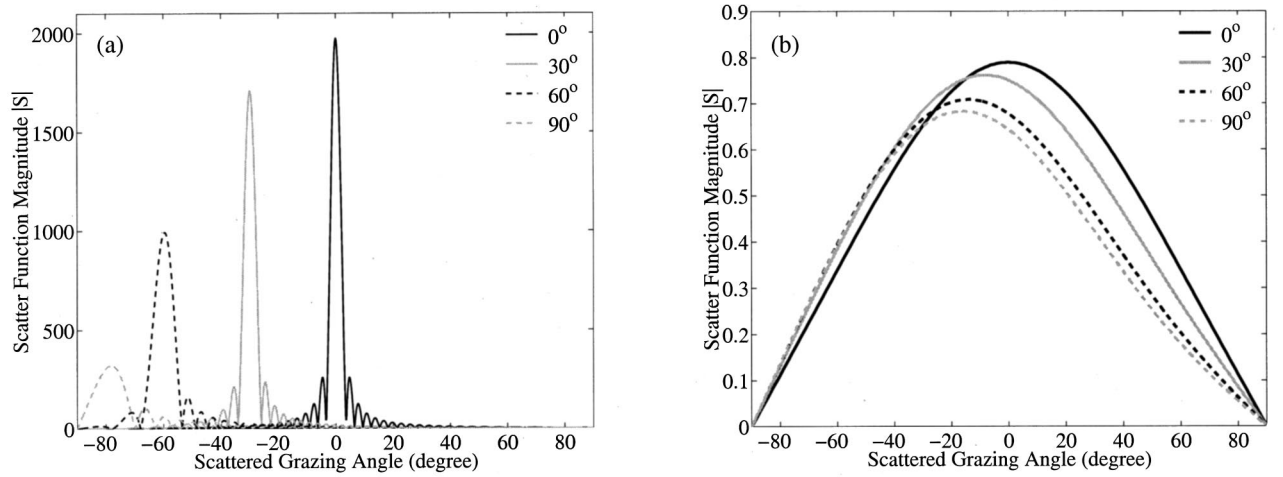


FIG. 7. Similar to Fig. 5, but for an upright rigid circular disk of (a) $ka=62.8$ and (b) $ka=1.3$. The width λ/L of the scatter function main lobe for broadside incidence $[(90°-\alpha_i)=0°]$ is 2.9° or 0.05 rad for $ka=62.8$ and 143° or 2.5 rad for $ka=1.3$.

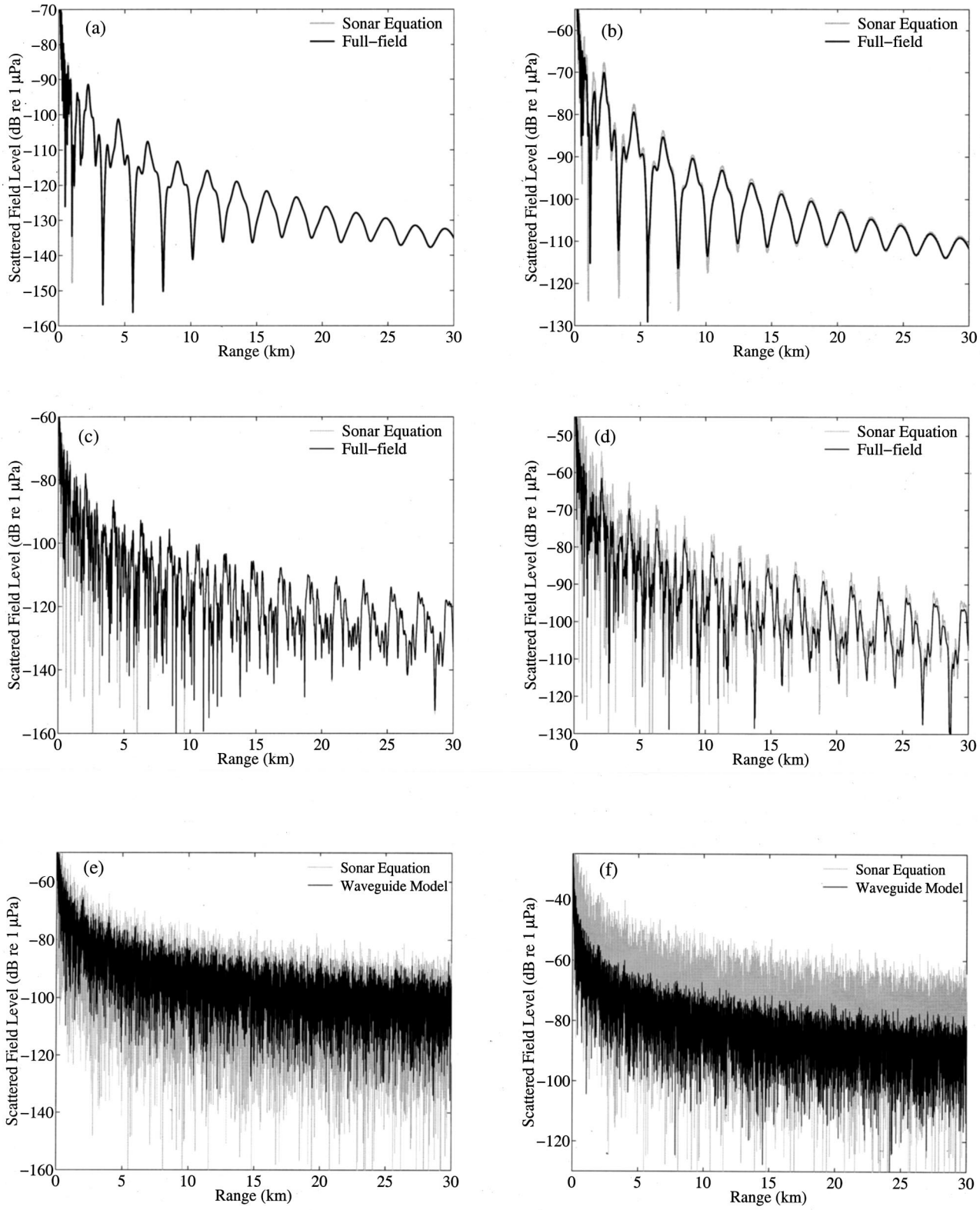


FIG. 8. Similar to Figs. 2(a)–(c) but for a 10 m radius pressure-release sphere at 300 Hz in the (a) back and (b) forward azimuths in Pekeris silt waveguide, the (c) back and (d) forward azimuths in Pekeris sand waveguide, and the (e) back and (f) forward azimuths in the perfectly reflecting waveguide. *ka* is 12.6 for this example.

slightly smaller than the full grazing angle span of the waveguide modes so the sonar equation provides only an order-of-magnitude approximation in backscatter.

In general, for pressure-release objects that are compact, small compared to the wavelength ($ka \ll 1$), the sonar equation approximation is always valid since the scattered field is effectively omnidirectional. For rigid objects that are compact, the sonar equation is still a good approximation, even though their scattered field always maintains some directionality as *ka* decreases.

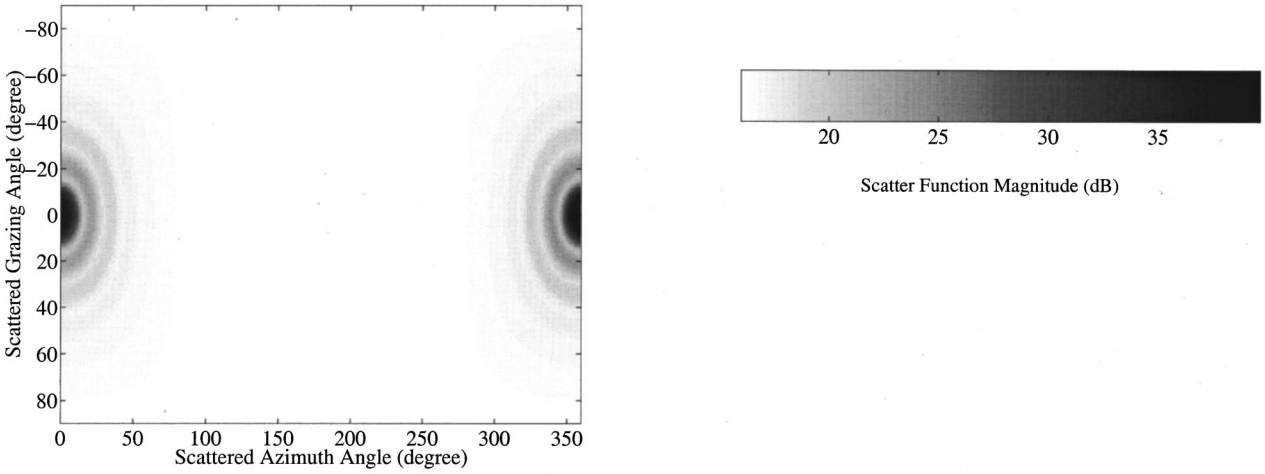


FIG. 9. Similar to Fig. 4, but for a pressure-release sphere of $ka=12.6$. The scatter function for the sphere is given by Eqs. (8) and (9) of Ref. 9 with $f(n)$ replaced by $(-1)^n f(n)$ to convert from Ingenito's definition to the standard one described in Ref. 8. It can also be obtained from Ref. 14.

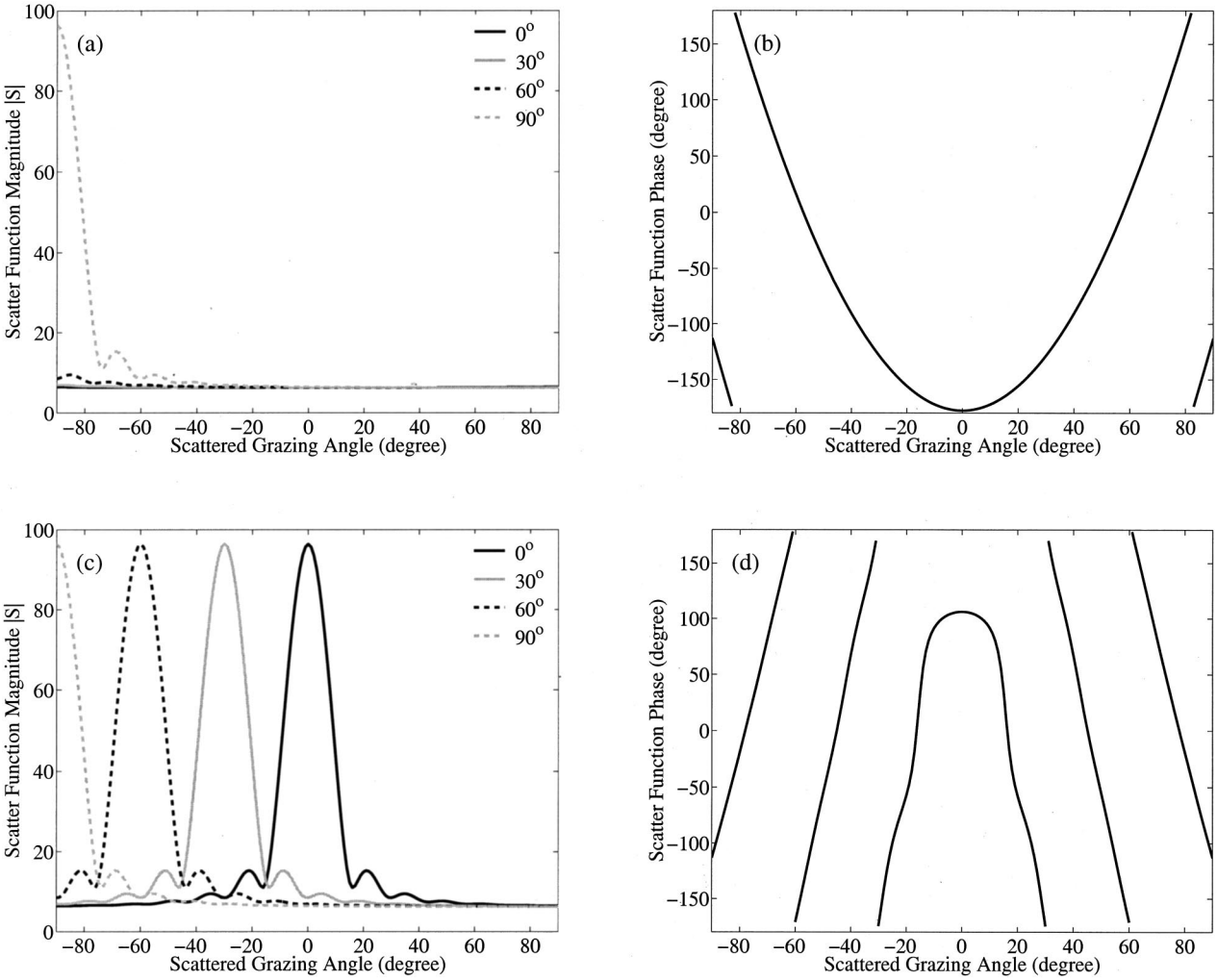


FIG. 10. (a) and (c) are similar to Fig. 5, but for a pressure-release sphere of $ka=12.6$ plotted in the back ($\beta=180°$) and forward ($\beta=0°$) azimuths, respectively, of the scattered plane waves. The solid curves in (a) and (c) for $(90°-\alpha_i)=0°$ are slices through Fig. 9 in the back and forward scatter azimuths, respectively. The width λ/L of the scatter function main lobe is 14.3° or 0.25 rad in (c). The phase of the scatter function for the solid curve [horizontal incidence, $(\alpha_i-90°)=0°$] in (a) and (c) is shown in (b) and (d), respectively, which are in the back and forward scatter azimuths.

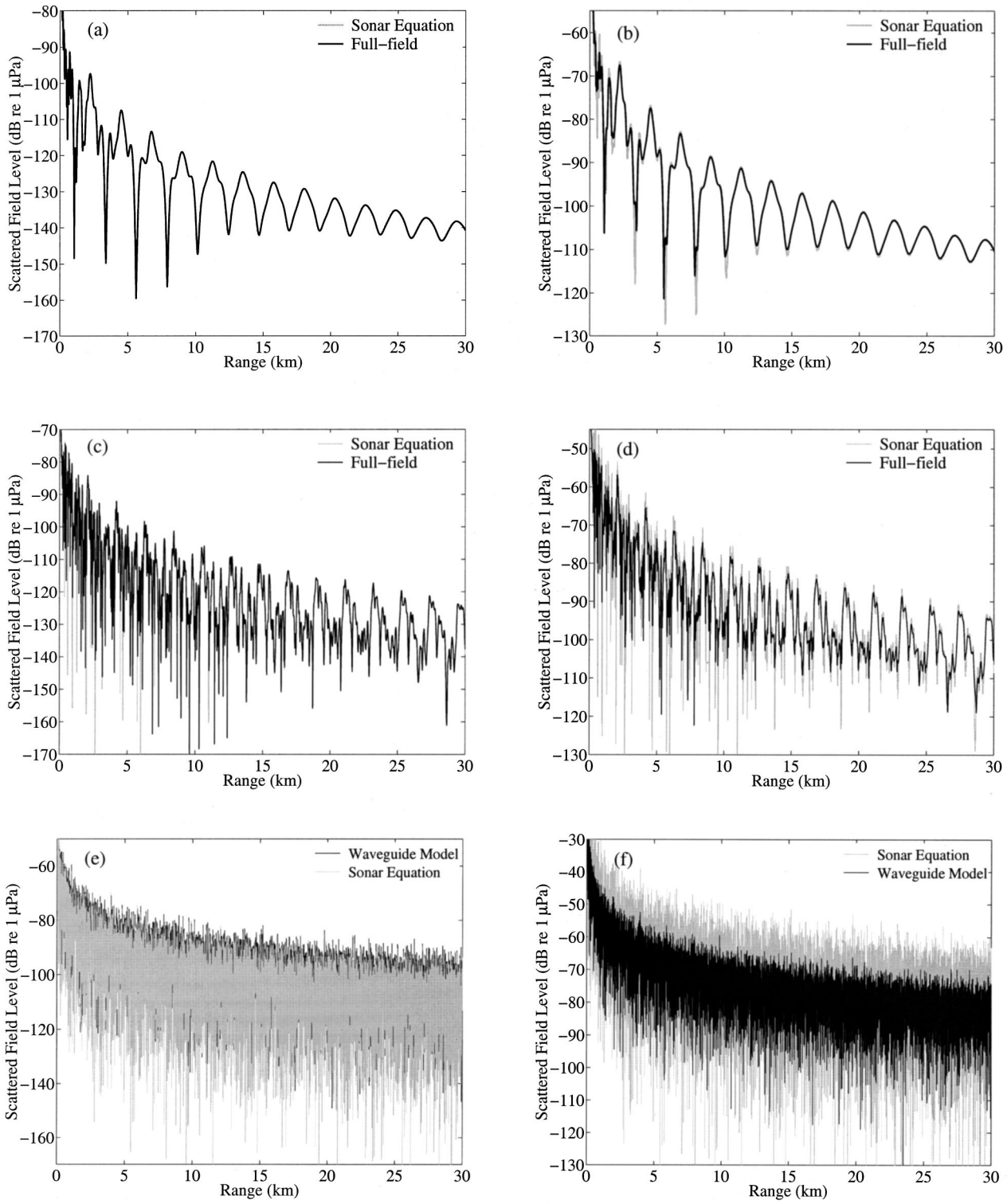


FIG. 11. Similar to Fig. 8, but for a pressure-release prolate spheroid of aspect ratio 2 with a major axis of 40 m and a minor axis of 20 m at 300 Hz. The spheroid is aligned such that the incoming plane waves are incident at the bow aspect where $ka = 12.6$.

C. Effect of orientation and shape of homogeneous convex targets on the validity of the sonar equation

In this section, we investigate the effect of orientation and shape on the validity of the sonar equation for homogeneous convex objects. The effect of orientation is investigated by rotating the upright disk ($ka = 12.56$) of Sec. IV A clockwise by 18° and 34°, respectively, about the y axis. This places the zero of the scatter function at the horizontal, as can be seen by the inspection of Fig. 5, so that the scatter function experiences a phase change about the horizontal. The difference between the scattered field and sonar equation for the two rotated cases is almost identical to the unrotated case shown in Figs. 2(a)–(c). The variation across the cases

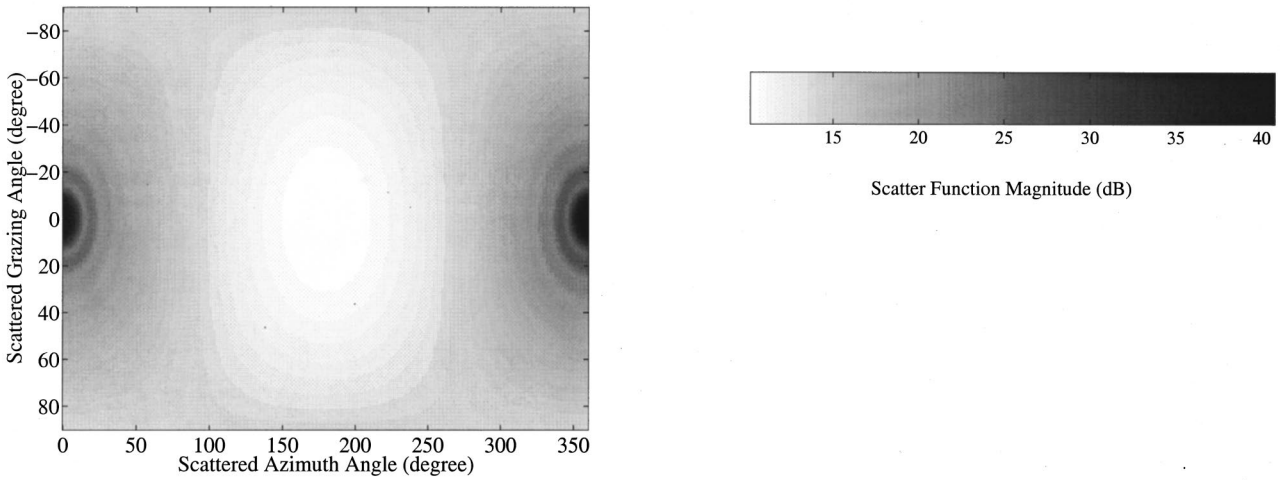


FIG. 12. Similar to Fig. 4, but for a pressure-release prolate spheroid of aspect ratio 2 with a major axis of 40 m and a minor axis of 20 m at 300 Hz. The forward scatter direction is along the major axis. At the bow aspect of the spheroid, $ka=12.6$. The scatter function for the spheroid is given by Eq. (11.38) of Ref. 14.

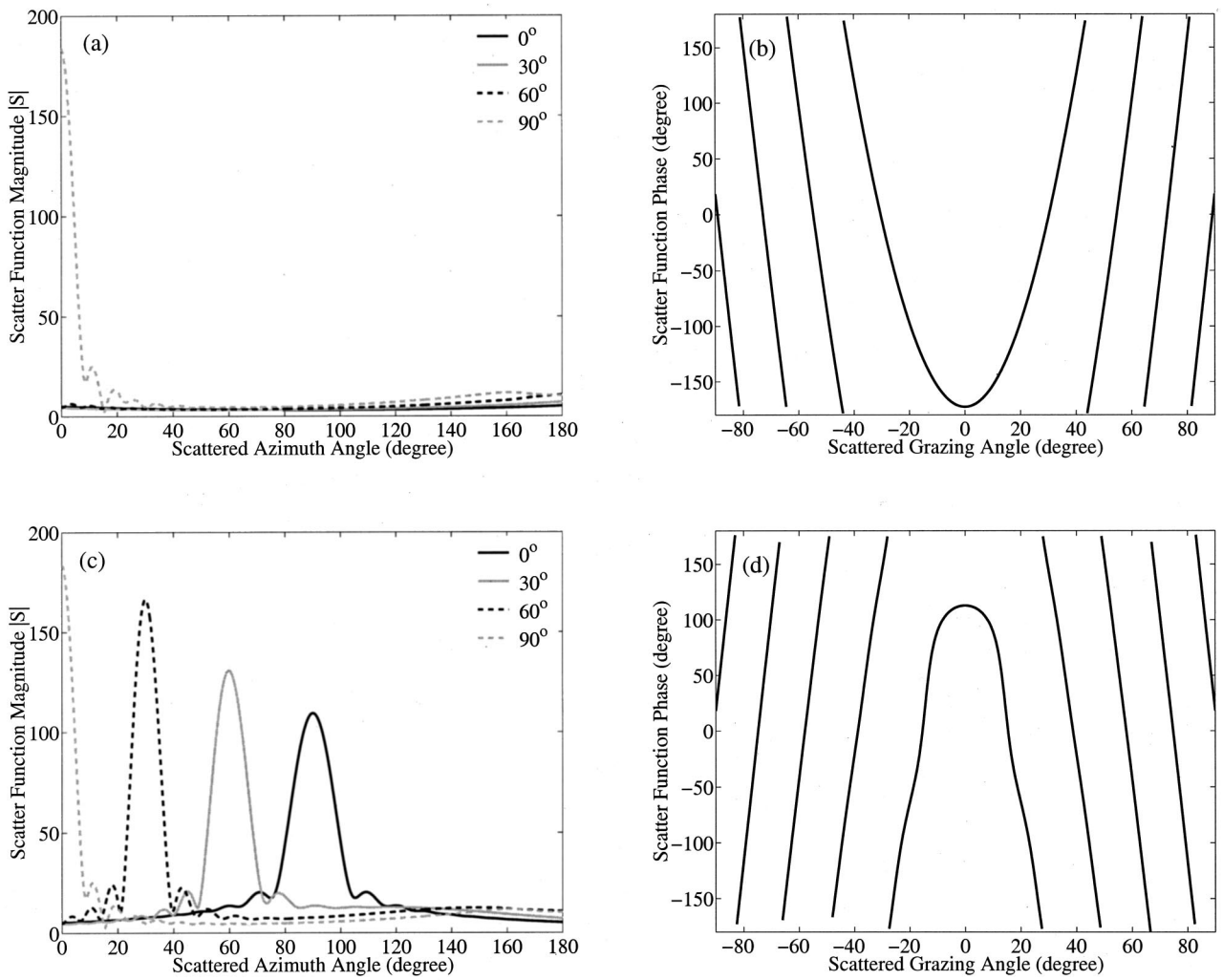


FIG. 13. Similar to Figs. 10(a)–(d) but for a pressure-release prolate spheroid of aspect ratio 2 with a major axis of 40 m and a minor axis of 20 m at 300 Hz. The forward scatter direction is along the major axis. At the bow aspect of the spheroid, $ka=12.6$. The solid curves in (a) and (c) for $(\alpha_i-90°)=0°$ are slices through Fig. 12 at the back $\beta=180°$, and forward $\beta=0°$ scatter azimuths, respectively, of the scattered plane waves. The width λ/L of the scatter function main lobe at horizontal grazing incidence $[(\alpha_i-90°)=0°]$ is 14.3° or 0.25 rad in (c).

is less than roughly 1 dB for ranges beyond 1 km, showing that the condition $2\Delta\psi<\lambda/L$, as shown in Fig. 2(d), holds, regardless of object orientation. This makes sense because the condition states that the sonar equation will be a good approximation if the scatter function undergoes no more than one oscillation within $\pm\Delta\psi$. This minimizes the destructive interference possible in multimodal propagation and scattering.

There are two limiting object shapes for homogeneous convex objects, flat and rounded ones. Flat objects, such as the disk examined in the previous section, are highly directional scatterers when *ka* is large. In free space, they scatter the strongest in specular and forward scatter directions. The sonar equation is only valid for flat objects when the condition $2\Delta\psi<\lambda/L$ holds, as demonstrated in the previous section for an upright disk and in this section for the rotated disk. We note that the scattered field level in the forward azimuth is identical to that given in the back azimuth in Figs. 2, 3, and 6 because the scatter function for the upright disk is antisymmetric about the plane of the disk, as can be seen in Eq. (24) and Fig. 4. The phase of the scatter function is constant over the main lobe, as can be seen in Eq. (24). This constancy over the main lobe is characteristic of flat objects, where the narrowest lobe is always the main lobe for broadside incidence. Rectangular plate examples are investigated in Ref. 15.

When the condition $2\Delta\psi<\lambda/L$ does not hold, the sonar equation can still be extremely accurate in nonforward scatter azimuths for spheres and certain other rounded or smoothly varying convex objects. For these objects, the scatter function is approximately uniform in nonforward directions but has a main lobe in the forward direction of width λ/L. This uniformity makes it possible to approximate the scatter function in nonforward azimuths as a factorable constant over $\pm\Delta\psi$, making the sonar equation valid. The sonar equation, however, will still only be valid in the forward scatter azimuth for spheres and rounded objects when $2\Delta\psi<\lambda/L$ holds because they behave like flat objects in forward scatter by Babinet's principle. These issues will be illustrated for two smoothly varying convex objects, the sphere and the prolate spheroid, in the same three waveguides examined in previous sections. Babinet's principle will be discussed further in Sec. V.

As expected, the sonar equation predicts the backscattered field much more accurately than the forward scattered field for both the sphere and the prolate spheroid, as shown in Figs. 8 and 11, respectively, for the various waveguides. This can be explained by examining the scatter function of a sphere, given in Ref. 14 and plotted in Figs. 9 and 10. Figure 10(a) shows that the magnitude of the scatter function in the backscatter azimuth for the sphere is approximately constant over $\pm\Delta\psi$ in the Pekeris silt and sand waveguides, where $\Delta\psi$ can be determined from Fig. 2(d). Figure 10(b) shows that the phase of the scatter function varies by less than $\pi/4$ over $\pm\Delta\psi$ and so can be considered effectively constant. This means that the scatter function can be factored from the modal sum in Eq. (1) so that propagation is decoupled from scattering. The sonar equation then becomes valid in the backscatter direction for the sphere in the Pekeris sand and silt waveguides.

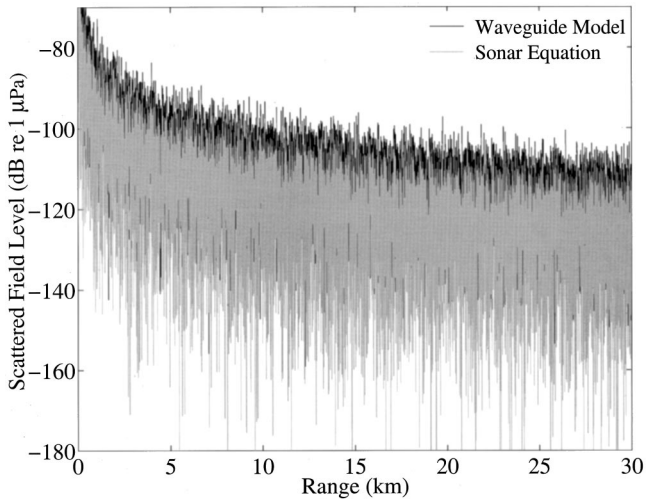


FIG. 14. Similar to Fig. 11(e), but for a pressure-release prolate spheroid of aspect ratio 5 with a major axis of 20 m and a minor axis of 4 m at 300 Hz. The spheroid is aligned such that the incoming plane waves are incident at the bow aspect where $ka=2.5$.

In the perfectly reflecting waveguide, $\Delta\psi$ is $\pi/2$ rad or 90°. In this case, some higher-order modes at very steep incident grazing angles near 90°, for example, scatter much stronger in the backscatter azimuth than the other modes of lower order that scatter more uniformly, as shown in Fig. 10(a). The phase of the scatter function also varies by more than 180° over $\Delta\psi$. As a result, scattering is not completely decoupled from propagation and so the sonar equation provides only a crude approximation in the backscatter azimuth, as shown in Fig. 8(e).

In the forward azimuth, the sphere scatters approximately as a flat object with the same projected area, by Babinet's principle, and so behaves like the disk of Fig. 2. Similar conclusions can be drawn for the prolate spheroid from Figs.

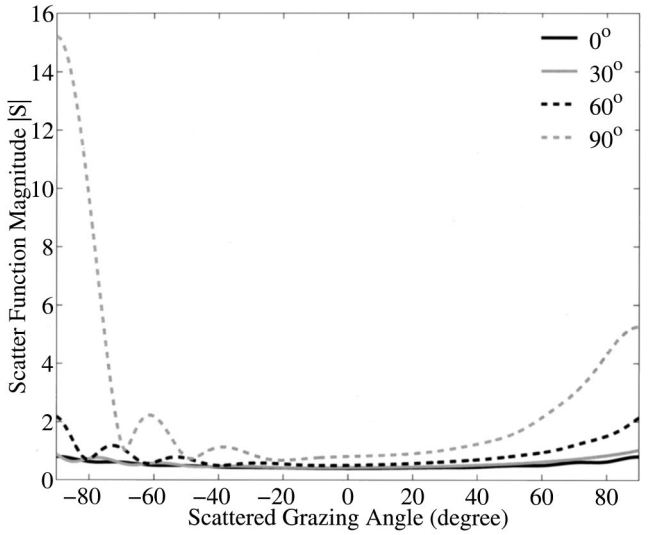


FIG. 15. Similar to Fig. 5 but for a pressure-release prolate spheroid of aspect ratio 5 with a major axis of 20 m and a minor axis of 4 m at 300 Hz in the backscatter azimuth $\beta=180°$. The forward scatter direction is along the major axis. At the bow aspect of the spheroid, $ka=2.5$.

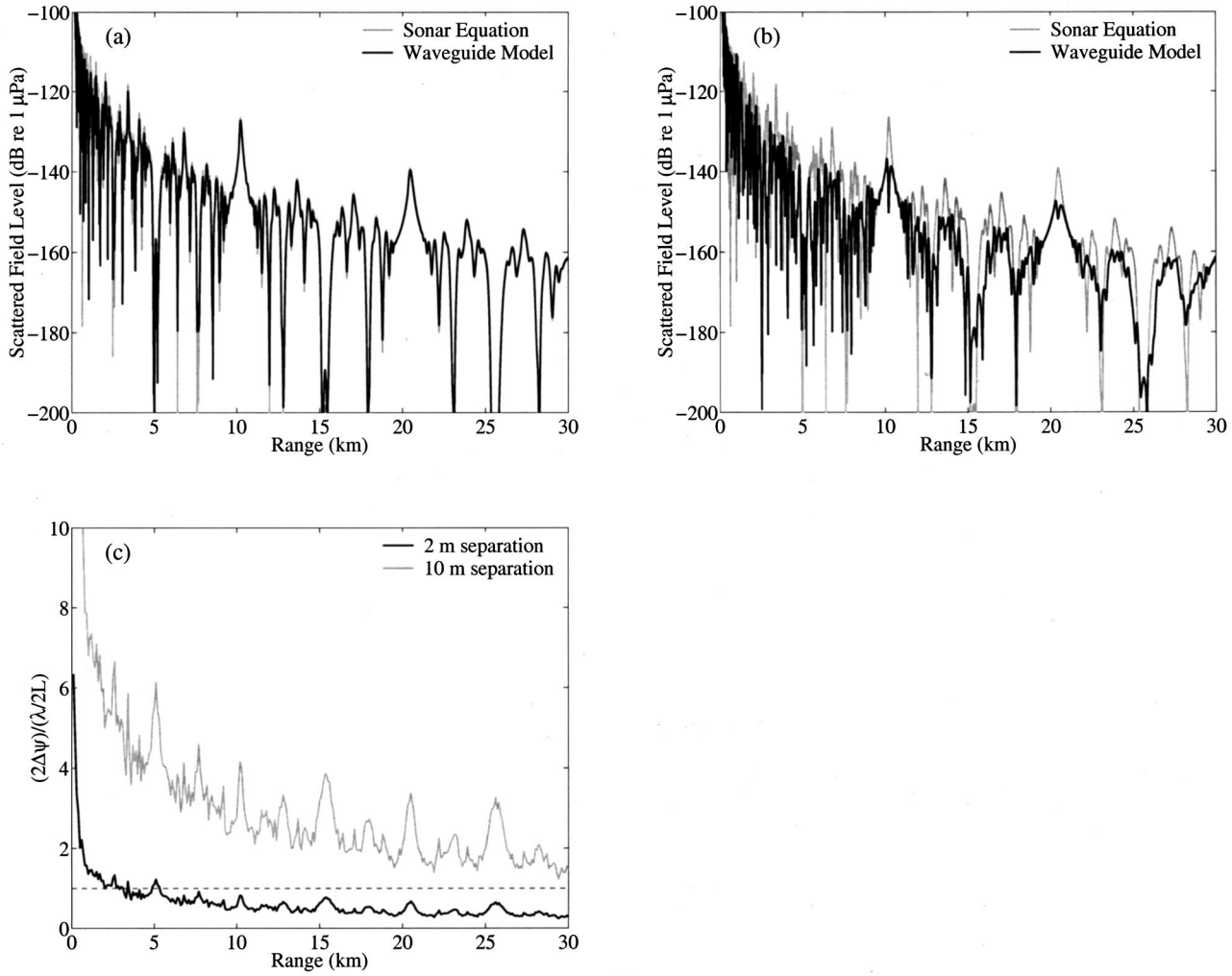


FIG. 16. Similar to Fig. 2(a) but for a ''single object'' comprised of two point scatterers with a separation of (a) $L=2$ m and (b) $L=10$ m, respectively, at 1500 Hz. (c) The ratio $2\Delta\psi/(\lambda/2L)$ for the examples given in (a) and (b). The sonar equation provides a good approximation to the scattered field in the waveguide when $2\Delta\psi/(\lambda/2L)<1$.

11–13. Its scatter function is given in Ref. 14.

The sonar equation can sometimes underestimate the scattered field depending on the object type, size, waveguide, and frequency. In Fig. 14, the sonar equation underestimates the backscattered field from a spheroid in a perfectly reflecting waveguide by a small amount, roughly 3 dB. The sonar equation assumes that the scatter function is constant and equal to that at the incident $(\pi/2-\alpha_i)=0$ and scattered $(\pi/2-\alpha)=0$ grazing angles in Fig. 15. This, however, corresponds to the backscatter minimum in the scatter function plot and is not representative of all values within $\pm\Delta\psi$.

D. Nonconvex objects and fluctuating objects

The scatter function lobes for homogeneous convex objects have minimum widths limited to λ/L by diffraction. For more general, potentially inhomogeneous, and nonconvex objects, the scatter function lobes are limited by both diffraction and interference from different parts of the object. For example, a situation may arise where it is convenient to consider an aggregate of disjoint scatterers as a ''single object'' with a single scatter function. In the limiting case of two point scatterers separated by length L normal to the incident wave, the main lobe of this ''single object's'' scatter function has the narrowest width possible $\lambda/(2L)$. The condition for the sonar equation to become valid in a waveguide then is the most stringent $2\Delta\psi<\lambda/2L$.

The scatter function for a ''single object'' comprised of two point scatterers separated in depth by length L is $2\cos(\pi L[\cos\alpha_i-\cos\alpha]/\lambda)$ if the scatter function of each point is unity. The field scattered from this ''single object'' is given in Figs. 16(a) and (b) for point separations of $L=2$ m and $L=10$ m, respectively, at 1500 Hz using both the sonar equation and waveguide scattering model. For $L=2$ m the sonar equation is valid and matches the waveguide model because the condition $2\Delta\psi<\lambda/2L$ is satisfied, as can be seen in Fig. 16(c). This condition is not satisfied for the $L=10$ m case, where the sonar equation is not valid, as can be seen in Figs. 16(b) and (c).

For an object whose orientation is unknown or constantly varying, like a fluctuating target, a conservative cri-

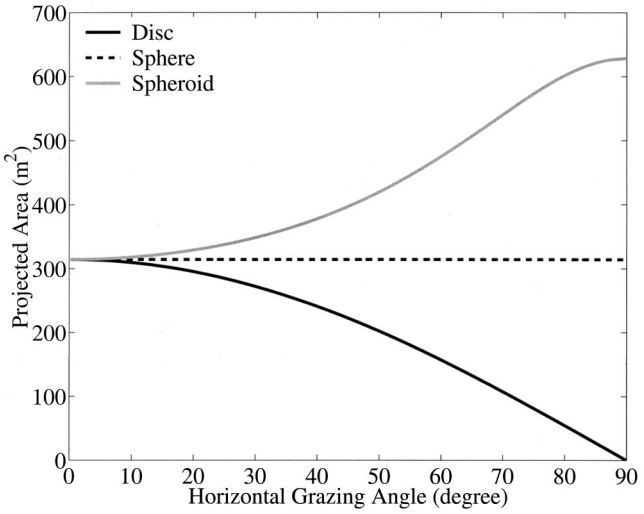


FIG. 17. Projected areas of a 10 m radius rigid circular disk, a 10 m radius pressure-release sphere and a pressure-release prolate spheroid of aspect ratio 2 with a major axis of 40 m and a minor axis of 20 m are plotted as a function of horizontal grazing angle $|\alpha_i - 90°|$ of the incident plane waves. The orientation of the disk and spheroid are similar to Figs. 2 and 11.

terion for the sonar equation's validity in a waveguide is $2\,\Delta\psi < \lambda/2L_{\max}$ where $L_{\max}$ is the largest spatial extent of the object, which may be composed of an aggregate of disjoint scatterers.

E. Forward scatter function and projected area for homogeneous convex objects

According to the extinction or forward scatter theorem in free space,[10,13,16] the scatter function of an object in the forward direction at a given frequency is proportional to the object's projected area normal to the direction of propagation of the incident wave for high *ka*. The projected area of the 10 m radius rigid circular disk is largest for plane waves incident on the disk at horizontal grazing $[(\pi/2 - \alpha_i) = 0]$, as shown in Figs. 5 and 7(a). At other incident grazing angles, the projected area of the disk is smaller and the peak value of the scatter function corresponding to forward scatter decreases. In addition, the width of the forward scatter lobe broadens since it is inversely related to the projected area. For a compact pressure-release object ($ka \ll 1$) the scatter function becomes omnidirectional.

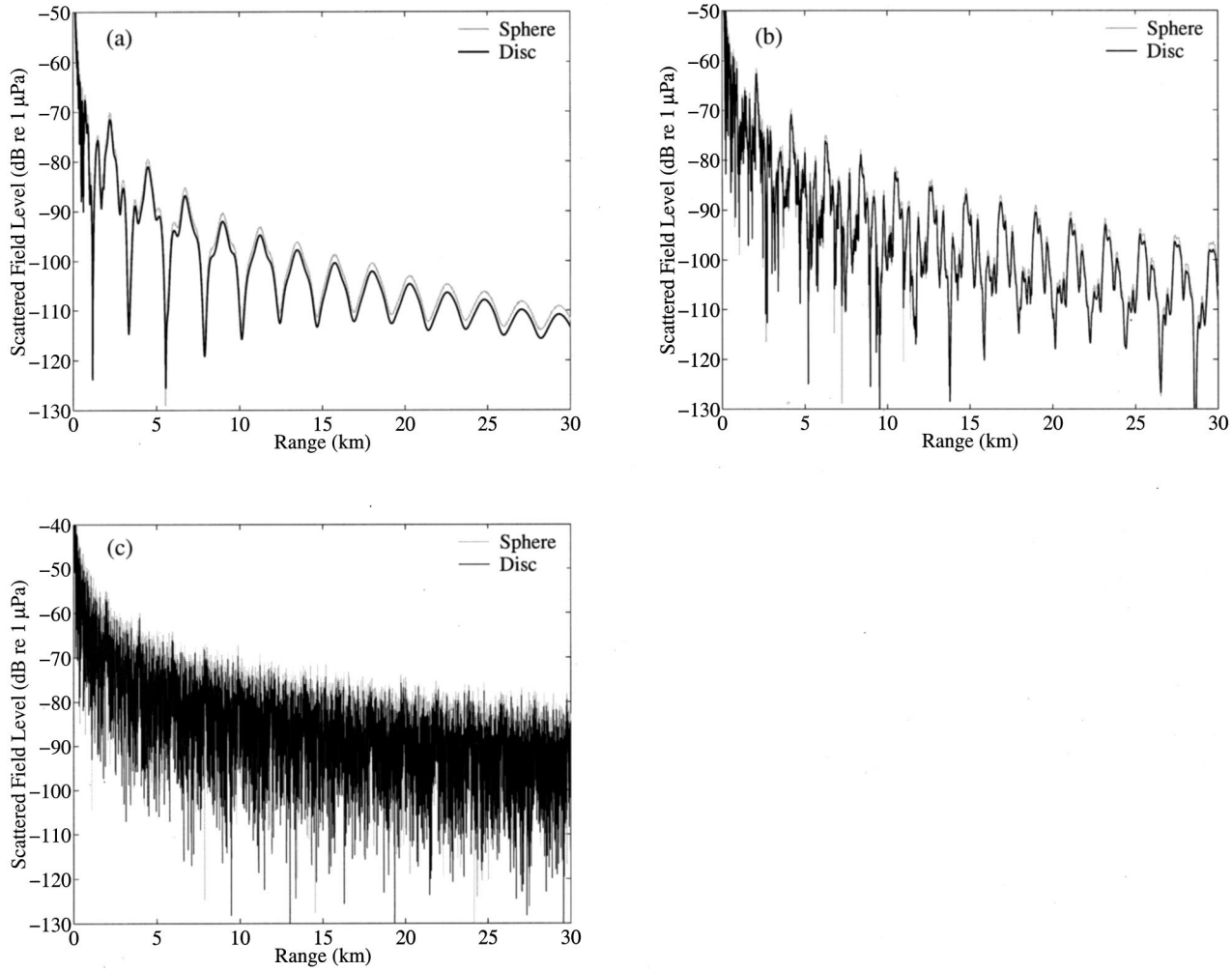


FIG. 18. The forward scattered field at 300 Hz calculated using the waveguide scattering model, Eq. (1), from a 10 m radius pressure release sphere is compared to that from a rigid circular disk of radius 10 m in (a) Pekeris silt, (b) Pekeris sand, and (c) perfectly reflecting waveguides. *ka* is 12.6 for these examples. The geometry of the setup is similar to Figs. 2 and 8, except that the receiver is in the forward azimuth.

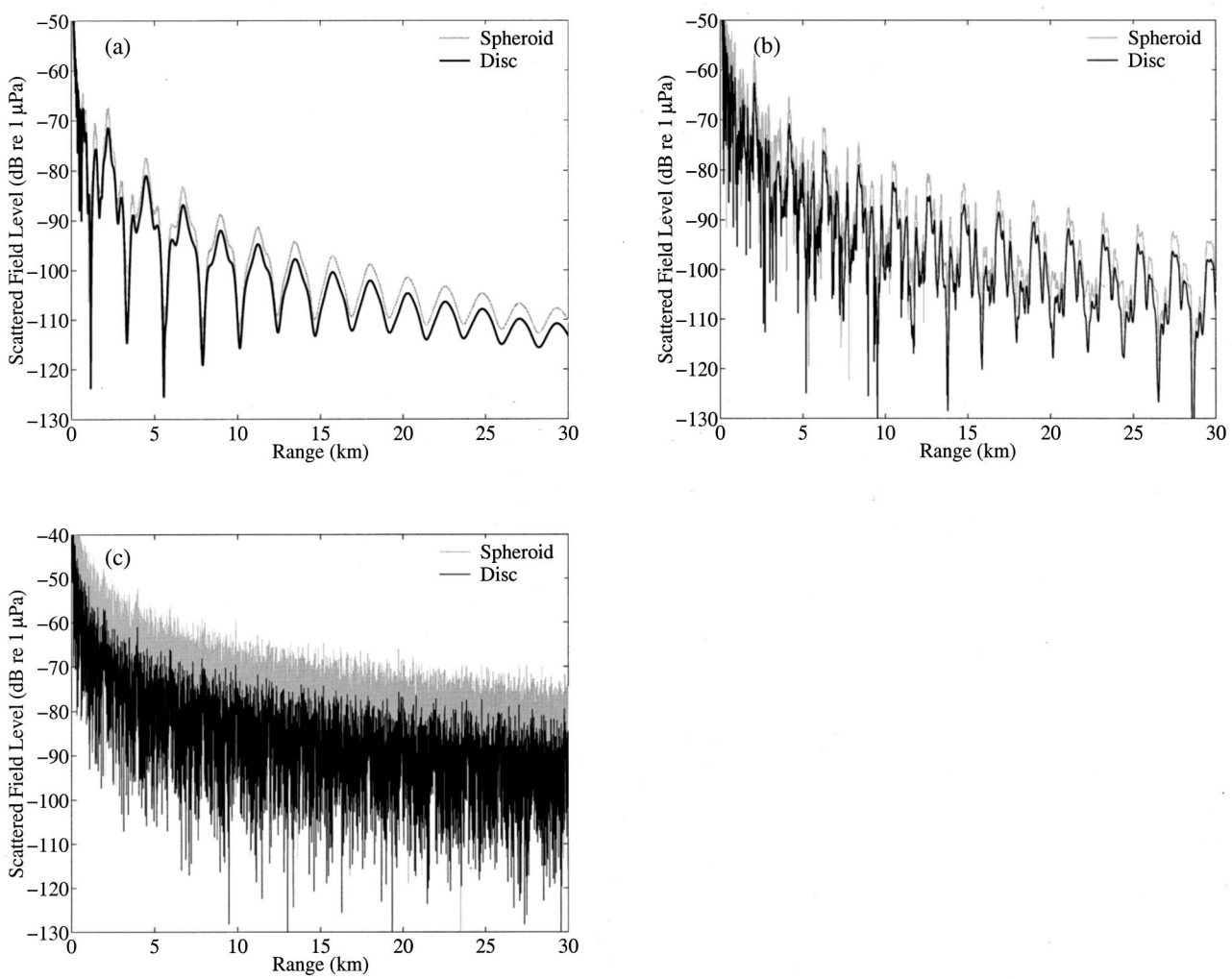


FIG. 19. Similar to Figs. 18(a)–(c) but for a pressure-release prolate spheroid of aspect ratio 2 with a major axis of 40 m and a minor axis of 20 m and a rigid circular disk of radius 10 m. *ka* is 12.6 for the disk and for the spheroid in the bow aspect in these examples. The geometry of the setup is similar to Figs. 2 and 11, except that the receiver is in the forward azimuth.

For the 10 m radius pressure-release sphere in Fig. 10(c), the forward scatter function's peak is constant because the projected area of the sphere is independent of the angle of incidence. This is not true for the pressure-release prolate spheroid in Fig. 13(c), where the projected area increases with the incident grazing angle. This causes the scatter function to have a forward scatter lobe with increasing peak amplitude and decreasing width.

V. BABINET'S PRINCIPLE AND FORWARD SCATTERED FIELD IN A WAVEGUIDE

Babinet's principle maintains that the forward scattered fields from impenetrable objects in free space with identical projected areas are equal for large *ka*.[10,13] This also holds true for some penetrable objects.[10] The forward scattered field from a large object in free space interferes destructively with the incident field to form a shadow directly behind the object. In the farfield, the intensity of the forward scattered field is the same pattern as that diffracted through a hole of the same projected area as the object in a rigid wall. The forward scattered field then only depends on the projected area of the object. It is therefore possible to replace the 3-D object by the largest 2-D cross section of the object normal to the incident wave. This is Babinet's principle. For example, in free space, a sphere, and a circular disk of the same radius, with the disk aligned normal to the incident wave vector, have the same projected area, and hence identical forward scattered fields when $ka \gg 1$.

We find that Babinet's principle is approximately valid in forward *azimuth* in a waveguide if the projected area of the object does not vary significantly for incident plane waves over $\pm \Delta\psi$. We stress that this holds in the forward *azimuth* because there is generally no unique forward direction in a waveguide. This result is illustrated in Figs. 17–20. The projected areas for an upright disk, sphere, and prolate spheroid are shown as a function of horizontal grazing angle in Fig. 17, where the spheroid is oriented as in Sec. IV C. All three objects have the same projected area for a horizontally propagating plane wave. The spheroid deviates the most from its flat counterpart, the disk, as the grazing angle changes. The corresponding values of $\Delta\psi$ over range can be determined from Fig. 2(d) for the sand and silt waveguides, and are typically less than 30°, where the variation in projected area is small for all three objects. For the perfectly

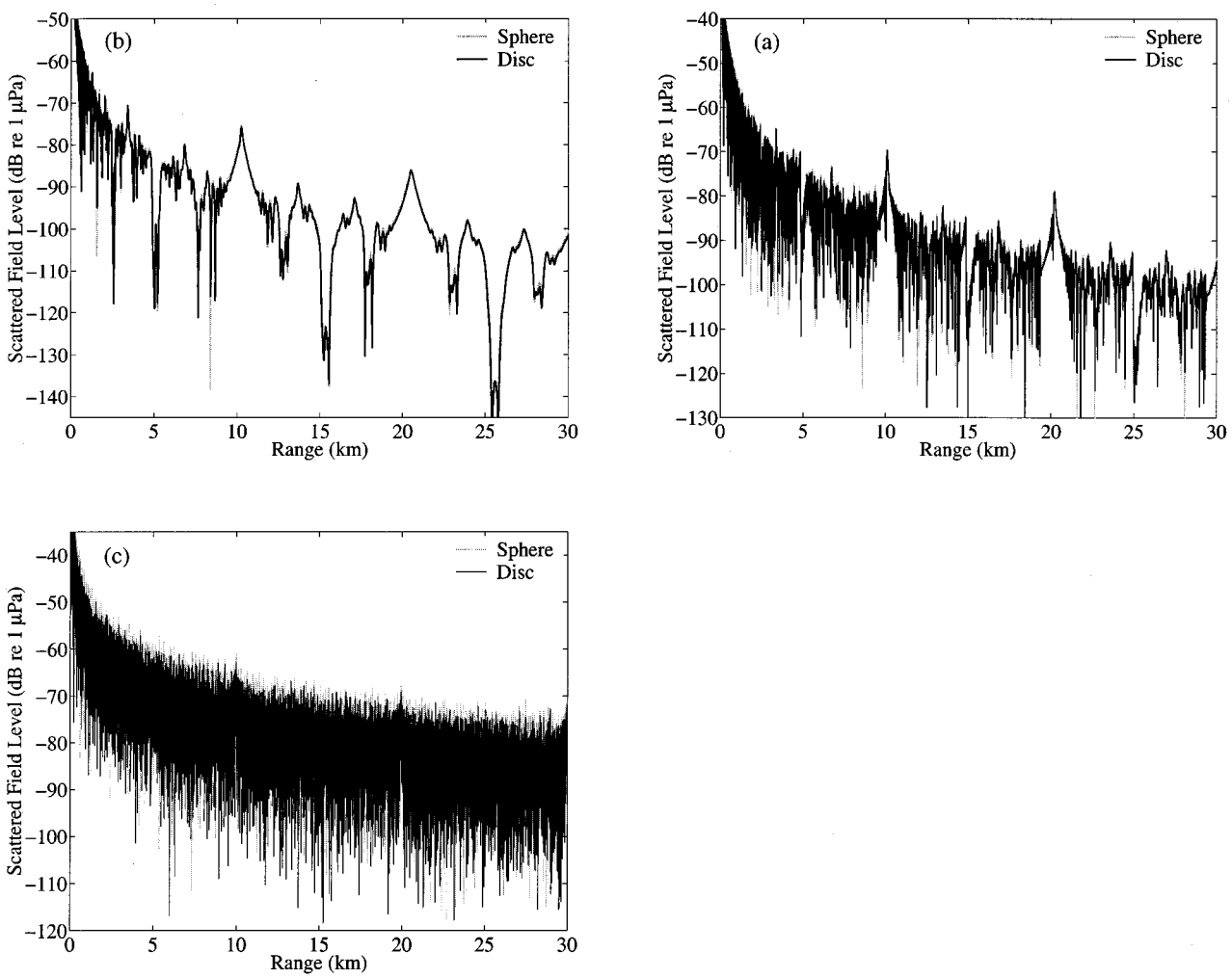


FIG. 20. Similar to Fig. 18 but at 1500 Hz. *ka* is 62.8 for these examples.

reflecting waveguide, the variation in projected area is very large for all three objects, since $\Delta\psi$ is 90°. Babinet's principle is then expected to be valid for these objects in the silt and sand waveguides at high *ka*, but not in the perfectly reflecting waveguide. This is found to be the case in Figs. 18–20, where in Figs. 18–19 *ka* is not large enough for Babinet's principle to provide a better than crude approximation, even in freespace. (The spheroid's scatter function is computed using spheroidal wave functions that are known to be numerically unstable for large arguments[17] and so the scattered field at the higher *ka* value of 62.8 for the spheroid is not shown.)

When Babinet's principle holds in a waveguide, the sonar equation will only be valid in the forward scatter azimuth if the condition $2\Delta\psi<\lambda/2L$ is satisfied, regardless of object shape. This means that rounded objects such as spheres have no advantage over flat objects in attaining the sonar equation in the forward azimuth in a waveguide.

VI. CONCLUSION

As a general conclusion, we find that *the sonar equation is valid when the target's scatter function is roughly constant over the equivalent horizontal grazing angles* $\pm\Delta\psi$ *spanned by the dominant waveguide modes*. This is approximately true (1) for all objects when $2\Delta\psi<\lambda/2L$ and (2) for spheres and certain other rounded objects in nonforward scatter azimuths, even when (1) does not hold. For homogeneous convex objects condition (1) is the less stringent $2\Delta\psi<\lambda/L$.

The sonar operator has the ability to lower the frequency of transmission until the target's scatter function becomes approximately constant over $\pm\Delta\psi$. The sonar equation then becomes valid when $f<c/(4L\ \Delta\psi)$. Operating in this frequency regime is desirable because when the sonar equation is valid, only a single-parameter, target strength, is necessary to characterize the scattering properties of the target. This greatly simplifies target classification in a shallow water waveguide by making the traditional approach[1,2] valid.

We find that Babinet's principle is approximately valid in the forward *azimuth* in a waveguide if the projected area of the object does not vary significantly for incident plane waves over $\pm\Delta\psi$. When Babinet's principle holds in a waveguide, the sonar equation will only be valid in the forward scatter azimuth if the condition $2\Delta\psi<\lambda/2L$ is satisfied regardless of object shape. This means that rounded objects such as spheres have no advantage over flat objects in attain-

ing the sonar equation in the forward azimuth in a waveguide.

[1] R. J. Urick, *Principles of Underwater Sound* (McGraw-Hill, New York, 1983).
[2] National Defence Research Committee, *Physics of Sound in the Sea* (Peninsula, Los Altos, 1989).
[3] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, Woodbury, NY, 1994).
[4] W. S. Burdic, *Underwater Acoustic System Analysis* (Prentice–Hall, Englewood Cliffs, NJ, 1991).
[5] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982).
[6] N. C. Makris and P. Ratilal, ''Validity of the sonar equation and Babinet's principle for object scattering in a shallow water waveguide,'' J. Acoust. Soc. Am. **106**, 2158 (1999).
[7] F. Ingenito, ''Scattering from an object in a stratified medium,'' J. Acoust. Soc. Am. **82**, 2051–2059 (1987).
[8] N. C. Makris and P. Ratilal, ''A unified model for reverberation and submerged object scattering in a stratified ocean waveguide,'' J. Acoust. Soc. Am. **109**, 909–941 (2001).
[9] N. C. Makris, ''A spectral approach to 3-D object scattering in layered media applied to scattering from submerged spheres,'' J. Acoust. Soc. Am. **104**, 2105–2113 (1998); **106**, 518 (1999) (Erratum).
[10] H. C. van de Hulst, *Light Scattering by Small Particles* (Dover, New York, 1981).
[11] N. C. Makris, F. Ingenito, and W. A. Kuperman, ''Detection of a submerged object insonified by surface noise in an ocean waveguide,'' J. Acoust. Soc. Am. **96**, 1703–1724 (1994).
[12] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1986), pp. 418–436.
[13] M. Born and E. Wolf, *Principles of Optics, Electromagnetic Theory of Propagation Interference and Diffraction of Light*, 6th ed. (Cambridge University Press, Cambridge, 1980).
[14] J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, in *Electromagnetic and Acoustic Scattering by Simple Shapes* (North-Holland, Amsterdam, 1969).
[15] P. Ratilal, ''Remote sensing of submerged objects and geomorphology in continental shelf waters with acoustic waveguide scattering,'' Ph.D. thesis, MIT, Cambridge, MA, 2002.
[16] P. Ratilal and N. C. Makris, ''Extinction theorem for object scattering in a stratified medium,'' J. Acoust. Soc. Am. **110**, 2924 (2001).
[17] S. Asano, ''Light scattering properties of spheroidal particles,'' Appl. Opt. **18**, 712–723 (1979).

Measurements and modeling of acoustic scattering from partially and completely buried spherical shells

A. Tesei,[a] A. Maguer,[b] and W. L. J. Fox[c]
SACLANT Undersea Research Centre, Viale San Bartolomeo 400, 19138 La Spezia, Italy

R. Lim
Coastal Systems Station/Dahlgren Naval Surface Warfare Centre, Code R21, 6703 West Highway 98, Panama City, Florida 32407-7001

H. Schmidt
Department of Ocean Engineering, MIT, Cambridge, Massachusetts 02139

(Received 27 August 2001; revised 28 May 2002; accepted 29 July 2002)

The use of low-frequency sonars (2–15 kHz) is explored to better exploit scattering features of buried targets that can contribute to their detection and classification. Compared to conventional mine countermeasure sonars, sound penetrates better into the sediment at these frequencies, and the excitation of structural waves in the targets is enhanced. The main contributions to target echo are the specular reflection, geometric diffraction effects, and the structural response, with the latter being particularly important for man-made elastic objects possessing particular symmetries such as bodies of revolution. The resonance response derives from elastic periodic phenomena such as surface circumferential waves revolving around the target. The GOATS'98 experiment, conducted jointly by SACLANTCEN and MIT off the island of Elba, involved controlled monostatic measurements of scattering by spherical shells which were partially and completely buried in sand, and suspended in the water column. The analysis mainly addresses a study of the effect of burial on the dynamics of backscattered elastic waves, which can be clearly identified in the target responses, and is based on the comparison of measurements with appropriate scattering models. Data interpretation results are in good agreement with theory. This positive result demonstrates the applicability of low-frequency methodologies based on resonance analysis to the classification of buried objects. [DOI: 10.1121/1.1509425]

PACS numbers: 43.30.Jx, 43.40.Fz, 43.20.Tb [DLB]

I. INTRODUCTION

At the high frequencies traditionally used for mine countermeasures (MCM), typically in excess of 30 kHz, the acoustic scattering by moderate-sized bottom targets ($\sim$1 m) is well described by the geometrical theory of diffraction. The reason for using these high frequencies is the strong dependence of prevailing technology on imaging resolution for detection, and more importantly, for classification. However, it is well established that at lower frequencies, man-made targets in particular may support the excitation of strong structural waves or resonances consistent with their structural symmetries. For target sizes $\sim$1 m, these resonances can be strongly excited in the 1–10 kHz frequency regime, well below the frequencies required for classification by imaging.

Analysis of wave scattering by underwater elastic objects in the $4<ka<30$ range (where k is the acoustic wave number and a is the characteristic dimension of the object) has identified several target classifying features that are reflected in the scattering, e.g., the radius, the wall thickness, and the elastic properties for spherical shells.[1,2] However, the question is whether features based on the spatial and temporal characteristics of the structural response can provide robust and reliable classification clues. If they can, this will provide a technology push for including lower-frequency sonars in operational systems.

Another benefit of moving to lower sonar frequencies is the improved seabed penetration. For insonification above the seabed critical grazing angle (assuming a flat seabed), the attenuation decreases approximately linearly with frequency, while below the critical grazing angle, the evanescent lateral wave field in the bottom penetrates deeper.[3] Also, the reverberation from surficial features such as small-scale sand ripples will decrease with decreasing frequency, with an increase in signal-to-reverberation as a result.[4]

Here, a coordinated experiment and modeling effort explores the potential of low- to medium-frequency sonar concepts for improved target detection and classification. Specifically the analysis aims at verifying through direct comparison of experimental data and wave theory scattering models whether the circumferential waves backscattered by an elastic, axisymmetric target remain detectable and identifiable whether the target is proud or buried. Further, the analysis evaluates how the burial depth, the sediment properties, and the seabed interface influence the dynamics of the generated structural responses.

[a]Electronic mail: tesei@saclantc.nato.int
[b]Present address: Thales Underwater Systems, 274 Victoria Road, Rydalmere NSW 2116, Australia.
[c]Present address: Applied Physics Laboratory, 1013 NE 40th St., Seattle, WA 98105-6698.

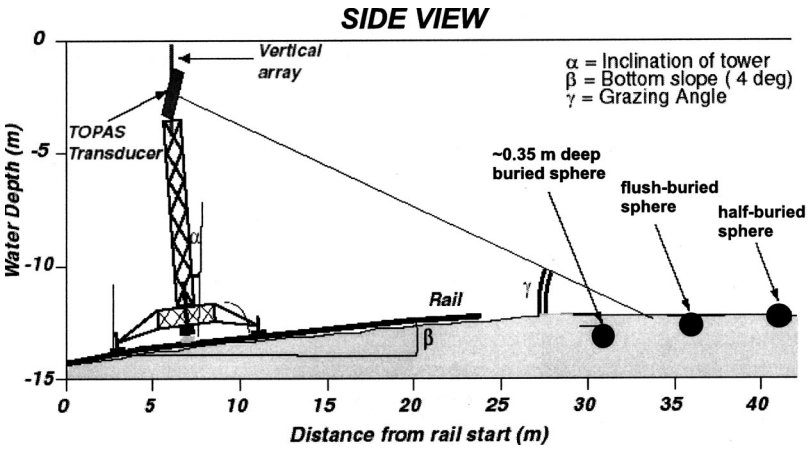


FIG. 1. GOATS'98 experimental configuration.

Past experimental and modeling work has provided evidence for the excitation of structural waves in completely buried spherical shells, including investigations on the sensitivity to burial depth, sediment type, and grazing angle.[5,6] The excitation of target resonances for flush and partially buried targets has more recently been investigated using new modeling capabilities,[6,7] but equivalent experimental work has been lacking.

The analysis presented here uses the comprehensive data set acquired during the GOATS'98 experiment.[8] Monostatic and bistatic measurements were conducted on a suite of man-made targets of different shapes and burial depths. The targets were insonified by a parametric projector that scanned over a wide range of grazing angles from below critical to above. This paper will be limited to the analysis of backscattering with insonification above critical grazing for three identical spherical shells: one half buried, one flush buried, and one completely buried. The spherical shells were relatively thin-walled and filled with air. Hence, the free-field scattering physics is relatively simple to model and well understood.[9–11] The analysis will use the suite of scattering models presented in Refs. 5–7.

II. EXPERIMENTAL CONFIGURATION

The GOATS'98 experiment[8] was carried out in May, 1998, on a sandy bottom in 12–15 m water near Marciana Marina, off the island of Elba, Italy. The TOPAS (Topographic Parametric Sonar) source was used to insonify the targets with a highly directional beam. The sonar is an end-fire parametric array of 24 horizontal staves. It covers the frequency range 2–16 kHz for the secondary frequency and 35–45 kHz for the primary frequency. The transmitting source level is about 238 dB *re*: 1 μPa at 1 m for the primary frequency and the source level obtained at the difference frequencies varies from 190 to 207 dB *re*: 1 μPa at 1 m (see Ref. 12). A short single pulse is obtained by transmitting a weighted burst at the primary frequency. At the secondary frequency the vertical beamwidth is 3°–5° and the horizontal is about 8°. As with all parametric sonars, there are no conventional sidelobe structures, and the beam pattern decays monotonically outside the main 3 dB down beamwidth. The far field is estimated to start at about 30–35 m in front of the transducer, while the volume of nonlinear interaction[13] is estimated to extend for the first 11 m (see Ref. 3). All the measurements were at ranges larger than the nonlinear interaction length of the sonar, so that nonlinear penetration effects can be ignored in the data analysis.

The experimental configuration is shown in Fig. 1. In order to acquire data from various source-receiver geometries, the transmitter was mounted on a 10 m tower. The tower in turn was mounted on a 24 m linear rail on the bottom, along which its position could be precisely controlled. The TOPAS transmitter was mounted in a pan-and-tilt assembly so that the transmission direction could be accurately measured. A linear receiving array of 16 equally spaced omnidirectional hydrophones was mounted vertically in a near-monostatic configuration. The array was designed to work at 8 kHz; hence, the distance between adjacent elements is 0.094 m generating an acoustic aperture of 1.41 m. The main lobe of its beam pattern is about 6° at 8 kHz, 26° at 2 kHz, and 3° at 15 kHz. The sidelobes are more than 13 dB lower than the main lobe. The sensitivity of all the hydrophones was approximately constant (around −163 dB) in the bandwidth 1–15 kHz.

Three identical spherical shells were deployed in line with the TOPAS rail at different burial depths (about 35 cm into the sediment, flush, and half-buried). One of the shells was measured before burial under quasi-free-field conditions suspended in the water column. They were air-filled, steel spherical shells nominally of 0.53 m outer radius and 3 cm wall thickness. As the relative thickness h (defined as the ratio between wall thickness, d, and outer radius, a) was equal to 0.0566, they are classified as thin (though others may consider this an intermediate thickness[14] tending to thin). They had a steel lug at the top which was used for suspension and transportation. The spherical shells were constructed by welding two hemispherical shells together. The nonuniformity of the resulting shell at the weld point was not measured, and could be a source of unknown acoustic effects in the following results.

Isospeed conditions were measured in the water column at 1520 m/s. The depth of the water over the target field was about 12 m. Grain size estimates of the bottom corresponded to medium sand. The average density of the sediment, measured over the volume of bottom cores, was 1.91 g/cm^3. The sediment propagation loss was estimated around 0.5 dB/λ, as demonstrated in Refs. 3 and 15. Sound speed was estimated

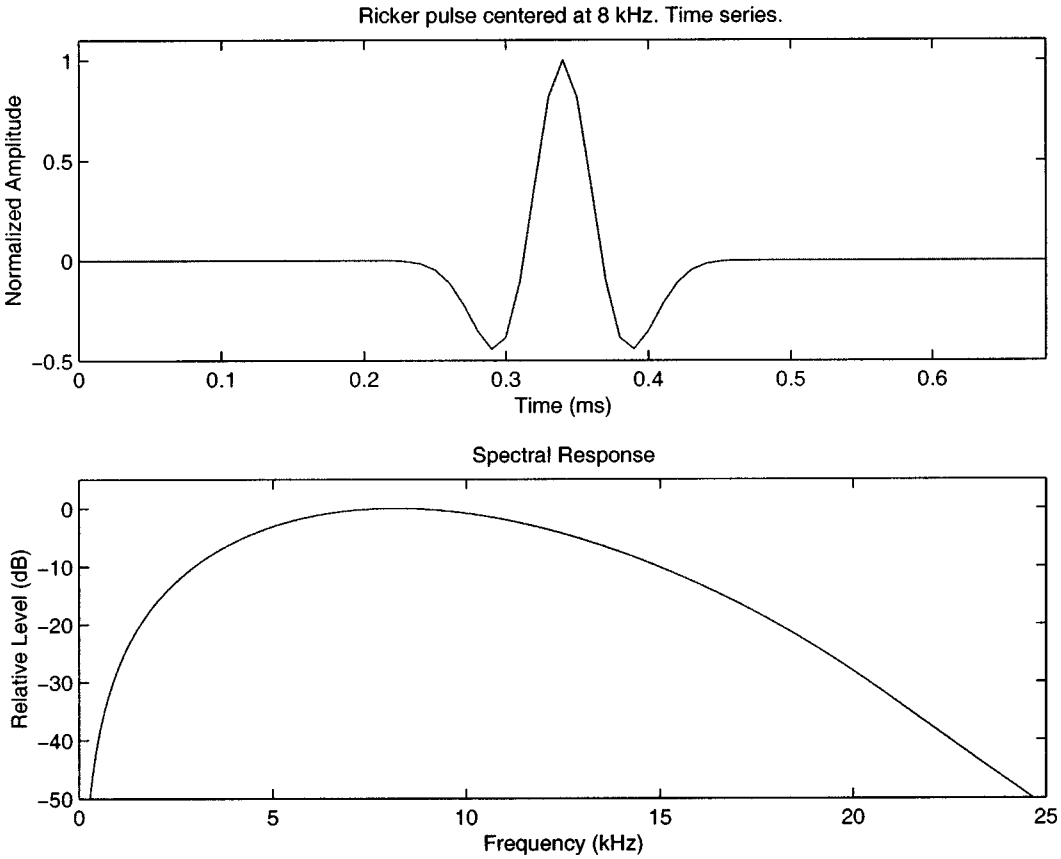


FIG. 2. Model of a Ricker pulse centered at 8 kHz.

to be on the order of 1720 m/s by pulsed travel time measurements on sediment cores centered at 200 kHz. However, the analysis of low-frequency sound waves emitted by the TOPAS and received on a vertical array of buried hydrophones in the immediate vicinity of the GOATS'98 target field has proved the strong dispersion of acoustic waves in a sandy seabed.[15] From those acoustic measurements the sediment sound speed was estimated to be approximately 1640 m/s at 8 kHz. This dispersion is consistent with the predictions of Biot's theory for propagation in porous media. These considerations also influence the estimation of the sediment critical angle which should consequently decrease as the frequency is reduced; thus, the nominal critical angle of the sediment is estimated at 28° from the measurements on cores at 200 kHz and at 22° if a sound speed of 1640 m/s is assumed.

III. TARGET SCATTERING MODELS

Scattering models for a spherical shell either in the free field (i.e., suspended in the water column) or buried in the seabed are briefly presented. Models are developed in the frequency domain and provide the scatterer transfer function. In order to compare simulated and real sonar data in the time domain, the simulations of the transfer function are multiplied in the frequency domain by the model of the incident pulse transmitted by the acoustic source, and the resulting products are inverse Fourier transformed.[16] A Ricker function, defined as the second derivative of a Gaussian,[17] is used for the incident pulse at a nominal center frequency of 8 kHz (Fig. 2). The time signal is normalized to unit peak amplitude and its spectral level is normalized to 0 dB maximum. The simulations are computed to a maximum frequency of 25 kHz, but the results are presented in the bandwidth 0–15 kHz and then analyzed between 2 and 15 kHz, where the signal-to-interference ratio of the at-sea data is sufficient to allow model-data comparison and elastic wave analysis. In the simulated and measured target scattering response plots, the time series will be normalized with respect to the maximum amplitude of the specular echo, and the signal spectral levels will be normalized to 0 dB maximum.

A. Spherical shell in free field

The free field case is used as for comparison of the results obtained when the spherical shells are partially or completely embedded in the sediments. Under free field conditions, the acoustic pressure p_s scattered by an elastic fluid-filled spherical shell, insonified by an incident plane wave with acoustic pressure $p_i = p_0 e^{i(kr-\omega t)}$, can be presented in a series form:[18]

$$p_s = p_0 e^{-i\omega t} \sum_{n=0}^{\infty} i^n (2n+1) A_n h_n^{(1)}(kr) P_n(\cos\theta), \qquad (1)$$

where p_0 is the pressure amplitude of the incident wave, k is the wave number, r is the range, θ is the scattering direction angle, n is the modal order, $h_n^{(1)}$ is the spherical Hankel function of nth order, A_n are coefficients to be determined from the boundary conditions,[18] and P_n is the Legendre polyno-

mial of nth order. The scattered field can be also expressed in terms of the form function $F(\theta)$:

$$p_s = p_i\left(\frac{a}{2r}\right)F(\theta), \tag{2}$$

where a is the spherical shell's outer radius. In the far field (i.e., for $r\to\infty$), the backscatter ($\theta=\pi$) form function $F(\pi)$, expressed as a sum of partial waves $F_n(\pi)$, becomes

$$F(\pi)=\sum_{n=0}^{\infty} F_n(\pi)=\frac{2}{ika}\sum_{n=0}^{\infty}(-1)^n(2n+1)A_n. \tag{3}$$

Scattering effects due to excitation of elastic waves on the shell are analyzed in terms of wave speed dispersion curves. The phase ($c_{\mathrm{ph},n}$) and group ($c_{g,n}$) speeds of the nth modal frequency f_n of a circumferential wave, taking a path of radius R, are defined, respectively, as[18–20]

$$\begin{aligned} |c_{\mathrm{ph},n}| &= \frac{2\pi R f_n}{n+1/2}, \\ c_{g,n} &= 2\pi R(f_n - f_{(n-1)}). \end{aligned} \tag{4}$$

In the low- to medium-frequency range, for thin air-filled spherical shells, the elastic contribution to scattering is due to the lowest-order flexural and compressional waves of the shell. These waves may be considered as the generalization of Lamb waves on a flat plate[9] for the case of a curved plate.

The symmetric (or compressional) S_0 *Lamb-type wave*[10,20] is supersonic, almost nondispersive, and travels in the shell with phase and group speeds almost coinciding and asymptotically tending to the shell material membrane speed, c^*_{shell}. It reradiates at each revolution with an angle corresponding to the shell material coupling angle:

$$\theta_c^* = \arcsin\left(\frac{c_{\mathrm{ext}}}{c^*_{\mathrm{shell}}}\right), \tag{5}$$

where c_{ext} is the sound speed in the medium loading the shell. The dynamics and energetic contribution to backscattering of this wave is expected to be only slightly influenced by the object outer medium (or media).

The antisymmetric (or flexural) A_0 Lamb-type wave of a spherical shell in vacuum bifurcates into two dispersive waves upon fluid loading. Of the two, the wave that more strongly influences the acoustic scattering amplitude of the target is the subsonic A_{0-} *Lamb-type wave*.[10,21] At low frequencies (until its phase speed approaches the outer-medium sound speed at the so-called coincidence frequency, f_c), it is flexural in nature,[21] and the effect of fluid loading on its dynamics is essentially inertial. In this frequency region it is only slightly influenced by the outer-medium properties, like the S_0 wave. Around the coincidence frequency the so-called midfrequency enhancement occurs as a result of trace velocity matching[22] when the flexural wave phase speed approaches that of the exterior diffractive field. In the midfrequency enhancement region the shell response amplitude reaches its maximum at the coincidence frequency. The A_{0-} wave starts to behave like a fluid-borne wave, becoming difficult to detect with increasing frequency because increased radiation damping makes it harder for the wave to travel all the way around the shell to radiate significantly in the back direction. Around the coincidence frequency it reradiates at each revolution approximately along the tangent to the sphere cross section. Its group speed reaches its maximum at the coincidence frequency. Hence, at frequencies around and beyond the coincidence frequency the A_{0-} wave dynamics and energy level are strongly influenced by the speed of sound in the exterior.

Around the coincidence frequency the A_{0-} wave is phase-coupled with the other antisymmetric Lamb-type wave, namely the A_{0+} *Lamb-type wave*,[10,20] the nature of which is opposite to the A_{0-} wave,[21] as for frequencies lower than the coincidence frequency it is fluid-borne and highly dispersive, hence highly radiation damped and almost undetectable. When its phase speed approaches the external sound speed, it changes nature from fluid-borne to flexural (shell-borne). Beyond the coincidence frequency the A_{0+} wave becomes supersonic and less dispersive, hence more easily detectable from backscattered data.

Figure 3 shows the simulated response of a water-loaded spherical shell to a Ricker pulse centered at 8 kHz. The interpretation of the simulated data is shown in terms of S_0, A_{0-}, and A_{0+} wave echoes in time and resonance modes in frequency. Here and in the following analysis, the modeled and measured spherical shells have the same nominal elastic parameters (steel compressional speed $c_p=5950$ m/s, shear speed $c_s=3240$ m/s, membrane speed $c^*_{\mathrm{steel}}=5435$ m/s, density $\rho=7.7$ g/cm^3), the density of the sea water is set to 1 g/cm^3 and its sound speed is set to the measured value 1520 m/s.

In order to maintain high signal-to-reverberation levels even in the case of buried targets, the data analysis focuses on a low- to intermediate-frequency range that stops approximately at the end of the region of midfrequency enhancement, hence extending over a bandwidth where the S_0 and A_{0-} waves are more easily detectable and only a few modes of the A_{0+} wave can be identified (as shown in Fig. 3). As a consequence, only the S_0 and A_{0-} waves will be considered in the following. Figure 4 shows the travel paths followed by the S_0 and A_{0-} Lamb-type waves revolving around the spherical shell in the free field. For thin-walled spherical shells the radius R of the circumferential wave path corresponds to the shell outer radius a for both the S_0 and the A_{0-} waves:

$$R_{S_0}=R_{A_{0-}}=a. \tag{6}$$

A simple empirical formula, which was shown[23] to be valid for steel, aluminum, and similar materials, allows the approximate estimation of the sound speed in the external medium loading the shell, c_{ext}, from f_c and the shell wall thickness d:

$$c_{\mathrm{ext}}\approx f_c 2\pi d. \tag{7}$$

This expression is related to the classical observation that the peak of the midfrequency enhancement on thin spherical shells scales with a constant frequency-thickness product. It will be used in lieu of local measurements of c_{ext} around the shell, which can be difficult in buried underwater configura-

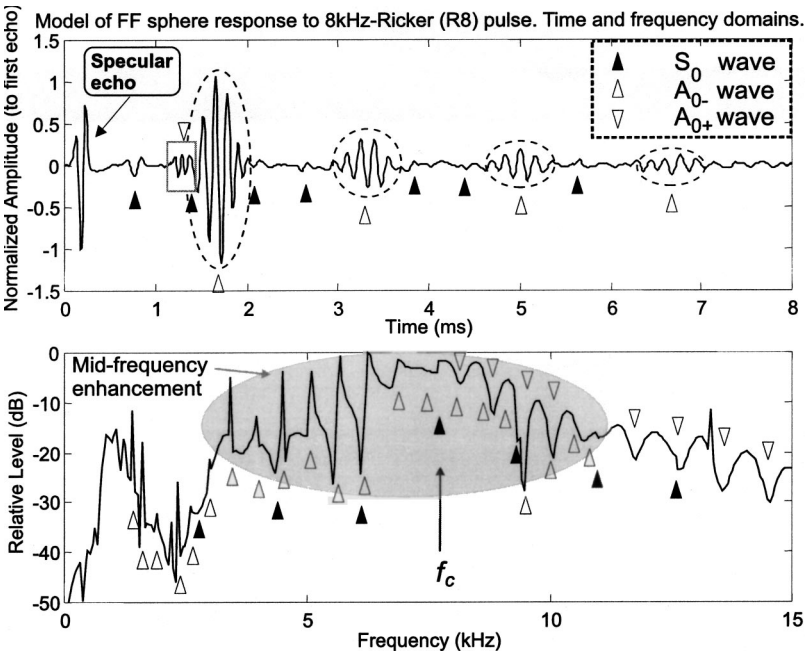


FIG. 3. Model of backscattering by free-field spherical shell. Elastic scattering analysis is applied.

tions. The viability of this approximation will be discussed in the following analysis.

B. Partially and completely buried spherical shells

Measurements of the fully and partly buried spherical shells are compared with model predictions based on transition- ($T-$) matrix solutions for the scattered field.[5,7] These solutions assume the bottom to be fluid with a flat water/sediment interface. While this assumption may not be strictly satisfied in the measurements, it is expected to provide better predictions of the scattering response than ignoring the environmental layering. The basic solutions are generalizations of the free-field spectral solution given by Boström[24,25] and take the form:

$$\mathbf{u}^{\mathrm{sca}}(\mathbf{r}) = \sum_{j=1}^{N} \sum_{pml} \boldsymbol{\Psi}_{pml}(\mathbf{r}-\mathbf{d}^{(j)})\,\gamma_{pml}^{(j)} = \sum_{j=1}^{N} \sum_{pml} \sum_{p'm'l'} \boldsymbol{\Psi}_{pml}(\mathbf{r}-\mathbf{d}^{(j)}) \times \mathbf{T}_{pmlp'm'l'}^{(j)} \alpha_{p'l'm'}(\mathbf{r}_s). \quad (8)$$

The field scattered by an obstacle embedded in N ($N \geqslant 1$) layers of a plane-stratified environment is formulated as a sum over a contribution from the segment in each layer. An $e^{-i\omega t}$ steady-state time dependence is assumed but suppressed. The T matrix of the obstacle segment penetrating the jth of the N host layers in contact with the obstacle is defined with the second equality as the operator, $\mathbf{T}^{(j)}$. It projects the vector of incident field coefficients, $\alpha(\mathbf{r}_s)$, onto the vector of expansion coefficients, $\gamma^{(j)}$, for the field scattered by the segment contained in layer j. Special basis sets defined in Ref. 26, $\boldsymbol{\Psi}_{pml}(\mathbf{r}-\mathbf{d}^{(j)})$, satisfying the host's boundary conditions and centered on each segment at $\mathbf{d}^{(j)}$, are used to expand the scattered field. The incident field coefficients for a point source at $\mathbf{r}_s$ are also proportional to this basis, i.e., $\alpha_{pml}(\mathbf{r}_s) \propto \boldsymbol{\Psi}_{pml}(\mathbf{r}_s)$.[26] Directional vertical line sources can be modeled by multiplying these coefficients by a phase factor to steer the beam with and applying a simple integral over the length of the source.[5] Using this basis to express the fields external to the scatterer, transmission through the seabed to and from the scatterer is properly accounted for.

One can accommodate an obstacle with a layered core such as a steel shell by expressing $\mathbf{T}^{(j)}$ via the following matrix formula:

$$\mathbf{T}^{(j)} = -[\hat{\mathbf{L}}^{(j)}(\hat{\mathbf{D}}+\mathbf{T}^0\mathbf{D})^{-1}(\hat{\mathbf{E}}+\mathbf{T}^0\mathbf{E})-\mathbf{M}^{(j)}] \times [\mathbf{L}(\hat{\mathbf{D}}+\mathbf{T}^0\mathbf{D})^{-1}(\hat{\mathbf{E}}+\mathbf{T}^0\mathbf{E})-\mathbf{M}]^{-1}, \quad (9)$$

where the matrices $\hat{\mathbf{D}}$, $\hat{\mathbf{E}}$, $\mathbf{L}$, $\hat{\mathbf{L}}^{(j)}$, $\mathbf{M}$, and $\hat{\mathbf{M}}^{(j)}$ are as defined in Ref. 26 and the matrices $\mathbf{D}$ and $\mathbf{E}$ differ from $\hat{\mathbf{D}}$ and $\hat{\mathbf{E}}$, respectively, only in replacing regular partial-wave functions with outgoing ones. The interior structure of the obstacle is contained in $\mathbf{T}^0$, the free-field scattering T matrix of the core embedded in an unbounded elastic medium specified with the material parameters of the outer layer of the obstacle. Thus, $\mathbf{T}^0 = \mathbf{0}$ if the obstacle is homogeneous. The form of $\mathbf{T}^0$

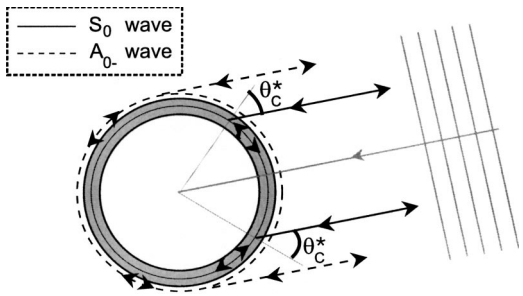


FIG. 4. Simplified scheme of travel paths for the S_0 and A_{0-} Lamb-type elastic waves backscattered by an air-filled spherical shell around the coincidence frequency. The ray diagram is based on the ray theory approximation for fluid-loaded elastic objects (Refs. 20 and 33). The travel path of the A_{0-} wave is drawn outside the shell in order to emphasize its fluid-borne nature around the coincidence frequency. The reradiation angles are outlined for the two wave types. The wall thickness is not in scale.

has been presented previously for elastically layered cores[27] as well as fluid cores.[28] When $N=1$, Eqs. (8) and (9) reduce to the formulation presented in Ref. 5 for fully buried obstacles. Translation operators may be used to extract from the **L** and **M** matrices terms that make up the rescattering matrix evident in the solution of Ref. 5. The rescattering matrix accounts for multiple interactions of the scatterer with the interfaces of the environment. For the present applications, we leave these interaction terms embedded in the **L** and **M** matrices to mitigate convergence problems that arise with partial burial configurations.

While the segmented *T*-matrix formulation contained in Eqs. (8) and (9) may be used to calculate the field scattered by layered inhomogeneous obstacles that penetrate a plane-layered host, the obstacle layering should be sufficiently smooth and obey certain convergence restrictions. In particular, for surface fields expanded in basis functions of spherical geometry, the origin-centered sphere that circumscribes the inner boundary of each homogeneous obstacle layer should never overlap the origin-centered sphere that inscribes its outer boundary. This restriction becomes a problem for elongated shells but the acoustic response of spherical shells can be calculated effectively.

For comparisons with measurements, Eqs. (8) and (9) were implemented using the numerical techniques described in Ref. 26. All surface and wave number integrations needed to evaluate the basis functions and the *T* matrix were performed by applying 32 point Gauss quadrature to a sufficient number of segments of the specified integration contours to attain a stable, converged result. Matrix products involving inversions were performed with Gaussian elimination. Double precision complex arithmetic was used in all calculations. A resolution of 400 frequency points was maintained in the predicted spectra and the scattered field at each frequency point was evaluated with a maximum truncation of 50 l and m values. The field calculations are scaled by the in-water source amplitude at 1 m. Only one bottom layer was assumed $(N=1)$.

A simplified scheme of the travel paths of the S_0 and A_{0-} elastic waves around the coincidence frequency is shown in Fig. 5 in the case of partial burial, which is the most complicated geometry. Under partial burial conditions, the reradiation angle for each wave in either medium can vary with the shell coupling, i.e., as the outer-medium sound speed varies, trace velocity-matching[20,22] determines the point at which the incident field couples to the shell and, consequently, the reradiation angle.

Theoretically,[5,7] the S_0 wave is not expected to be greatly influenced by burial, except for a slight shift toward lower frequencies of its first resonance modes, due to the significantly greater inertial loading of the shell in the sediment, having a larger relative density than water. The shift is expected to be more evident for the first mode and to decrease as the modal order (and frequency) increases. Further, the frequency shift should increase with the percentage of target surface in contact with sediment. The A_{0-} wave is expected to behave like the S_0 wave at low frequencies, where it is subject to inertial loading. Therefore, its first free-field modes should also tend to shift to lower frequencies. As

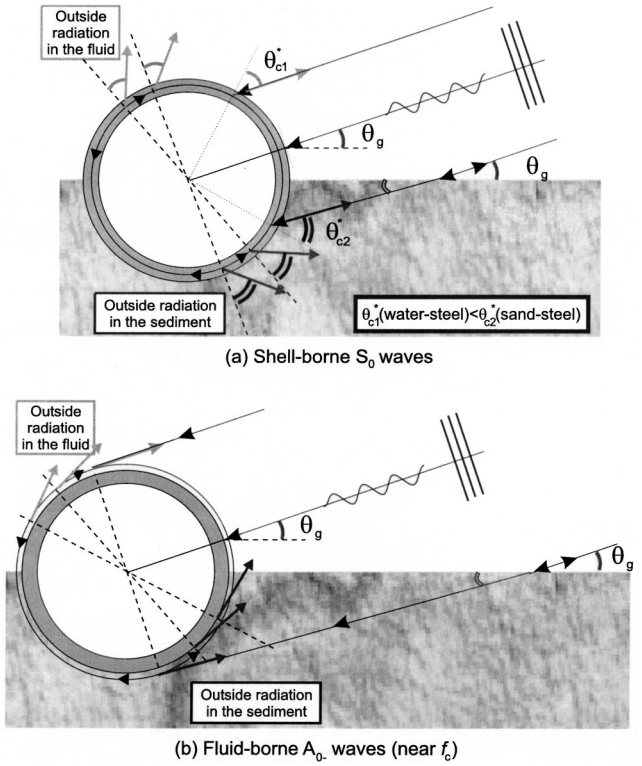


FIG. 5. Simplified travel paths of the shell-borne S_0 (a) and fluid-borne A_{0-} (b) waves in the case of partial burial. As in Fig. 4 the travel path of the A_{0-} wave is drawn outside the shell in order to emphasize its fluid-borne nature around the coincidence frequency. The wall thickness is not in scale.

the A_{0-} wave becomes fluid-borne in nature, its phase speed increases to approach the exterior sound speed, which is higher when the target is completely buried. From Eq. (4), this means the modes of the A_{0-} wave that are excited in a midfrequency enhancement, as well as the coincidence frequency f_c, are expected to experience an upward shift upon burial. For partial burial, we hypothesize that these modes will again shift an amount consistent with the percentage of target surface in contact with sediment.

In line with the above-given expectations, we will treat the target exterior as an effective homogeneous medium if the shell is partially buried so that the exterior sound speed and exterior density needed to determine wave dispersion on the shell can be easily specified. Thus, an effective exterior sound speed $c_{\text{ext}}^{\text{eff}}$ is determined by generalizing Eq. (7) as

$$c_{\text{ext}}^{\text{eff}} \approx f_c 2\pi d. \tag{10}$$

This assumption should be reasonable if the dominant contribution to the backscatter from the S_0 and A_{0-} waves are from complete circumnavigations of the shell, i.e., contributions from paths around the shell that include reflections at the interface are small. Under these circumstances, $c_{\text{ext}}^{\text{eff}}$ may be expected to be comparable to a weighted harmonic average of the sound speeds of the two exterior fluids in order to account for propagation of the exterior diffracted field through both fluids.[29]

Similarly, an external effective density $\rho_{\text{ext}}^{\text{eff}}$ will be defined in the case of partial burial as an average of the densities of the two external fluids, weighted according to the fraction of the volume of the sphere loaded by each fluid.

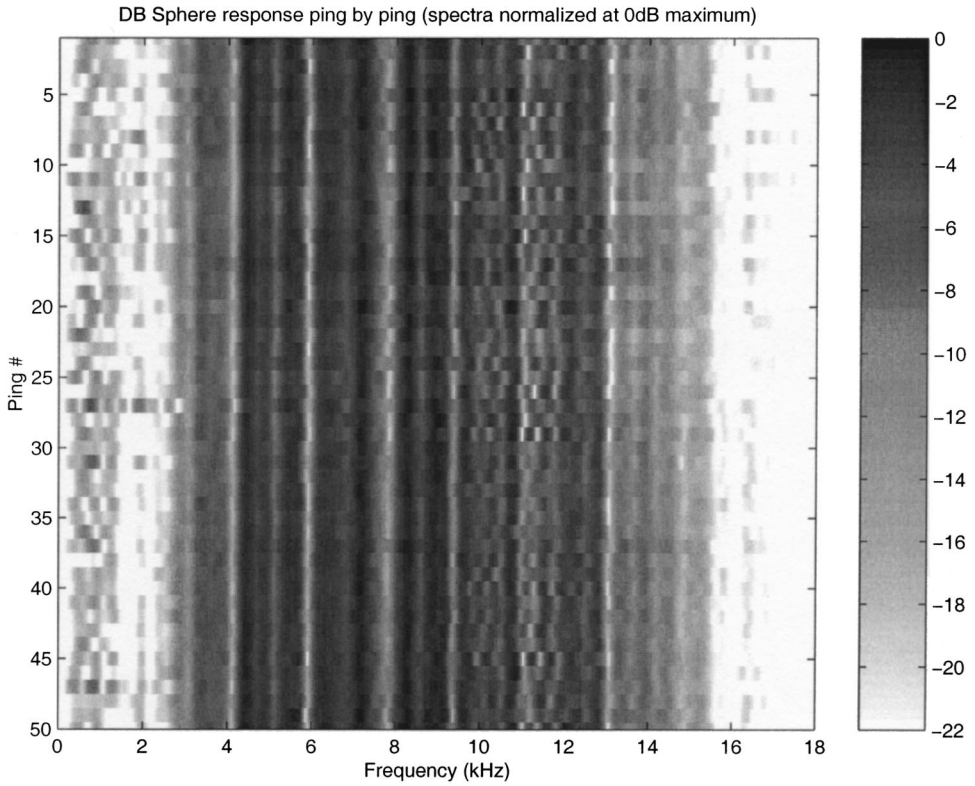


FIG. 6. Typical example of ping-to-ping fluctuation of the measured data sets. Spectrum of the 50 pings scattered from the deeply buried target.

This definition of effective density is particularly appropriate at low frequency, where the more significant impact of inertial (i.e., mass) loading suggests a weighting based on mass fraction (density multiplied by volume fraction).

Of course, these effective parameters are not meant to represent the environment in the scattering calculations for a partially buried shell. They will be used only in the algorithm outlined in the Appendix for determining the dispersion of surface guided waves on the shell from the measured or calculated backscatter signals. These definitions for $c_{\text{ext}}^{\text{eff}}$ and $\rho_{\text{ext}}^{\text{eff}}$ are convenient because they allow the determination of the dispersion properties as if the environment of the shell were homogeneous. The result should be understood to be an estimate of the average dispersion of the A_{0-} and S_0 waves experienced upon propagation around the buried and unburied sections of the shell. The rationale for expecting the resulting dispersion curves to be representative of the average dispersion is based on the intuitive expectation that the timing and coupling of waves going around the shell at any point on its surface depends on the local exterior parameters at that point only. Therefore, the time for a wave to go around the top of the shell in water is independent of the time it takes going around the bottom in the sediment. In an analogous fashion, for a pulse transmitted through layers of materials of different speed, one can determine the transmission time by summing the thickness/sound-speed ratios of each layer or treat the stack of layers as a single layer with an effective sound speed given by a harmonic average of the layer sound speeds. Both treatments give identical results for the pulse transmission time. As in that simple example, both definitions for the effective parameters in our problem have been found to yield good results in simulation.

IV. EXPERIMENTAL RESULTS

The data selected are the aligned coherent averages of 50 pings of the beamformed acquisitions by the vertical array. From a preliminary analysis of ping-to-ping data fluctuation over each data set, the measurement variability was estimated. As a typical example of the studied data sets, Fig. 6 shows the ping-to-ping variability of the spectrum of the response measured from the deeply buried spherical shell. The signal appears stable in the bandwidth 2–15 kHz, whereas below 2 kHz the signal is almost completely hidden by noise and, consequently, data analysis cannot be applied. In particular, in the range 2–15 kHz, the standard deviation of the location of the dips or peaks, corresponding to the resonance frequencies of the generated elastic waves, was estimated to be approximately 1%–3% of the mean value. As a consequence, the aligned coherent average can be used in this range for data analysis as a robust measurement of each target response.

The geometry of each measurement was selected in such a way that the target could be completely illuminated by the TOPAS central beam. However, while the free-field target could be measured in the far field of the source so that the far-field model presented in Sec. III A is applicable, the proud and buried targets were measured in the near field of the TOPAS. As the *T*-matrix model described in Sec. III B can simulate directional sound beams in the near field of the source, the TOPAS vertical beam pattern was approximated by the beam pattern of a linear vertical array with a 5°-wide main lobe in the applications of the model to these cases. Despite the imperfect simulation of the actual TOPAS beam pattern, the *T*-matrix model should provide the correct target response in terms of the resonance frequency locations and

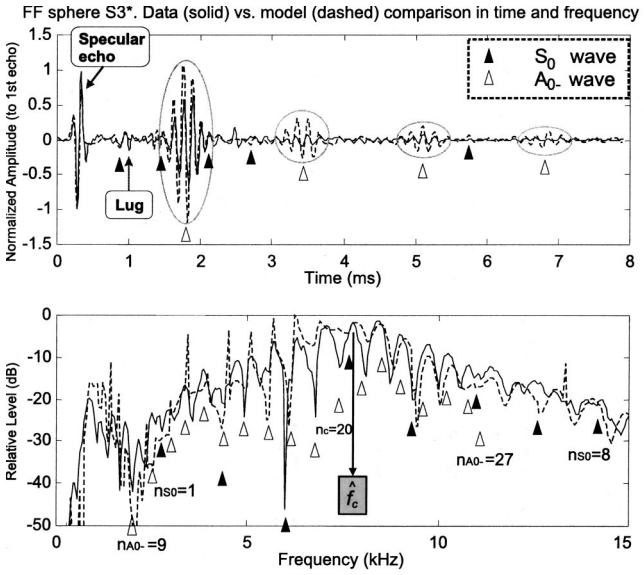


FIG. 7. Free-field spherical shell. Model-data comparison in time and frequency domains. Elastic scattering analysis by identification of S_0 and A_{0-} wave echoes in time and resonance modes in frequency. Resonance frequency modes are labeled by n_{S0} and n_{A0-}, respectively.

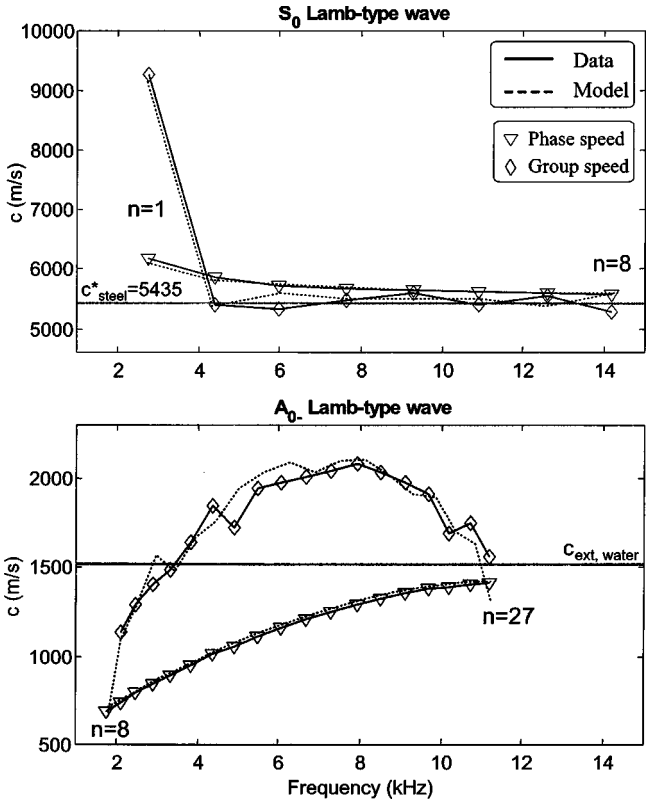


FIG. 8. Free-field spherical shell. The S_0 (top) and A_{0-} (bottom) wave dynamics is described by phase and group dispersion curves.

radiation damping of the elastic waves. However, a mismatch with the scattering amplitude can be expected.

A. Free field spherical shell (FF)

The spherical shell suspended in midwater was measured at a range of about 35 m, which was in the far field of the source and uniformly illuminated notwithstanding the high directivity of the TOPAS. The measured target response is presented in Fig. 7, where comparison with the computed model is outlined and the elastic wave analysis is superimposed. Model-data agreement is generally good in terms of signal shape, levels, and feature identification, except for an evident mismatch in the level of the A_{0-} wave features. In the frequency domain, a mismatch in the level of most of the measured A_{0-} resonance modes and a decay of about 5 dB of the midfrequency enhancement are estimated with respect to expectation. A corresponding mismatch is also evident in all the A_{0-} wave echoes in the time domain. The sharpness of the A_{0-} resonance features is reduced in the data spectrum, probably as a result of signal time windowing.[30] However, the broader mismatch in the levels of the midfrequency enhancement in frequency and of the wave echoes in time is likely caused by nonuniformities of the shell wall thickness and/or by the presence of the steel lug attached to the top of the target. The same effects will be seen also in the following, particularly in the model-data comparison of data from half- and flush-buried spherical shells.

Further, two arrivals recorded in the data time series at about 1 and 2.5 ms, respectively, were not predicted by the model. On the basis of geometrical considerations the first echo corresponds to the time of arrival predicted for the top lug of the spherical shell; the second remains unidentified.

At low frequency ($f<2$ kHz) the model-data fit is poor, as the data spectrum is 5–10 dB lower than the model prediction. This may be explained if the TOPAS transmitted a Ricker pulse with a low-frequency component lower than expected by the ideal wave form shown in Fig. 2 and used for simulations. This possibility is suggested by the high ping-to-ping variability below 2 kHz in Fig. 6. However, this hypothesis cannot be confirmed by measurements because the direct echo transmitted by the source could not be recorded under the same geometry and by the same system used for measuring the target. A low-frequency mismatch will be seen in the following model-data comparisons for all the cases considered.

The modeled and estimated dynamics of the S_0 and A_{0-} waves are compared in Fig. 8 by means of the respective dispersion curves, which have been computed from the identified wave modes (labeled n) according to the formulas in Eq. (4). The wave dynamics are in good agreement with theory for both waves. Also, unlike the scattering amplitude, wave dispersion appears quite insensitive to shell imperfections even in the case of the A_{0-} wave.

The S_0 and A_{0-} wave modes are extracted in the frequency domain, and identified on the basis of model-data comparison. The approach used here and in the following for the extraction and identification of resonance frequencies from data is described in the Appendix.[1] The estimated coincidence frequency, $\hat{f}_c$, is localized at 8.12 kHz by the peak of the envelope of the spectral midfrequency enhancement, which roughly corresponds to the maximum of the A_{0-} wave group speed curve.[1] By using Eq. (7) the estimated sound speed in water is 1530 m/s, which is in good agreement with the measured value of 1520 m/s.

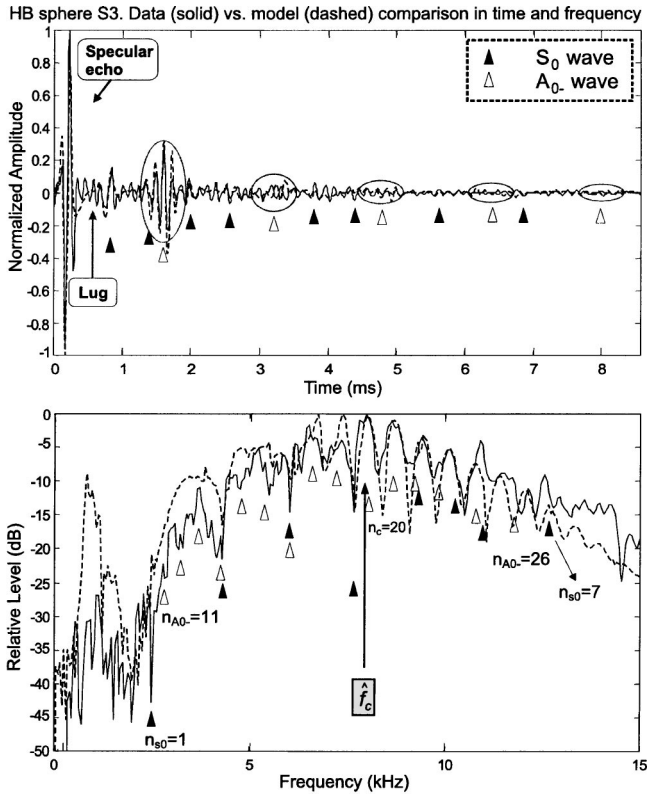


FIG. 9. Half-buried spherical shell. Model-data fit and elastic scattering interpretation.

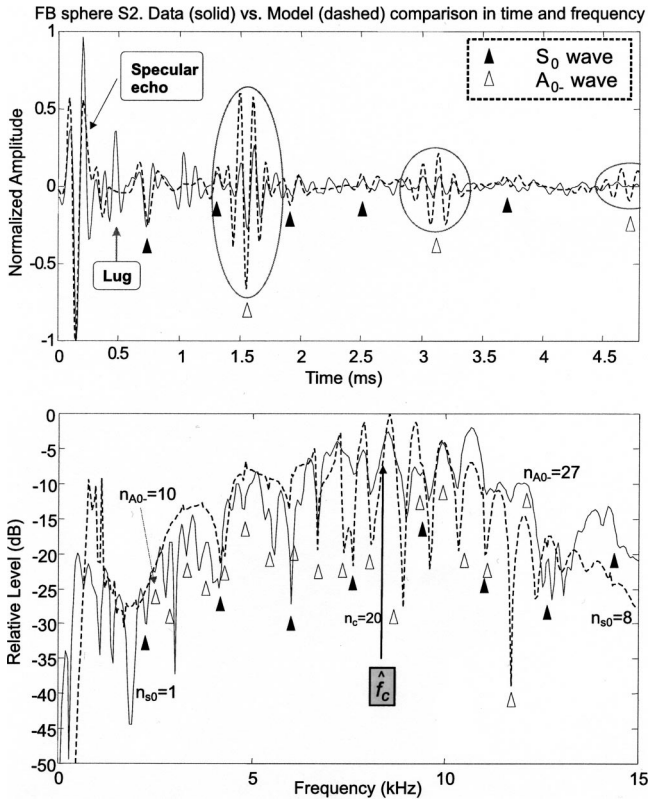


FIG. 10. Flush-buried spherical shell. Model-data fit. Elastic scattering analysis.

B. Half-buried spherical shell (HB)

In Fig. 9 the measured response from the half-buried (HB) spherical shell is compared with the *T*-matrix-based simulation described in Sec. III B. The measured grazing angle is about 26° and the range 22 m. The best fit is obtained for a burial depth of 10.6 cm below the sphere equator and a sand sound speed set to 1647 m/s. This result is in agreement with the theoretical predictions of a reduced speed in the porous media as frequency is reduced, demonstrated by the at-sea data analysis[15] mentioned in Sec. II.

This case is the most complicated to interpret, due to the presence of two outer media (i.e., water and sand). However, both the waves are still identified in time and in frequency and have relative levels comparable with the free-field case. The mismatch in the level of the wave, which was noticed in the free-field data analysis, is much less evident here in the time domain and in the midfrequency enhancement. This is probably the result of a more optimal orientation of the shell. (In tank tests with welded spherical shells, it has been observed that echo amplitudes due to the A_{0-} wave are significantly affected by the orientation of the equatorial weld relative to the incident wave.) However, the mismatch noticed at low frequency (i.e., for $f<2$ kHz) in the free-field data as discussed previously is even more significant here. The reason for the larger discrepancy is unknown.

Above 13 kHz the model-data fit is poorer because of an enhancement of the reverberant noise, which superimposes several unmatched echoes on the predicted target response in the time domain. Among those unpredicted echoes, one, localized in Fig. 9, should correspond to the return from the target's support lug, but its level is unexpectedly low relative to the specular echo when compared to that measured in the free field.

From $\hat{f}_c$ localization ($\hat{f}_c=8.25$ kHz), the estimate of the effective external sound speed obtained from Eq. (10) is 1555 m/s, which is not far from the result (1580 m/s) obtained by applying the formula of $c_{\text{ext}}^{\text{eff}}$ proposed in Ref. 29. This result validates the hypothesis that even when the A_{0-} wave becomes fluid-borne in nature, it continues to revolve around the whole spherical shell and is not limited to travel only in water, reflected back by the bottom boundary. This effect might also be a result of the relatively small impedance difference between water and sand.

C. Flush- and deeply-buried spherical shells (FB/DB)

The response of the flush-buried spherical shell is plotted in Fig. 10, where it is compared with the simulated data and analyzed in terms of elastic wave extraction and identification. The measured grazing angle is 34.5° and the range is 18 m. The best fit model for the flush-buried spherical shell corresponds to a burial depth of 2 cm from the top of the target, and sand sound speed set to 1642 m/s.

As shown in Fig. 10, both the waves outlined in free-field data are evident and their scattering levels relative to the specular echo are comparable with the free-field case. The mismatch in the A_{0-} wave levels is particularly evident in the time domain. As in the previous data, a strong mismatch is evident at low frequency. Above 13 kHz the model-data fit is poor because of a significant decrease of the signal-to-reverberation level, hence the elastic scattering analysis be-

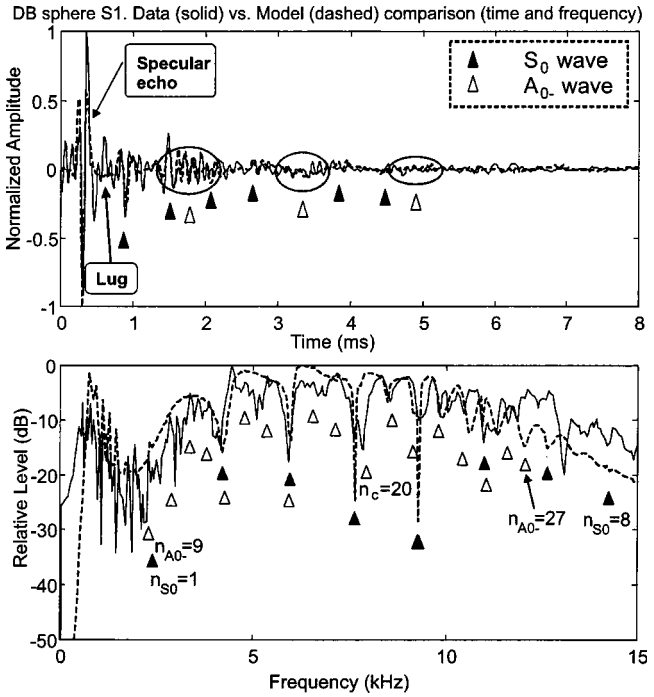


FIG. 11. Deeply-buried spherical shell. Model-data fit. Elastic scattering analysis.

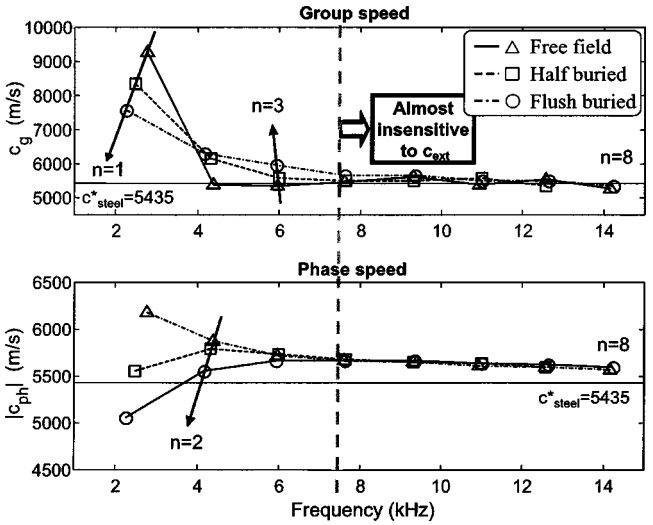


FIG. 12. Comparison of S_0 wave dispersion curves as the burial depth changes, for the free-field, half-buried, and flush-buried cases. The estimated resonance modes are localized on the respective curves by triangles, squares, and circles.

comes unreliable. This effect is more evident than in the half-buried case, in which the signal is maintained at a higher level by the scattering contribution from the water-loaded part of the target. The reverberation, which may be due mainly to volume inhomogeneities, affects the data time series with several unmatched echoes. The strong return following the specular echo is identified as a reflection by the shell's support lug. Its considerably strong level is justified by the relatively short travel path through the sediment needed to account for this echo. From $\hat{f}_c$ localization ($\hat{f}_c = 8.765$ kHz), Eq. (7) provides the outer medium speed estimate of 1652 m/s.

For the measurements of the deeply buried spherical shell, the grazing angle is about 42° and the range is about 16 m. The best fit with the model was found by setting the burial depth to 35 cm from the top of the target and the sand sound speed to 1652 m/s. Figure 11 shows the model-data comparison in time and frequency and the superimposed analysis of elastic waves. As in the flush-buried case, the fairly strong arrival from the top lug can be easily identified on the basis of geometrical considerations. Its level, being comparable to the predicted elastic wave contribution, is justified by the relatively short path length through the sediment needed to account for it.

As expected from theory, the level of the S_0 wave remains essentially unchanged with respect to the specular echo as the burial depth varies. Therefore, the S_0 echoes and resonance modes are easily identified. However, due to the attenuation caused by propagation through the sand, the A_{0-} wave level decreases significantly with respect to the flush-buried case. The low signal-to-interference ratio only allows a subset of these resonance modes to be unambiguously identified in the frequency domain. The significant increase of the reverberation level is already evident at 12 kHz and causes many spurious echoes to appear in the data time series with levels comparable to the predicted elastic response of the target. Under these conditions the midfrequency enhancement cannot be detected and localized with sufficient precision to allow a robust estimation of the coincidence frequency.

Further, a model-data mismatch appears at low frequencies. Here, the discrepancy is less severe than in the half-buried and flush-buried cases, but it is of the same order (5–10 dB) as that measured in the free field. Nevertheless, it is notable that the sand sound speed values that result in good agreement between the scattering models and data for both the flush-buried and deeply buried spherical shells are very close to each other and to that selected in the half-buried case. When compared to the sand sound speed measured in core samples (1720 m/s) by transmitting a 200 kHz pulse, there is a significant difference. As outlined in Sec. IV B, this result confirms those presented in Ref. 15.

D. Analysis of elastic wave dynamics

The behavior of elastic waves *versus* frequency as burial depth changes is predictable using the simple theoretical arguments described in Sec. III B. The elastic analysis is performed in terms of the wave speed dispersion curves estimated from the averaged data presented for each target in Secs. IV A–IV C. For supercritical insonification at ranges beyond the nonlinear volume interaction region of the parametric source, the selection of different grazing angles and ranges for measurements on the three buried targets is not expected to influence the characteristics of the generated elastic waves in terms of resonance frequencies. Hence, the data sets are assumed comparable.

Due to the robustness of the scattering measurements discussed in Sec. IV, the deviation of the dispersion curves of the S_0 and A_{0-} waves estimated from each ping is small compared to the mean values on each data set (for both waves, the standard deviations on the phase speed range

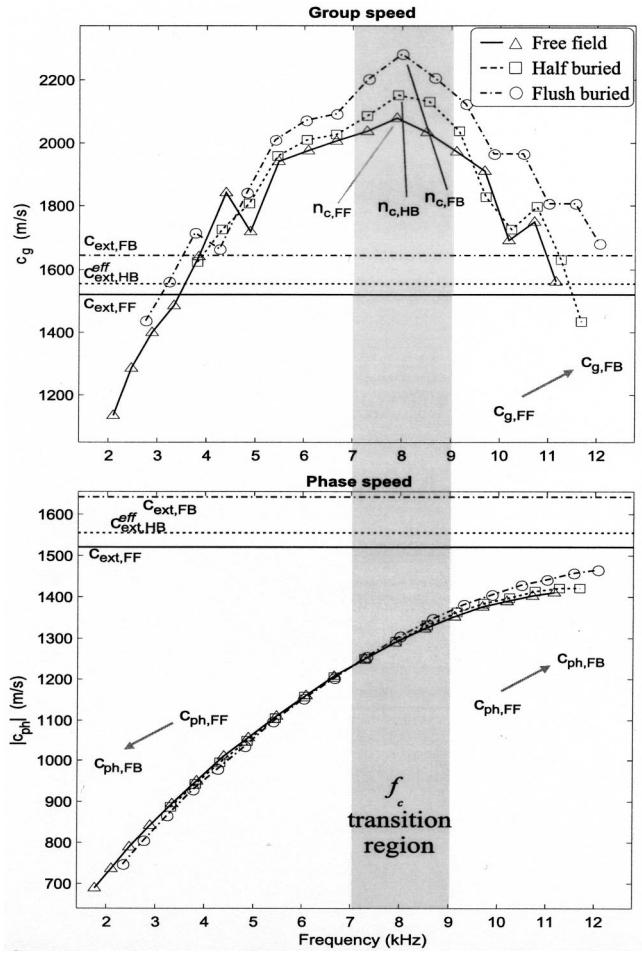


FIG. 13. Comparison of A_{0-} wave dispersion curves among the free-field, half-buried, and flush-buried cases. The estimated resonance modes are localized on the respective curves by triangles, squares, and circles. The gray bar localizes the transition region around the coincidence frequency where the A_{0-} wave is expected to change its nature from a shell-borne to a fluid-borne wave.

from 2 to 6 m/s from target to target, on the group speed from 30 to 90 m/s). As a result, the dispersion curves for individual pings are quite close to those estimated from ping averaged data. Given the data stability in the range 2–15 kHz and the number of pings (50) averaged in each data set, the estimation of the dispersion curves computed from the averaged data sets is assumed to be statistically consistent.

The dispersion curves of the waves estimated in the free-field, half-buried, and flush-buried cases are compared in Figs. 12 and 13. These three cases are compared first in order to analyze the wave dynamics as the percentage of shell surface loaded by the sediment changes from 0 to 100. Due to the very small standard deviations computed on the estimated dispersion curves, the comparison among curves belonging to different burial cases can be considered statistically significant. The smoothness of the experimental curves that will be shown in the following helps confirmation of the robustness of their estimation.

As emphasized by the arrows in Fig. 12, the first three modes (but particularly the first one) of the S_0 wave shift down in frequency with burial depth, from the free field to the totally buried case, which is a consequence of the predicted greater inertial loading caused by sand compared to water. For higher frequencies, dispersion due to the sediment becomes negligible in all cases; hence, all three curves almost coincide and tend to the membrane speed for steel, c^*_{steel}.

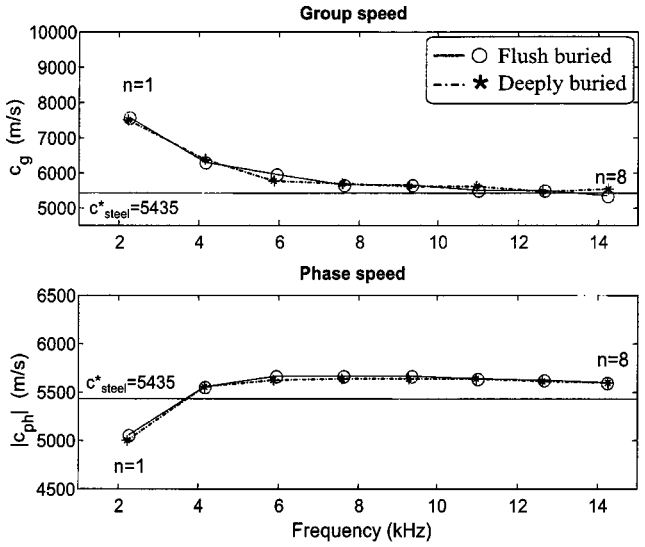


FIG. 14. Comparison of S_0 wave dispersion curves as the burial depth changes between the flush-buried and deeply buried cases. The estimated resonance modes are localized on the respective curves by circles and stars.

In accordance with theory, the phase speed of the A_{0-} wave has opposite trends at frequencies lower and higher than the coincidence frequency, while the group speed increases with burial across the whole bandwidth. At low frequency, where the wave behaves as shell-borne, its modes tend to shift to lower frequencies due to sand inertial loading (as noticed for the first S_0 modes). As the percentage of surface area loaded by sediment increases, these modes exhibit a larger shift down in frequency. Consequently, the wave phase speed tends to decrease. A concurrent increase of the group speed is observed at these frequencies because sediment loading shifts the resonance modes more as the mode order decreases causing the frequency gap between modes to increase with burial.

Around and beyond the coincidence frequency, where the wave behaves as a fluid-borne wave, the wave characteristics are opposite, namely the A_{0-} mode locations shift to higher frequencies and the phase speed tends to increase with burial depth and frequency. Also the group speed increases with burial and frequency because, at higher frequencies, the upward shift of the resonance modes again widens the gap between the modes. This is expected from theory. The shift to higher frequency is related to the increased speed of the effective outer medium in which this wave travels as the burial depth increases. At the same time, the whole midfrequency enhancement region shifts up with burial, together with the coincidence frequency, which is linearly related to the effective external sound speed according to the empirical formula of Eq. (10).

In Figs. 14 and 15 the S_0 and A_{0-} wave dynamics are analyzed as the burial depth changes between two cases of completely buried spherical shells. Between the flush-buried and deeply buried cases the changes in wave dynamics are

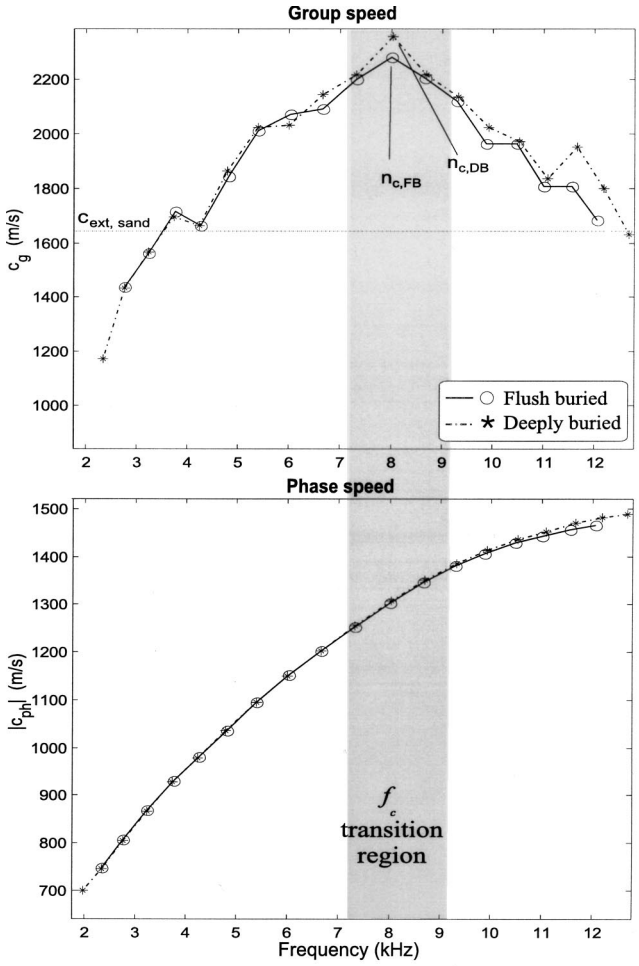


FIG. 15. Comparison of A_{0-} wave dispersion curves between the flush-buried and deeply buried cases. The estimated resonance modes are localized on the respective curves by circles and stars. As in Fig. 13, the gray bar localizes the transition region around the coincidence frequency where the A_{0-} wave is expected to change its nature from a shell-borne to a fluid-borne wave.

slight, as expected,[5] and the respective phase and group speeds of the S_0 and A_{0-} waves almost coincide. In particular, a minute low-frequency shift of the first S_0 and A_{0-} modes of the deeply buried shell is predicted by theory[5] with respect to the flush-buried one, due to an increasing effective inertial mass loading of the shell as its distance from the water-sand boundary increases. However, this effect is detectable only in the first and third modes of the measured dispersion curves of the S_0 wave, and cannot be resolved for the A_{0-} wave. The frequency shifts are expected to be so small that the measurement accuracy becomes inadequate. The A_{0-} wave modes around and beyond the coincidence frequency are almost unchanged as the properties of the loading fluid where the wave is supposed to travel are approximately the same as the depth changes. Only a slight shift toward higher frequencies of the deeply buried A_{0-} highest-order modes is detectable, which is in agreement with the slightly higher sand sound speed chosen to optimize the model-data fit for the deeply buried target with respect to the flush-buried case.

V. CONCLUSIONS

Measurements of elastic wave scattering from partially and completely buried spherical shells have been analyzed using a suite of wave theory target scattering models. The significance of burial depth was investigated by analyzing backscattering data from three identical spherical shells, respectively, half-, flush-, and completely buried within a sandy seabed. Free-field backscattering measurements obtained by suspending one of the spherical shells in the water column were used as a reference for verifying the geometry and the elastic parameters of the spherical shells, and as a benchmark for the target scattering models. With the objective of investigating the potential of using structural resonances for classification of buried targets, special emphasis was placed on investigating the excitation of the fundamental symmetric and antisymmetric Lamb-type waves (S_0 and A_0, respectively).

The benchmark analysis of the backscattering from the free-field spherical shell demonstrated the capability of the scattering models to capture the dispersion characteristics of the A_0 and S_0 elastic components. On the other hand, the models consistently overestimated the scattered levels associated with the midfrequency enhancement of the antisymmetric modes, probably caused by shell imperfections or the presence of a lifting lug welded to the top of the spherical shells. The reflection from this top lug could also be identified in the data time series. A model-data mismatch was also detected at $f<2$ kHz, common to all the data sets but particularly strong in the half-buried and flush-buried cases. The common effect may be due to a mismatch between the transmitted pulse and its model used in simulation. However, the greater discrepancy in the two cases mentioned remains unexplained.

Despite the low-frequency mismatch, the data analysis of the buried spherical shells showed good agreement between the measurements and theory,[5,7,10] especially in terms of elastic wave dynamics. The elastic waves identified in the free-field case could be identified at all burial depths, and their dispersion characteristics were accurately represented by the models. As was the case for the free-field spherical shell, the level of the symmetric S_0 returns were accurately modeled as well, while the A_0 returns were overestimated in amplitude by the models. On the other hand the relative reduction of these arrivals relative to the half-buried spherical shell was accurately predicted. Due to their stronger interaction with the surrounding medium, the antisymmetric modes are more affected by burial into the sediment than the symmetric ones, an effect which is accurately captured by the models.

Another interesting result of the analysis is the independent confirmation of the strong dispersion of acoustic waves in a sandy seabed, recently demonstrated on the basis of acoustic data analysis by Maguer *et al.*[15] Thus, the present analysis suggests a sediment sound speed of about 1640 m/s, whereas the direct core measurements at 200 kHz yielded 1720 m/s. As described in the earlier work, the Biot model was needed to interpret the experimental results.

The analysis of elastic wave scattering from the buried targets used in GOATS'98 is continuing. The analysis is con-

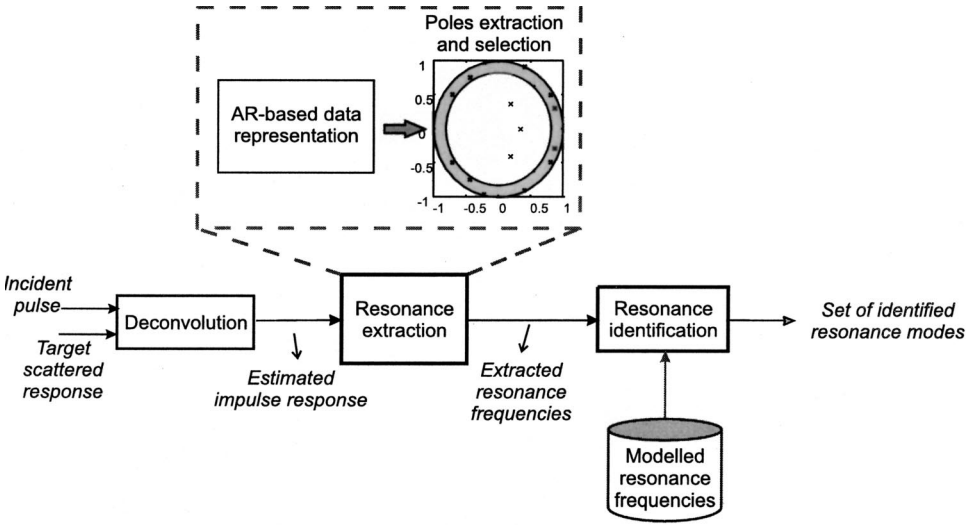


FIG. 16. Block diagram of the resonance extraction and identification procedure.

centrating on the subcritical insonification regime. Effective insonification below the critical grazing angle of the seabed would allow for higher area search rates, and low frequencies penetrate better into ocean sediments in these geometries. The analysis will be extended to the more complicated aspect-dependent scattering properties of buried cylindrical targets and bistatic receiver geometries.

ACKNOWLEDGMENTS

The authors wish to thank all the SACLANTCEN and MIT staff members involved in GOATS'98 experiment for their invaluable contribution to its success. They especially want to thank Per Arne Sletner for establishing and running the TOPAS source facility from the shore laboratory at Marciana Marina, and Marco Paoli for leading the extensive diving effort. Piero Guerrini excelled as Engineering Coordinator. The outstanding effort of the crews of R/V ALLIANCE and T/B MANNING is highly appreciated. The comments provided by Reginald Hollett contributed significantly to the quality of the paper and are highly appreciated. A special thanks to the Management at SACLANTCEN for squeezing GOATS'98 into their program of work at short notice, for their continued support for the entire project. The U.S. Office of Naval Research supported the MIT components of the experiment and the analysis is performed at the Costal Systems Station.

APPENDIX: APPROACH OF RESONANCE FREQUENCY EXTRACTION AND MODEL-BASED IDENTIFICATION

The method used for extracting and identifying the elastic resonance modes from the measured response of a spherical shell is outlined in Fig. 16.

The approach used for the resonance extraction can be applied to any kind of elastic target. The identification method is model-based, with the reference model being the response of an air-filled, thin-walled spherical shell. However it can be extended to any other object, as long as a model is available.

The scatterer is assumed to be well represented by a linear time invariant (LTI) time-discrete system with a rational transfer function (TF) model.[1,18] Unattenuated resonances are defined as those poles of the TF of the LTI system that lie on the unit circle in the z domain. Furthermore, according to resonance scattering theory, the modal resonances have been proven to correspond to the poles of the rational transfer function of the scatterer.[18] The most straightforward method for estimating the poles of a system consists of representing its TF with a parametric autoregressive (AR) model.[31] As a real impulse response is to be modeled, all complex poles (or zeros) included in the TF rational pattern imply the presence of their respective complex conjugates among the poles (or zeros).[1]

AR modeling is applied to the measured data after a preliminary nonparametric deconvolution of the target response which provides the impulse response of the scattering system. In theory, only the poles on the unit circle in the z plane are resonances of nondispersive waves. With the S_0 wave being almost nondispersive in the whole frequency range and the A_{0-} wave having relatively insignificant dispersion below the coincidence frequency, the corresponding modes are expected to lie within a finite annulus defined in the z plane. Each couple of complex conjugated poles and each real-valued pole estimated by the AR model which lies within the annulus is a potential resonance mode of the two waves. Here the $0.90 \leqslant |z| \leqslant 1$ annulus is assumed.

The set of resonance frequencies $\{f_k^{\mathrm{AR}}, k=1,...,K\}$ extracted from data are identified by the following model-based technique. A model pattern of expected resonance modes belonging to the A_{0-} and S_0 waves is numerically computed from the approximated formulas (9.4.9) and (9.4.13) presented in Ref. 32. This method was developed under free-field conditions,[32] but has been adapted here to targets interacting with the seabed. The external sound speed and density required by the original algorithm are substituted with the effective quantities $c_{\mathrm{ext}}^{\mathrm{eff}}$ and $\rho_{\mathrm{ext}}^{\mathrm{eff}}$ defined in Sec. III B (see Ref. 29). This generalization is justified because of the contrast between the elastic properties of the target and the sediment which makes the effect of replacing water by sediment relatively insignificant in regard to the spectral locations of the modes, as confirmed by the results of the analysis.

The model produces a set of resonance frequencies

$$\{f_{l,n_l},\ \ n_l=1,2,...,N_l,l\in L\}, \tag{A1}$$

where l is the wave index ($L=\{S_0,A_{0-}\}$) and n_l is the resonance modal order of the l wave type. The AR-based extracted set of frequencies is identified by a simple algorithm of model-data pattern association.

The association problem is solved by a distance-based approach. The subset of measured resonances $\{f^{\mathrm{AR}}_{l,\hat{n}_l},\hat{n}_l \in\{1,...,\hat{N}_l\}\}$ which best fits the model pattern [Eq. (A1)] is found for each wave type l in terms of minimum Euclidean distance between model-data pairs of frequencies:

$$\{f^{\mathrm{AR}}_{l,\hat{n}_l}\}=\{f^{\mathrm{AR}}_k:\min_{k,n_l}(f^{\mathrm{AR}}_k-f_{l,n_l})^2<\mathrm{thr}\}. \tag{A2}$$

Hence, for the l wave type, if the distance between the measured frequency f^{AR}_{k*} and a certain modeled resonance mode f_{l,n_l^*} is the minimum over all the n_l values, but is higher than a fixed threshold thr, then f^{AR}_{k*} cannot be identified with any of the resonance modes predicted for the l wave.

[1] A. Tesei, W. L. J. Fox, A. Maguer, and A. Løvik, "Target parameter estimation using resonance scattering analysis applied to air-filled, cylindrical shells in water," J. Acoust. Soc. Am. **108**, 2891–2900 (2000).

[2] C. E. Baum, "Discrimination of buried targets via the singularity expansion," Inverse Probl. **13**, 557–570 (1997).

[3] A. Maguer, W. Fox, H. Schmidt, E. Pouliquen, and E. Bovio, "Mechanisms for subcritical penetration into a sandy bottom: Experimental and modeling results," J. Acoust. Soc. Am. **107**, 1215–1225 (2000).

[4] H. Schmidt, and J. Lee, "Physics of 3-d scattering from rippled seabeds and buried targets in shallow water," J. Acoust. Soc. Am. **105**, 1605–1617 (1999).

[5] R. Lim, J. L. Lopes, R. H. Hackman, and D. G. Todoroff, "Scattering by objects buried in underwater sediments: Theory and experiment," J. Acoust. Soc. Am. **93**, 1762–1783 (1993).

[6] J. A. Fawcett, "Summary of object/waveguide scattering modeling at SACLANTCEN (1995–1998)," SR-298, NATO SACLANT Undersea Research Centre, La Spezia, Italy, June 1998.

[7] R. Lim, "Scattering by partially buried shells," in *Proceedings of the ICA/ASA Meeting* (AIP, New York, June 1998), Vol. 1, pp. 501, 502.

[8] H. Schmidt, A. Maguer, E. Bovio, W. Fox, K. LePage, N. Pace, P. Guerrini, P. Sletner, E. Michelozzi, B. Moran, and R. Grieve, "Generic oceanographic array technologies (GOATS)'98-Bi-static seabed scattering measurements using autonomous underwater vehicles," SR-302, NATO SACLANT Undersea Research Centre, La Spezia, Italy, October 1998.

[9] I. A. Viktorov, *Rayleigh and Lamb Waves. Physical Theory and Applications* (Plenum, New York, 1967).

[10] G. S. Sammelmann, D. H. Trivett, and R. H. Hackman, "The acoustic scattering by a submerged, spherical shell. I. The bifurcation of the dispersion curve for the spherical antisymmetric Lamb wave," J. Acoust. Soc. Am. **85**, 114–124 (1989).

[11] M. Talmant, H. Überall, R. D. Miller, M. F. Werby, and J. W. Dickey, "Lamb waves and fluid-borne waves on water-loaded, air-filled thin spherical shells," J. Acoust. Soc. Am. **86**, 278–289 (1989).

[12] O. Bergem, and N. G. Pace, "Calibration of the TOPAS PS040. I. Measurements recorded with TOPAS acquisition system," SM-119, NATO SACLANT Undersea Research Centre, La Spezia, Italy, Jan. 1996.

[13] D. J. Wingham, N. G. Pace, and R. V. Ceen, "An experimental study of the penetration of a water-sediment interface by a parametric beam," J. Acoust. Soc. Am. **79**, 363–374 (1986).

[14] E. Breitenbach, H. Überall, and K.-B. Yoo, "Resonant acoustic scattering from elastic cylindrical shells," J. Acoust. Soc. Am. **74**, 1267–1273 (1983).

[15] A. Maguer, E. Bovio, W. Fox, and H. Schmidt, "*In situ* estimation of sediment sound speed and critical angle," J. Acoust. Soc. Am. **108**, 987–996 (2000).

[16] J. A. Fawcett, W. L. J. Fox, and A. Maguer, "Modeling of scattering by objects on the seabed," J. Acoust. Soc. Am. **104**, 3296–3304 (1998).

[17] E. A. Robinson, and S. Treitel, *Geophysical Signal Analysis* (Prentice–Hall, Englewood Cliffs, NJ, 1980).

[18] N. D. Veksler, *Resonance Acoustic Spectroscopy* (Springer, Berlin, 1993).

[19] K. F. Graff, *Wave Motion in Elastic Solids* (Dover, New York, 1991).

[20] S. G. Kargl, and P. L. Marston, "Observations and modeling of the backscattering of short tone bursts from a spherical shell: Lamb wave echoes, glory, and axial reverberations," J. Acoust. Soc. Am. **85**, 1014–1028 (1989).

[21] G. Maze, F. Leon, J. Ripoche, and H. Überall, "Repulsion phenomena in the phase-velocity dispersion curves of circumferential waves on elastic cylindrical shells," J. Acoust. Soc. Am. **105**, 1695–1701 (1999).

[22] A. D. Pierce, *Acoustics. An Introduction to its Physical Principles and Applications* (ASA, New York, 1989).

[23] F. Magand, and P. Chevret, "Time frequency analysis of energy distribution for circumferential waves on cylindrical elastic shells," Acust. Acta Acust. **82**, 707–716 (1996).

[24] A. Boström, "Scattering of stationary acoustic waves by an elastic obstacle immersed in a fluid," J. Acoust. Soc. Am. **67**, 390–398 (1980).

[25] A. Boström, "Scattering of acoustic waves by a layered elastic obstacle in a fluid-An improved null field approach," J. Acoust. Soc. Am. **76**, 588–593 (1984).

[26] R. Lim, "Acoustic scattering by a partially buried three-dimensional elastic obstacle," J. Acoust. Soc. Am. **104**, 769–782 (1998).

[27] A. Boström, "Multiple scattering of waves by bounded obstacles," J. Acoust. Soc. Am. **67**, 399–413 (1980).

[28] B. Peterson, V. K. Varadan, and V. V. Varadan, "Scattering of elastic waves by a fluid inclusion," J. Acoust. Soc. Am. **73**, 1487–1493 (1983).

[29] If a wave diffracts around the partially buried sphere in a time t at speeds in each medium not too far from the external sound speeds in these media, then we can assume $t\approx(L_w/C_w)+(L_s/C_s)=L/c^{\mathrm{eff}}_{\mathrm{ext}}$. Here, L_w is the path length around the sphere in water, L_s is the path length around the sphere in sediment, $L=L_w+L_s$, C_w is the water sound speed, C_s is the sediment sound speed, and $c^{\mathrm{eff}}_{\mathrm{ext}}$ is the effective speed. Then $c^{\mathrm{eff}}_{\mathrm{ext}}$ could be solved for as $c^{\mathrm{eff}}_{\mathrm{ext}}=L/(L_w/C_w+L_s/C_s)$.

[30] G. Kaduchak, C. S. Kwiatkowski, and P. L. Marston, "Measurement and interpretation of the impulse response for backscattering by a thin spherical shell using a broad-bandwidth source that is nearly acoustically transparent," J. Acoust. Soc. Am. **97**, 2699–2708 (1995).

[31] S. M. Kay, *Modern Spectral Estimation: Theory and Application* (Prentice-Hall, New York, 1988).

[32] J. D. Kaplunov, L. Y. Kossovich, and E. V. Nolde, *Dynamics of Thin Walled Elastic Bodies* (Academic, San Diego, 1998).

[33] P. Marston, and N. Sun, "Resonance and interference scattering near the coincidence frequency of a thin spherical shell: An approximate ray synthesis," J. Acoust. Soc. Am. **92**, 3315–3319 (1992).

Normal incidence reflection loss from a sandy sediment

Nicholas P. Chotiros
Applied Research Laboratories, The University of Texas at Austin, Austin, Texas 78713-8029

Anthony P. Lyons,[a] John Osler,[b] and Nicholas G. Pace
Saclant Undersea Research Centre, La Spezia, Italy

(Received 23 August 2000; revised 4 December 2001; accepted 26 June 2002)

Acoustic reflection loss at normal incidence from a sandy sediment, in the Biodola Gulf on the north side of the island of Elba, Italy, was measured in the frequency band 8–17 kHz, using a self-calibrating method. The water depth was approximately 11 m. The mean and standard deviation of the sand grain diameter were 2.25 (0.21 mm) and 0.6 ϕ, respectively. The reflection loss was measured using an acoustic intensity integral method, which is insensitive to roughness effects within the selected frequency band. The measured value of reflection loss was 11 dB, ± 2 dB. The result is consistent with previous measurements in the published literature. The computed reflection loss for a flat interface between water and a uniform fluid or visco-elastic medium with the same properties is 8 dB, ± 1 dB. The theoretical and experimental values do not significantly overlap, which leads to the conclusion that the visco-elastic model is inappropriate. The Biot model is suggested as a better alternative but more work is needed to ascertain the appropriate parameter values. [DOI: 10.1121/1.1511198]

PACS numbers: 43.30.Ma, 43.30.Ky, 43.30.Gv [DLB]

I. INTRODUCTION

Normal incidence reflection of sound from ocean sediments is relatively easy to detect. It is used in depth-sounders to give water depth measurements in countless ships and boats. Mariners and fishermen have observed that the signature of the bottom echo changes with location, and some have claimed to be able to interpret them in various ways. Yet, there has been little experimental verification of acoustic models with respect to the measured and predicted reflection loss. Some have assumed that the ocean sediment behaves like an elastic medium and proceeded to use reflection measurements to invert for sediment properties,[1] or to use sediment properties to predict reflection loss,[2] with varying degrees of success. It will be shown in this study that the elastic or visco-elastic medium assumption is inadequate for sandy ocean sediments. Since sand is a porous medium rather than a uniform elastic solid, Biot's model[3] of sound propagation in a porous elastic medium is likely to be a better alternative. Using Biot's theory as formulated by Stoll and Kan,[4] it is possible to compute the reflection coefficient given the relevant input parameters. The disadvantage of the Biot model is that it requires a large number of input parameters, some of which are difficult to measure.

The scope of this study is limited to the measurement of the normal incidence reflection loss over sandy sediments and a simple comparison of the results to elastic and visco-elastic media theory. This simple comparison is complicated by the roughness of the sand surface. The theoretically computed reflection loss is applicable to a flat surface of infinite extent but the sand surface at the experiment site is definitely not flat. This problem has been extensively studied.[5] The study included numerical examples computed for a sand surface with a roughness spectrum very similar to the surface at this site. With regard to the possibility of relating the measured reflection loss of a rough surface to that of a flat surface, there are a couple of issues. (1) The Kirchhoff approximation is necessary to make the connection between rough and flat surfaces, and, from the numerical simulations, it was concluded that the Kirchhoff approximation is valid above 10 kHz. (2) The surface roughness randomizes the reflected signal, necessitating a statistical approach. The signal intensity, averaged over several pings, is a possible statistic. The numerical modeling showed a reduction in the peak value of the average reflected intensity, relative to that of a flat surface, due to time spreading of the acoustic energy caused by the roughness scattering, that increased with both height above the sand and signal frequency. Interpolating from the numerical results to a height above bottom appropriate to this experiment, the estimated reduction is less than 2 dB, for frequencies below 20 kHz. Furthermore, the reduction due to time spreading may be overcome by integrating the average signal intensity over a window that contains the significant part of the reflected and scattered energies. Using a wide-beam receiver to capture the energy over all significant scattering angles, this statistic is a very reliable measure of the reflection loss. The numerical simulations also indicated that the volume scattering component, due to inhomogeneities within the sand, would be insignificant compared to that of the surface reflection and scattering. Therefore, it may be safely neglected. Further discussion of sediment surface roughness issues may be found in the Appendix.

In the following sections, the process of collecting the data will be described followed by the data processing procedures and the method used to compute the reflection loss, which is not trivial. Finally, the results are compared to pre-

[a]Present address: Applied Research Laboratory, Pennsylvania State University, Box 30, State College, PA 16801.
[b]Present address: Defense Research Establishment Atlantic, 9 Grove St., P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada.

viously published data and visco-elastic model predictions. Comparison with the poro-elastic model is briefly mentioned, but a proper treatment of that subject is beyond the scope of this study.

II. DATA COLLECTION

Normal incidence reflection loss data were gathered, nominally between 9 and 19 kHz, by the Supreme Allied Commander Atlantic Undersea Research Center (SACLANTCEN) in cooperation with the Applied Research Laboratories, University of Texas at Austin (ARL:UT) over a sandy sediment, in the Biodola Gulf on the north side of the island of Elba, Italy. This site was chosen for its uncluttered sand bed of well to moderately sorted sand. A self-calibrating approach was adopted, in which a cylindrically symmetric hydrophone (ITC Type 8101 S/N 1361232) was positioned between a sound projector (ITC Type 3013 S/N 2092) and the sediment. The source was a hemi-spherical cap design, constructed of a 1-3 composite material, capable of broadband sound projection, uniformly over a wide range of angles about its axis of symmetry. According to the calibration data provided by the manufacturer, the variation in hydrophone sensitivity in the axial plane of the transducer was within $+0.5$ and -1.5 dB. The hydrophone recorded both the direct arrival from the projector and the reflected pulse from the sediment, in addition to several other reflections.

The water depth was 11 m. From previous experiments,[6] the sand bed was estimated to be 1.8 m thick, with a density of 1920 kg/m^3. A *p*-wave speed of 1720 m/s was measured at 200 kHz from diver cores. The mean and standard deviation of the sand grain diameter were 2.25 (0.21 mm) and 0.6 ϕ, respectively. The sound speed in the water was 1530 m/s. The projector was mounted in a hoop that was suspended from a single buoy 3 m below the water surface. The hydrophone was suspended 1.3 m beneath the projector in a second hoop, as shown in Fig. 1. The position of the hydrophone was offset from the center of the hoop to avoid scattered energy from the hoop, which would tend to converge at the center. The assembly was floated and tethered to an anchor some distance from the research ship, the RV Manning, to avoid possible interference by reflections from the hull of the ship. The transducers were cabled to the data collection equipment aboard the ship.

The surface condition, visible in the background of Fig. 1, was calm, and remained so throughout the experiment. The site was in a relatively sheltered bay. There was only a small amount of source motion induced by wave activity, and it was mainly in the vertical direction. Consequently, changes in source angle were very small. Given the wide source beam width, any resulting changes in the bottom insonification would have been a very small fraction of a decibel. This is confirmed by the almost imperceptible variation from one ping to the next of the direct path signal from projector to hydrophone.

The choice of source and hydrophone depths was dictated by a number of considerations. The source had to be sufficiently removed from the surface so that the surface bounce signal would not interfere with the direct arrival at the hydrophone. The projector to hydrophone separation (ph) had to be greater than the Rayleigh distance of the source, which is estimated to be less than a meter at 20 kHz. The bottom reflection (pbh) must be sufficiently separated in time from the surface reflection (psh) as well as the surface-bottom multiples (psbh,pbsh). The expected arrival times of the various signals are shown in Fig. 2.

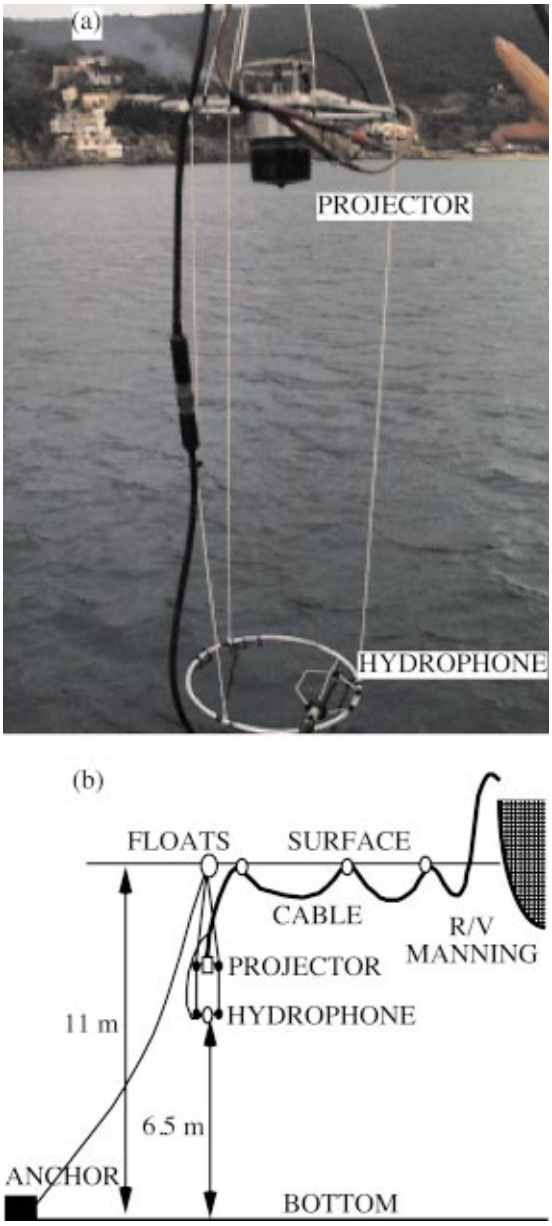


FIG. 1. (a) Photograph of projector and hydrophone assembly and (b) diagram of deployment over the sediment.

The projected signal was a gated sine-wave of exactly three cycles. This was a compromise arrangement, found by experimentation, that provided a workable balance between the need to maximize signal pulse length to obtain a high signal-to-noise ratio after filtering, and the need to limit the pulse length to allow the multiple arrivals to be separable in time. With the given source transducer, it was found that a pulse shorter than three cycles caused excessive ringing and did not carry enough energy. In order to cover the band from 10 to 20 kHz, it was determined that three cycle pulses at 2 kHz intervals would provide adequate coverage. The ampli-

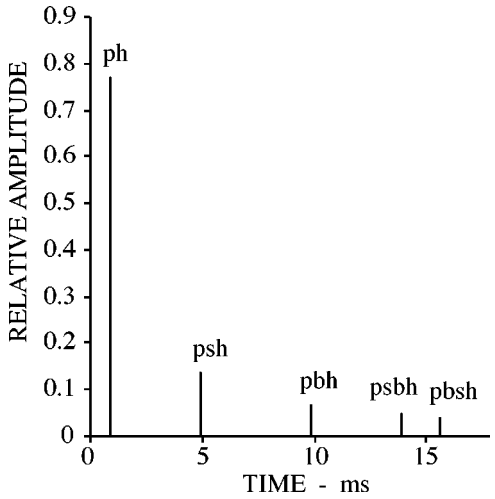


FIG. 2. Expected arrival times of various signal paths.

fier gain had to be adjusted so that the direct path signal did not saturate the system, and yet allow the bottom reflected signal to be clearly measurable above the background noise. An example of a received raw signal is shown in Fig. 3; the plot axes scales were adjusted to show the bottom and other reflected signals and not the direct signal from the projector. The projector to hydrophone (ph) and the bottom bounce (pbh) signals could be separated by time gating. In both cases, a window length of 4 ms was found to be adequate. In addition to the acoustic signals, there were a couple of electrical pulses: the feed-over from the projector power amplifier (e1), and a timing pulse (e2). Both e1 and e2 are electrical cross-talk that normally occur in experiments of this nature. They occur at definite time intervals after pulse transmission and they are completely separable by time gating from the desired acoustic signals. Expanded views of the windowed signals are shown in Figs. 4 and 5. The dynamic range was between + and −10 V with 16-bit resolution. The direct path signal was safely within these limits and, although the bottom reflection was considerably weaker, it was clearly measurable.

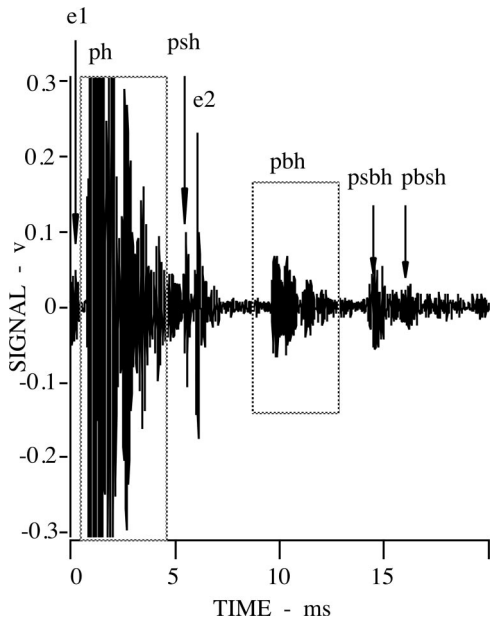


FIG. 3. Example of raw signal measured at the hydrophone.

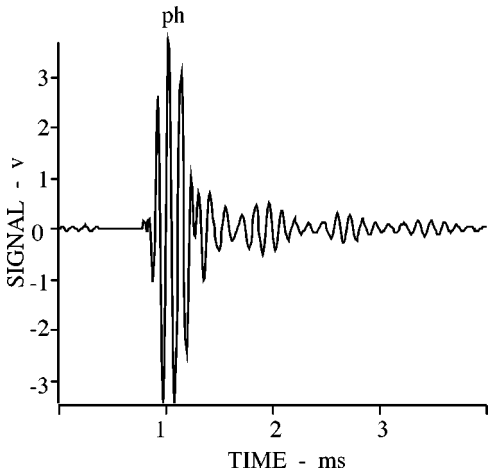


FIG. 4. Expanded view of a direct path signal.

On a ping-to-ping basis, the bottom reflection was a random process, because the projector and hydrophone were constantly in motion about their equilibrium positions under the influence of the waves and currents, and because the bottom was not a perfectly flat surface. The water-sediment interface was dominated by ripples of the order of a centimeter in amplitude, and by mounds caused by biological activity, as shown in the photograph in Fig. 6.

Data were recorded at carrier frequencies 9, 11, 13, 15, 17, and 19 kHz. At each frequency, 1000 pings were recorded, at a rate of 2 pings per second. This was necessary because the reflected signal is a random process, and a large number of pings were needed in order to obtain accurate statistics. An example of the average intensity, in decibels, as a function of time, computed from 250 pings, is shown in Fig. 7. The direct (ph) and bottom reflected (pbh) signals are clearly detectable and separable from the electrical interferences (e1,e2) and the other acoustic arrivals, including the surface reflection (psh) and the surface-bottom multiples (psbh,pbsh).

III. SCREENING

At each center frequency, the 1000 pings were divided into 4 groups of 250 pings. The averaged signal intensity of

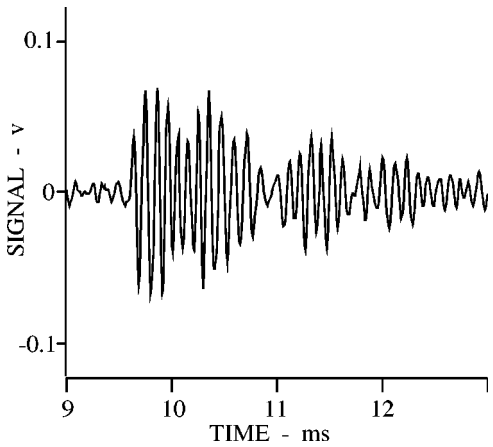


FIG. 5. Expanded view of a raw bottom reflected pulse.

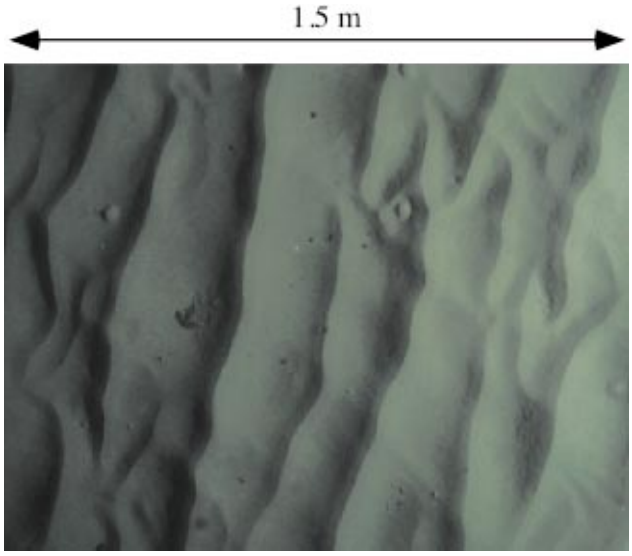


FIG. 6. A photograph of the sediment.

each group was compared to those of the other three. Rayleigh statistics would be appropriate since the reflected signal is an ensemble of many scattered signals. Assuming Rayleigh statistics, the average of 250 pings would have a standard error of 6.3% or 0.27 dB. At each frequency, the averages should agree to within a few standard deviations.

For a truly stationary random process, the differences between the averages should be consistent with the estimated standard deviation at all time samples. In reality, this was not the case. The differences between averages were smallest at the peak of the direct path pulse (ph), at less than 0.01 dB, confirming the stability of the direct path signal. Regarding the direct path signal, small changes in the path length due to relative motion of projector and hydrophone caused changes in arrival time, which tended to increase the differences at time samples along the slopes on either side of the peak. The differences were less than 0.2 dB on the side preceding the peak, but reached as much as 4 dB on the trailing side, suggesting that, in addition to the arrival time perturbation, there

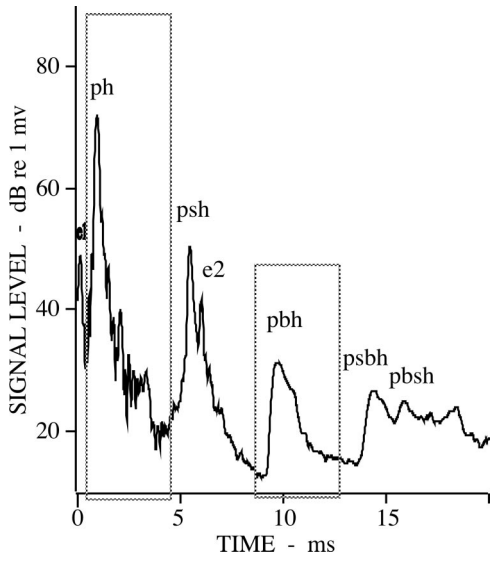


FIG. 7. Average intensity from a group of 250 pings.

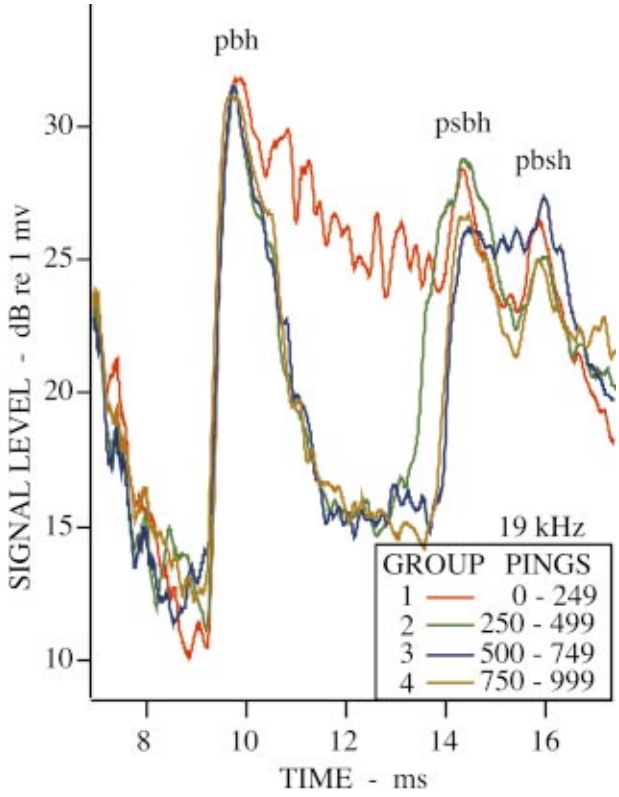


FIG. 8. Average signal intensities from the 4 groups of 250 pings at 19 kHz.

was a significant volume reverberation contribution. Regarding the bottom reflection signal, the difference between averages were greater than the expected standard deviation at all points within the selected time window. The dominant cause is the vertical motion of projector and hydrophone assembly relative to the bottom, which could be several centimeters. The range and statistics of the vertical motion may vary from one group of pings to the next, depending on the surface conditions, causing the average shape of the peaks to differ significantly. Therefore, the averages, at each time sample, differed by a value considerably larger than the expected standard deviation.

Some anomalous deviations in the bottom reflected signals were found. In the 19 kHz data, as shown in Fig. 8, groups 2, 3, and 4 were mutually consistent, but group 1 had an anomalously large decay tail, possibly due to an approaching school of fish or the equipment drifted too close to a patch of seagrass which were known to be in the periphery of the experimental area. Another example is shown in Fig. 9, in which only groups 1 and 2 of the 13 kHz data were mutually consistent. Groups 3 and 4 appear to contain additional scattered energy both before and after the main bottom reflected signal, possibly due to a school of fish passing beneath the transducers. In general, it was found that the averaged signal intensities naturally divided themselves into two categories: (a) Within each set of four, there was often a subset of two or more averaged intensities that were mutually consistent to within a few decibels. They differed by 4 dB or less, consistent with the direct path variations. (b) The remainder would differ from all the others by several decibels, often more than 10 dB. Therefore, the subset of two or

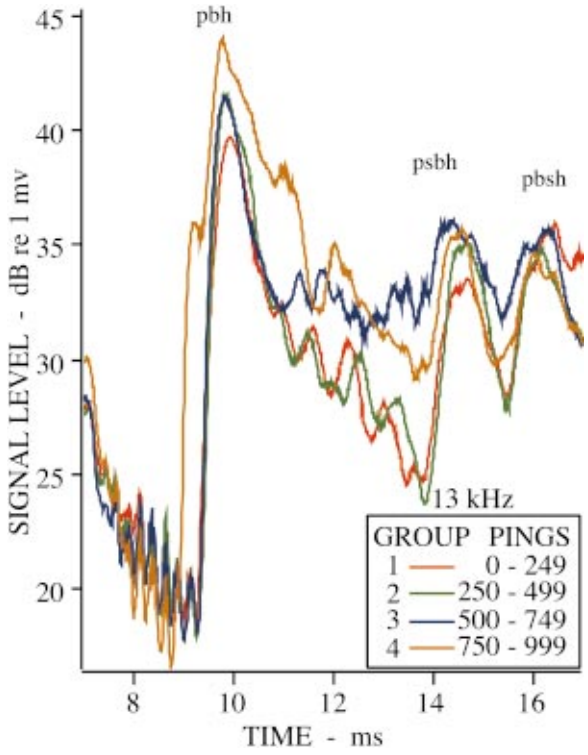


FIG. 9. Average signal intensities from the 4 groups of 250 pings at 13 kHz.

more mutually consistent averaged intensities, if it existed, would be accepted, and the remainder rejected. This screening process was applied to the bottom reflection signal at each frequency. A summary of the screening results are shown in Table I. The data set obtained at a center frequency of 17 kHz was completely rejected because there was very little mutual consistency. The accepted data were used in the subsequent processing to compute the reflection loss.

The critical issue is the acoustic energy integral within the selected time windows. In the case of the direct path signal, the integrated energy differences were less than 0.1 dB, with most of the variations coming from the volume scattering in the decay tail. In the case of the bottom reflection, within the subset that passed the screening process, the differences were consistent with the expected standard deviation of 0.27 dB—the largest measured was 0.6 dB.

IV. MEAN INTENSITY PROFILES

To observe the average intensity profile as a function of frequency, the accepted signals were filtered in 2 kHz bands, centered at frequencies 1 kHz apart, starting at 8.5 and ending at 17.5 kHz. The pass-band windows are shown in Fig. 10. Each pass-band is a cosine tapered window 2 kHz wide at the half-amplitude (−6 dB) point. The 8.5 and 9.5 kHz bands were filtered from the pings that were generated by three cycles of a 9 kHz carrier, because the signal spectrum had continuous coverage from 6 to 11 kHz. The 10.5 and 11.5 kHz bands were filtered from the 11 kHz signal for similar reasons. The 15.5, 16.5, and 17.5 kHz bands were filtered from the 19 kHz signal, because its spectrum had continuous coverage from 13.5 to 19.5 kHz, and also because all the 17 kHz data had been rejected. Higher frequency bands were not available because there was a low pass filter in the receiving circuitry that eliminated all significant spectral components above 19.5 kHz.

TABLE I. Reflection data, in groups of 250 pings, identified by center frequency and start time; *x*=rejected, =accepted.

Center frequency kHz	Start time UTC	Block 1	2	3	4
9	13:53:31		*x*	*x*	
11	13:44:05				*x*
13	14:03:00			*x*	*x*
15	14:12:00	*x*			
17	14:21:00	*x*	*x*	*x*	*x*
19	14:30:00	*x*			

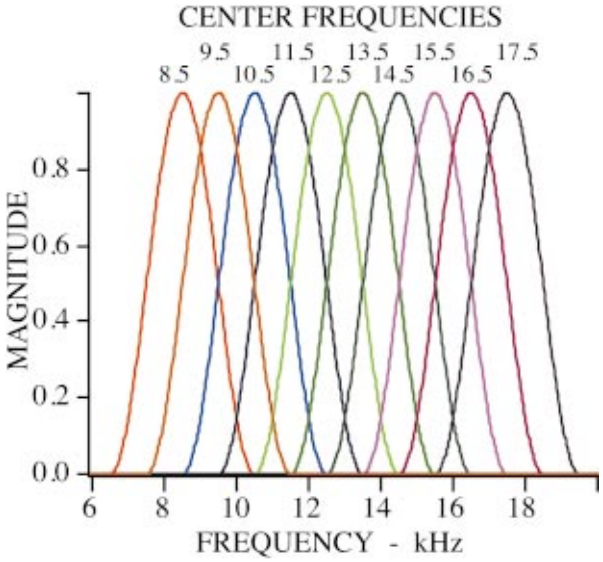


FIG. 10. Filter pass-bands.

The filtered and averaged signal intensity profiles were normalized, by setting the peak of the direct path component to 0 dB. The results are compared in Fig. 11. The main lobe of the direct path signal (ph) appears to be independent of frequency, which is expected, since the bandwidths were identical. The differences appear to be in the subsequent decay tail. The 8.5 and 9.5 kHz signals have the largest decay tails, probably because the projector has a tendency to ring at these frequencies. The electrical feedover interference is also visible (e1).

The surface reflected signal (psh) is the next strongest feature. This is the signal that is launched from the back of the projector, bounced from the water–air surface and received by the hydrophone. The levels appear to be increasing with frequency, which probably reflects the frequency dependence of the front-to-back ratio of the projector. The shape of its intensity profile is distorted due to the presence of an unwanted electrical interference pulse (e2).

The third feature is the bottom reflection signal (pbh). An expanded view of the bottom reflection profile is shown in Fig. 12. If the intensity level at the 9 ms mark may be used as an upper bound estimate of the background noise, then it appears that the bottom reflection peak is approximately 20 dB above the noise floor, except at 8.5 kHz where the ringing of the transducer has significantly raised the noise floor.

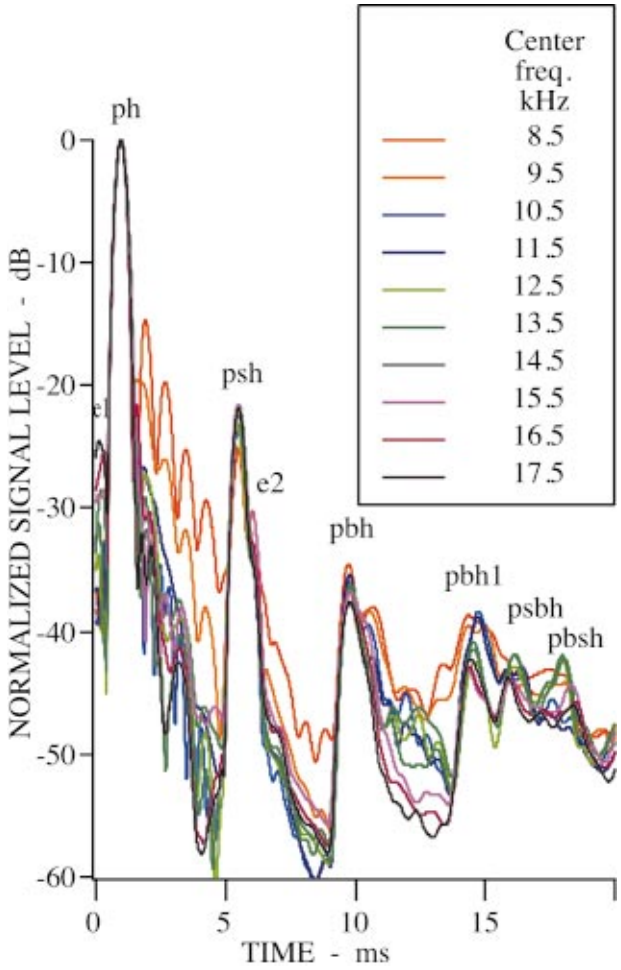


FIG. 11. Average intensity profiles in overlapping bands.

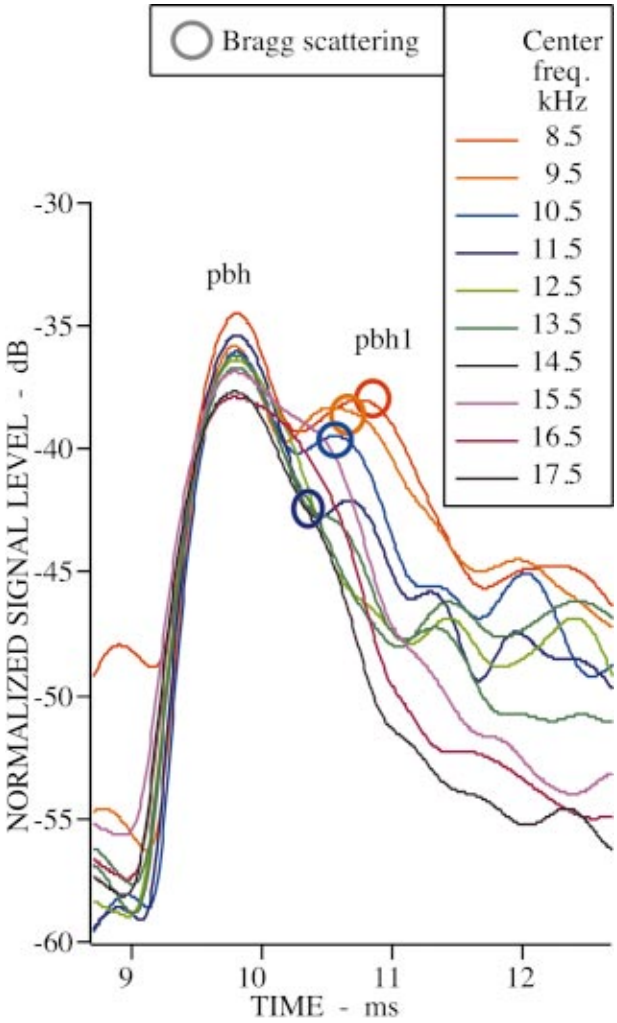


FIG. 12. Expanded view of the bottom reflected signal intensity profile.

The signals at 8.5, 9.5, 10.5, and 11.5 kHz show a second peak, labeled ''pbh1.'' It could be due to number of causes, but the two most likely are either (1) one or more reflections from a sub-bottom layer, or (2) scattering from surface roughness. If it were the former, the time delay would suggest that it came from a depth of approximately 0.7 m. Given that the attenuation of sound in sand is typically 0.5 dB per wavelength, the two-way attenuation loss at 8.5 kHz would be 4 dB. The sound level at the pbh1 peak is only 7 dB below that of pbh, therefore the reflection loss at the sub-bottom layer would have to be only 3 dB. This reflection loss is too low for an unconsolidated sediment layer. It would have had to be solid rock or a gas layer. At 17.5 kHz the two-way attenuation would be 8 dB, which is about 4 dB more than that at 8.5 kHz, therefore it should be detectable at all frequencies within the band, but in fact it is only detectable at the four lowest frequencies. For these reasons, the sub-bottom layer hypothesis was rejected.

If the peak at pbh1 were due to surface scattering, it should have some connection with the surface roughness spectrum of the sand, particularly the ripple structure. From the photograph in Fig. 6, it appears that the sand ripples are roughly 0.2 m apart. Given the height above bottom of the projector and receiver, it is possible to estimate the conditions for Bragg scattering. Constructive interference would occur when the scattered signal paths differ by an integral number (n) of wavelengths. At 8 kHz, this would occur at a range (R_n) of 3.46 m from the point directly beneath the transducers, and the corresponding time delay (T_n) would be 1.08 ms relative to the normal incidence bottom reflection path, as illustrated in Fig. 13. The corresponding values at the other signal frequencies are also shown. The time delay values (T_n) are marked on the respective signals in Fig. 12. It is predicted that T_n would decrease with increasing frequency, and this appears to be consistent with the data at the three lowest frequencies, 8.5, 9.5, and 10.5 kHz. At higher frequencies, T_n is expected to be so small that the Bragg peak would not be separable from the main arrival. Therefore, it is likely that the second and subsequent peaks are due to surface roughness scattering effects.

V. REFLECTION LOSS COMPUTATION

The concept of a reflection coefficient is a simple one. The reflection loss is the difference, in decibels, between the intensities of the incident and reflected waves. The magnitude of the reflection coefficient, in decibels, omitting the negative sign, is equal to the reflection loss.

In practice, the measurement is more complicated because the transducers were placed at some distance from the interface, the wave fronts were not planar, and the water-

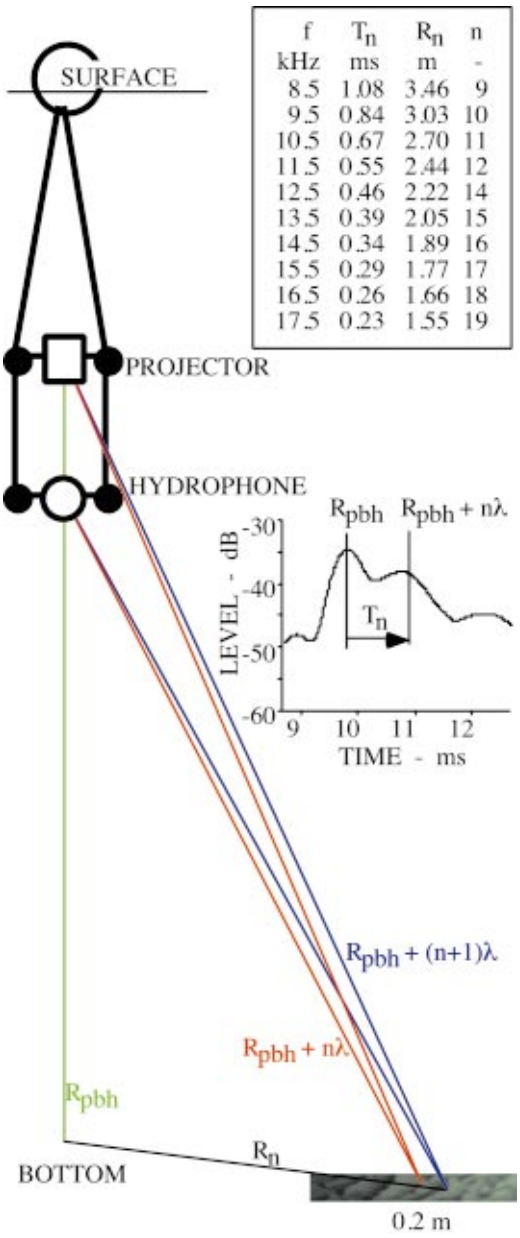


FIG. 13. Illustration of Bragg scattering by sediment surface ripples.

sediment interface was not perfectly flat. The following steps were taken to alleviate these practical limitations. (1) Spreading and absorption loss corrections were used to compensate for the direct and bottom reflected paths between the projector and hydrophone. (2) The wave front curvature problem was reduced by placing the projector as far as practicable from the interface. For normal incidence, the water depth was a practical limitation. It was necessary to assume that the reflection coefficient changes insignificantly over the range of significant angles, i.e., the range of angles subtended by the significant Fresnel zones at the interface. Given the frequency band and height above bottom, the range of angles works out to be approximately within $\pm 10°$ up to the third Fresnel zone. (3) Finally, there is a randomness in the reflected signal, which may be due in part to the surface roughness, but also to spatial variability of the reflection coefficient itself. This randomness precludes any meaningful measurement of the complex reflection coefficient. However, by averaging the intensity of the incident and reflected signals, it is possible to obtain repeatable statistics. Using pulses, rather than continuous signals, the different signals were separable according to travel time. With reference to Fig. 11, it is evident that the pulse duration of the incident intensity (ph) is shorter than that of the averaged reflected signal (pbh). The time spreading of the reflected signal may be due to a number of factors, including scattering by sediment surface roughness and from within the sediment volume, and random changes in height above bottom of the transducers during the course of the experiment.

Given the above conditions, there are a number of ways that the reflection loss may be computed. Two were considered: (1) The reflection loss may be computed from the intensity difference measured at the peaks. In this case, the loss may be overestimated due to the time spreading of the reflected energy, which tends to suppress the peak of the averaged reflected intensity. (2) The reflection loss may be computed from the difference between the time integrals of the intensity pulse, i.e., the energy densities. This result would recapture most of the reduction in the peak value due to the time spreading. The resulting value may slightly underestimate the true reflection loss because the energy in the tail of the reflected signal pulse would likely include additional energy scattered from inhomogeneities within the volume beneath the sediment interface. On balance, method (2) is expected to be more accurate because the effect of time spreading is quite large, but the additional energy from volume scattering beneath the interface is expected to be relatively insignificant. This problem is further discussed in the Appendix.

Using the intensity profiles in Fig. 11, both methods were applied. The integration windows chosen were 0.4–4.4 ms for the direct signal and 9–13 ms for the bottom reflection. The window for the direct signal included the main peak and several range lobes of the direct path signal. The window for the bottom reflected signal included the main peak and the decay tail with all of its minor lobes, but excluding the multipath components psbh and pbsh. The decay tail was included because it is an integral part of the energy reflected from the bottom. Due to bottom roughness, some of the energy that would have been received, had the bottom been perfectly flat, was scattered out of the receiver path and lost, resulting in a reduction in the main peak. This is expected to be largely compensated by energy from surrounding areas scattered into the receiver path, which shows up in the decay tail.[5] The spreading loss differences were estimated from the arrival times at the peaks. The results are shown in Fig. 14. With method (1), the reflection loss values were between 13 and 15 dB. Method (2) gave values between 9 and 13 dB. The latter values are considered to be the best estimates for the reasons given above. These values are consistent with the numerical simulations in Ref. 5 which indicated that method (1) would over estimate the reflection loss by about 2 dB. The actual difference is closer to 3 dB, most likely due to the vertical motion of the transducers, which was not included in the simulations.

The total measurement error E, in decibels, is composed of the following components, added logarithmically:

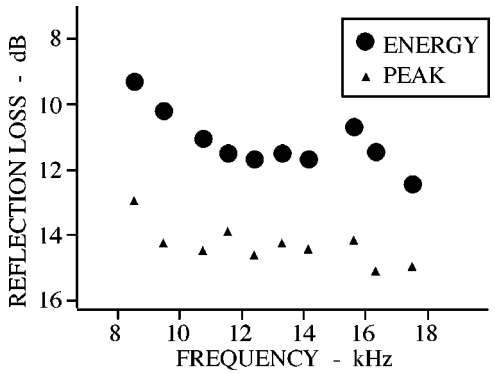


FIG. 14. Reflection loss as a function of frequency.

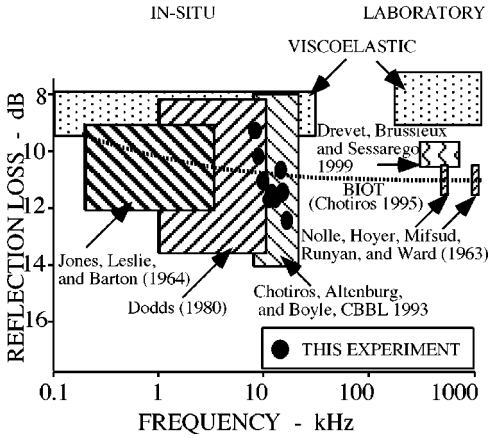


FIG. 15. Comparison with previous published measurements.

$$E = H + N + A, \tag{1}$$

where H is the difference in the hydrophone receiving sensitivity between the upward and downward directions, N is the background noise contribution, and A the contribution due to errors in transmission loss estimation.

As stated earlier, the variation in hydrophone sensitivity in the axial plane of the transducer was within $+0.5$ and -1.5 dB. Accordingly, in subtracting the upward traveling reflected energy density, in decibels, from the that of the downward traveling direct path, the error H due to hydrophone sensitivity variation cannot exceed 2 dB. From Fig. 11, the worse case noise floor, which is mainly due to transducer ringing, occurs at 8.5 kHz and it is approximately 13 dB below the average signal peak, both at ph and pbh. As a result, the peaks at ph and pbh are both expected to be overestimated by 0.2 dB or less. Accordingly, the noise contribution, N, is estimated to be less than 0.2 dB. Since the spherical spreading loss can be accurately computed from the pulse travel times, the contribution from transmission loss estimation is mainly due to the error in estimating the absorption loss. Allowing that the water temperature and salinity at the site were between 10 °C and 15 °C, and 30 and 35 ppt, the error, A, due to the uncertainty in absorption loss, according to the Francois and Garrison formula, is estimated to be less than 0.023 dB. Thus the total measurement error, $H+N+A$, is expected to be less than 2.223 dB.

VI. CONCLUSION

A self-calibrating method was used to measure the acoustic reflection loss at a water–sand interface, in which the same hydrophone is used to measure the incident and reflected signals. The results show reflection loss values between 9 and 13 dB. These values are consistent with previously published values, both *in situ* and in the laboratory, as shown in Fig. 15.

It is instructive to examine the laboratory results before making concluding remarks about the *in situ* measurements. The laboratory measurements by Nolle *et al.*[7] were made at 500 kHz and 1 MHz, using graded, cleaned, degassed, and packed sands: Four samples, with mean grain diameters, 0.64, 0.4, 0.17, and 0.12 mm, were used. The reflection loss was found to be 11 dB, with an error of ± 0.5 dB, independent of grain size and frequency. The porosity and density of the water-saturated sand were 36% and 2069 kg/m^3, respectively, and the sound speed ratio between sand and water was 1.16, independent of grain size. The density of the water was 1000 kg/m^3. The measured sound attenuation in the sand, a, in dB/m, was found to increase as the square root of frequency and inversely with grain diameter,

$$a = 0.016 \quad f/d \quad \text{dB/m}, \tag{2}$$

where f is frequency in Hertz and d is grain diameter in millimeters. The normal incidence reflection coefficient at the interface between two elastic media is simply given by the impedance ratio,

$$R_{\text{elastic}} = (z_s - z_w)/(z_s + z_w). \tag{3}$$

In the lossless case, the impedances of the sand and water, z_s and z_w, are simply equal to the product of the density and the sound speed. Given the above values, the elastic medium reflection coefficient was computed to be 0.41, which represents a reflection loss of 7.7 dB. Using a visco-elastic medium model, a similar calculation that takes account of the attenuation in the sediment gave the same value within a tenth of a decibel. Within the known range of values of the sound attenuation coefficient in sand, which is 1 decibel per wavelength or less at all relevant frequencies, it makes no difference if one models the visco-elastic medium as a causal model with one or more relaxation times or as a noncausal model with a constant complex bulk modulus. The computed normal incidence reflection loss will be within a tenth of a decibel of the lossless case. If a generous allowance of 5% is made for the accuracy of the measured density and sound speed, then the predicted values range from 7.2 to 9.1 dB. Clearly there is a discrepancy between the measured and theoretical values. This result is also consistent with a similar laboratory measurement on sand with a grain diameter of 0.25 mm recently performed by Drevet *et al.*[8] in the frequency band 300–700 kHz.

The measurements by Dodds[9] and Jones *et al.*[10] were made *in situ*, but little environmental information was provided other than that the sediment was sand. *In situ* measurements by Chotiros *et al.*[11,12] over a sandy sediment, near Panama City, Florida, in 1993, under the Coastal Benthic Boundary Layer (CBBL) program, were supported by extensive environmental measurements.[13] In that experiment, the measured values of sediment density and sound speed ranged

from 1907 to 2096 kg/m^3, and 1650 to 1750 m/s, respectively. The sound speed was measured in core samples at 400 kHz. The corresponding parameter values for the water overlying the sediment were 1020 kg/m^3 and 1530 m/s, respectively. Using these values, a reflection loss between 7.9 and 9.4 dB was computed. The theoretical and measured values are compared in Fig. 15. Although there is some overlap between the predicted and measured range of values of bottom reflection loss, the discrepancy cannot be adequately explained by experimental error.

Sound speeds and densities associated with the data reported here are similar to those of Chotiros *et al.* off Panama City, Florida, with one exception. The acoustic measurements by Maguer *et al.*[6] in Biodola Gulf suggested that there was significant dispersion, specifically that the sound speed measured in cores at 200 kHz may be significantly higher than the speed at lower frequencies where the reflection measurements were made. They deduced from propagation measurements that the sediment sound speed in the frequency range 2–15 kHz was 1685 m/s instead of the 1720 m/s measured in cores at 200 kHz. These values still fall within same the range of sound speed values used by Chotiros *et al.*[11,12] and therefore the same range of computed visco-elastic reflection loss values should be applicable. The measured values reported here are consistent with the values from Chotiros *et al.*, but with a smaller spread about the mean value. The data points are plotted as ovals in which the vertical extent represents the expected one standard deviation excursion of ± 0.27 dB and the horizontal extent covers the bandwidth. All but one data point fell outside the range of reflection loss values predicted by the visco-elastic model. The remaining variations in the measured values are likely due to variations in the front-to-back ratio of the hydrophone sensitivity as a function of frequency, and are unlikely to contain any useful information about the frequency dependence of the sediment properties.

In all known cases, the measured reflection loss at normal incidence is inconsistent with the elastic and visco-elastic media models above and beyond expected experimental errors. The measurements come from six independent experiments, two in the laboratory and four *in situ*, in the band from 1 kHz to 1 MHz. The experiment reported herein is particularly reliable due to the self-calibration approach adopted. Since sand is a porous medium rather than a uniform elastic solid, Biot's model[3,4] of sound propagation in a porous elastic medium is likely to be a better alternative. One notable difference between the visco- and poro-elastic models is that the latter specifies two compressional waves. The second wave, though rarely observed, readily absorbs energy, and could account for the enhanced reflection loss. Using parameters for sand put forward by Chotiros,[14] agreement with the measured values can be obtained, as shown in Fig. 15. This computation can be accomplished with readily available computer programs from various sources, including BOGGART3 from N. P. Chotiros, at ARL:UT, and OASES 2.1 from H. Schmidt, at MIT. However, more recent publications[15,16] suggest that aspects of this model, specifically the procedure for estimating certain input parameter values used by Chotiros[14] and the constitutive equations in Stoll and Kan,[4] are not entirely satisfactory. This agreement between model and experiment may be purely coincidental, and more research is needed to establish a physically sound model for sandy sediments.

In conclusion, the Biot model is put forward as a candidate solution for two reasons. (1) It has a loss mechanism, i.e., the slow wave, that can account for a reflection loss in excess of what the visco-elastic model can accommodate. (2) There is a tentative set of parameter values[14] that is consistent with the measured values of reflection loss, wave speeds, and attenuations, and sediment density. The Biot model represents a step forward in our understanding of the underlying physical processes in the acoustics of water-saturated sand, but specifics of the model require further research.

ACKNOWLEDGMENTS

The experiment was conducted by the SACLANTCEN as part of the Acoustic Penetration Experiment (APEx). Thanks are due to the commander and crew of the Research Vessel MANNING, and the engineering staff of the SACLANTCEN who provided excellent support. The authors would like to express their appreciation to the three anonymous reviewers who have significantly contributed to the quality of the final revision. The analysis of the reflection data was funded by the Office of Naval Research Code 321 OA, under the Broadband High-Frequency Sound Interaction with the Sea Floor Program, under the management of Dr. Jeff Simmen.

APPENDIX: REFLECTION LOSS AND SURFACE ROUGHNESS

Sediment surface roughness is usually presented as a power spectral density function. An assumption of stationarity is implicit in this approach. The power spectral density function, $S(f)$, is usually defined as the Fourier transform of the cross-correlation function of the sediment height, $B(r)$.

$$S(f) = \int_{-\infty}^{\infty} B(r) e^{i2\pi fr}\, dr. \tag{A1}$$

In the isotropic case, r is a horizontal distance, and f is a spatial frequency with units of cycles per unit distance. In the anisotropic case, the same approach may be used but it is necessary to specify the direction; r and f would be distances and spatial frequencies in a particular direction. This was the approach used by Briggs[17] when he published a collection of power spectra from ten sites around the world, measured by stereo photogrammetry. All but three were sandy sediments. The power spectral densities from Briggs are shown as the gray background in Fig. 16. It is evident that the curves have similar trends. They decrease with frequency at a rate of approximately 20 dB per decade, and their values differ by no more than 20 dB at any frequency. For the experiment reported here, stereo photographs were taken, but the turbidity of the water was too high to permit sediment roughness measurement. For comparison, the power spectral density from another site[18] in Biodola Gulf are shown as the solid curves in Fig. 16. Initially a two-dimensional spectral density was computed. Two orthogonal cuts were taken, converted to

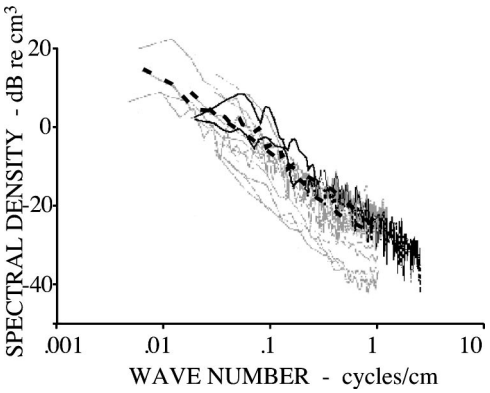


FIG. 16. Comparison of sediment roughness spectral density measurements: (a) gray background: measurements from Briggs (Ref. 17), (b) solid curves: measurements from a site in Biodola Gulf, and (c) broken curves: spectra used in numerical studies in Ref. 5.

one-dimensional spectra, and scaled for direct comparison with those from Briggs. Differences in the low frequency cutoff, i.e., in the region of 0.01 cycles/cm, are due to differences in equipment. Briggs used a mosaic of several photographs to extend spatial coverage and obtain spectral densities extending into the low frequency region. Lyons *et al.*[18] did not have that facility. The sand ripples that these spectra represent were rather bigger than those at site of the experiment reported here, but the general trends should be similar. The ripples have a definite orientation, and the two curves represent the spectra along and across the ripples, respectively. The spectrum across the ripples is the one that has a large peak at 0.05 cycles/cm, which corresponds to an average wave length of 0.2 m. It is evident that these spectra fall within the same range of spectral values measured by Briggs.

One purpose of measuring the reflection loss is for model verification and testing. A single measurement is not sufficient, but when several different measurements are brought together, it will be possible, in a future study, to test the validity of certain models. The models generally predict the reflection coefficient of a flat surface. Since the sand surface is rough, a measurement that is insensitive to the roughness would be very desirable. This problem was studied in some detail in Ref. 5. It included numerical studies, using two roughness power density spectra from the Briggs collection. The two spectra are shown as broken curves in Fig. 16. They are representative of the majority of the spectra in the Briggs collection. The two spectra from Biodola Gulf, running along and across the ripples, represent upper and lower extremes in orthogonal directions. Since the reflection measurement geometry is cylindrically symmetric, the effective one-dimensional roughness spectrum that determines the statistics of the acoustic signal would be an average that lies somewhere between the two extremes. The spectra used in the numerical study in Ref. 5 lie in this vicinity. Therefore, the results of that study should be applicable.

Very briefly, the essential points from Ref. 5 are as follows: (1) In order to make any connection between the reflection losses from rough and flat surfaces, the Kirchhoff approximation has to be applicable. For the roughness spectra used in the numerical study, the Kirchhoff approximation is valid at frequencies in the region of 10 kHz and higher. (2) The roughness randomizes the reflected signal. If one coherently averages the reflected signal, the result would be much smaller than that of a flat surface. (3) If one incoherently averages the reflected signal the reduction in signal level is ameliorated. For frequencies below 20 kHz and a height above bottom of 10 m or less, the numerical study predicted that the reduction would be in the region of 2 dB. The short fall is due to time spreading of the reflected signal energy, which is another effect of rough surface scattering. (4) By using the time integral of the average incoherent signal intensity, instead of its peak level, the shortfall due to time spreading may be removed. The remaining residual errors are due to (a) scattering from the within the sediment, and (b) time spreading outside the integration window. (a) From Ref. 5, scattering from within the sediment is expected to change the result by less than 0.01 dB. (b) From the intensity curves in Fig. 12, the intensity levels at the edge of the integration window are at least 10 dB below the peak level within the window. Assuming a constant decay rate, the failure to collect the energy outside the integration window would cause an error of less than 0.4 dB.

In summary, for a typical sediment roughness spectrum and a height above bottom of 10 m or less, the optimum frequency band is from 10 to 20 kHz. For the signals obtained in this experiment and the integration window chosen, the expected error due to roughness scattering is less than 0.4 dB.

[1] G. J. Heard, ''Bottom reflection coefficient measurement and geoacoustic inversion at the continental margin near Vancouver Island with the aid of spiking filters,'' J. Acoust. Soc. Am. **101**, 1953–1960 (1997).
[2] E. L. Hamilton, ''Reflection coefficients and bottom losses at normal incidence computed from pacific sediment properties,'' J. Acoust. Soc. Am. **35**, 995–1004 (1970).
[3] M. A. Biot and D. G. Willis, ''The Elastic coefficients of the theory of consolidation,'' J. Appl. Mech. **24**, 594–601 (1957).
[4] R. D. Stoll and T. K. Kan, ''Reflection of acoustic waves at a water-sediment interface,'' J. Acoust. Soc. Am. **70**, 149–156 (1981).
[5] N. P. Chotiros, ''Reflection and reverberaton in normal incidence echo-sounding,'' J. Acoust. Soc. Am. **96**, 2921–2929 (1994).
[6] A. Maguer, W. L. J. Fox, H. Schmidt, E. Pouliquen, and E. Bovio, ''Mechanisms for subcritical penetration into a sandy bottom: Experimental and modeling results,'' J. Acoust. Soc. Am. **107**, 1215–1225 (2000).
[7] A. W. Nolle, W. A. Hoyer, J. F. Mifsud, W. R. Runyan, and M. B. Ward, ''Acoustical properties of water-filled sands,'' J. Acoust. Soc. Am. **35**, 1394–1408 (1963).
[8] C. Drevet, M. Brussieux, and J. P. Sessarego, ''High frequency acoustic wave reflection on the surf zone seafloor,'' Acustica **5**, 701–706 (1999).
[9] D. J. Dodds, ''Attenuation estimates from high resolution subbottom profiler echoes,'' in *Bottom-Interacting Ocean Acoustics*, edited by W. A. Kuperman and F. B. Jensen, NATO Conference Series IV; Marine Sciences (Plenum, New York, 1980), pp. 525–540.
[10] J. L. Jones, C. B. Leslie, and L. E. Barton, ''Acoustic characteristics of underwater bottoms,'' J. Acoust. Soc. Am. **36**, 154–157 (1964).
[11] N. P. Chotiros, R. A. Altenburg, and F. A. Boyle, ''Reflection coefficient of sandy ocean sediments in shallow water,'' EOS, **75**, 181 (1994).
[12] N. P. Chotiros, ''Inversion and sandy ocean sediments,'' in *Full Field Inversion Methods in Ocean and Seismic Acoustics*, edited by O. Diachok, A. Caiti, P. Gerstoft, and H. Schmidt (Kluwer Academic, New York, 1995).
[13] M. D. Richardson and S. R. Griffin, ''*In situ* sediment geoacoustic properties. Coastal Benthic Boundary Layer Special Research Program: First Year,'' Report No. NRL/MR/7431-94-7099, Michael D. Richardson, Ed., pp. 146–148, Naval Research Laboratory, Stennis Space Center, MS 39529-5004 (1994).

[14] N. P. Chotiros, ''Biot model of sound propagation in water-saturated sand,'' J. Acoust. Soc. Am. **97**, 199–214 (1995).
[15] R. D. Stoll, ''Comments on 'Biot model of sound propagation in water-saturated sand [J. Acoust. Soc. Am. **97**, 199–214 (1995)],' '' J. Acoust. Soc. Am. **103**, 2723–2725 (1998).
[16] K. B. Briggs, M. D. Richardson, K. L. Williams, and E. I. Thorsos, ''Measurement of grain bulk modulus using sound speed measurements through liquid/grain suspensions,'' J. Acoust. Soc. Am. **104**, 1788 (1999).
[17] K. B. Briggs, ''Microtopographical roughness of shallow-water continental shelves,'' IEEE J. Ocean. Eng. **14**, 360–367 (1989).
[18] A. P. Lyons, W. L. J. Fox, T. Hasiotis, and E. Pouliquen, ''Characterization of the two-dimensional roughness of wave-rippled sea floors,'' IEEE J. Ocean. Eng. (submitted).

Orientation effects on linear time-reversing array retrofocusing in shallow water

Michael R. Dungan[a] and David R. Dowling[b]
Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109-2121

(Received 14 February 2001; revised 17 July 2002; accepted 29 July 2002)

A time-reversing array (TRA) can retrofocus acoustic energy, in both time and space, to the original sound-source location without any environmental information. This paper presents results from an analytical and computational investigation into the effects that array orientation has on linear TRA retrofocusing in shallow water environments. A linear TRA has three limiting orthogonal orientations with respect to a distant sound source in a shallow water waveguide: vertical, endfire, and broadside. Here, TRA retrofocus characteristics are predicted for monochromatic sound propagation in a Pekeris waveguide using a modal sum Green's function and in a more realistic sound channel having vertical sound speed variation using a computed Green's function. Both analytical and computational results are compared for the three orthogonal array orientations with vertical arrays performing best. Differences in TRA retrofocusing performance in the three orientations are primarily determined by geometrical considerations and the extra mode-shape weighting inherent in the back-propagated field of horizontal TRAs. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1508787]

PACS numbers: 43.30.Vh, 43.30.Yj, 43.30.Ft [DLB]

I. INTRODUCTION

Hydrophone arrays commonly used in underwater acoustic experiments are linear and vertical, and are typically moored from the ocean floor and suspended by a float, or dipped from an anchored vessel and vertically aligned by a weight. However, many important sonar applications involve towed arrays that are more nearly horizontal. This paper presents and compares time-reversing array (TRA) retrofocusing performance for different array orientations in range-independent waveguides and provides illustrations of predicted horizontal TRA performance.

The performance of TRAs in waveguides has been the subject of much recent theoretical, computational, and experimental attention (Kuperman *et al.*, 1998; Song *et al.*, 1998, 1999; Hodgkiss *et al.*, 1999; Dungan and Dowling, 2000, 2001; Roux and Fink, 2000; Khosla and Dowling, 2001; Kim *et al.*, 2001; Sabra *et al.*, 2002). In particular, the oceanic experiments (the first four references above plus Kim *et al.*, 2001) show that acoustic time reversal is a very promising technology for retrofocusing sound in the shallow ocean. However, these previous studies have only considered linear vertical arrays even though practical implementation of TRA concepts in the ocean may require more nearly horizontal or towed array deployments. The purpose of this paper is to illustrate the inherent limitations of the three limiting cases of linear TRA orientation (vertical, endfire, and broadside) for TRAs responding to a single remote time-harmonic source in a shallow ocean waveguide. Specific results are drawn from a Pekeris waveguide using a modal sum Green's function, and, in a more realistic sound channel, having vertical sound speed variation and bottom losses, using a Green's function computed from the parabolic-approximation acoustic propagation code RAM (Collins, 1993, 1994, 1998). Taken together, these results illustrate important generic features of a linear TRA's back-propagated field that will be present in any waveguide environment.

A variety of array geometries and orientations have been considered for other transducer array technologies with the most notable being matched field processing (MFP), a powerful technique for determining of the location of an underwater source [see Tolstoy (1993) or Jensen *et al.* (1994)]. In particular, Bogart and Yang (1994) describe how MFP results depend on array orientation, transducer spacing, and transducer number. Yang and Yates (1998) describe MFP with horizontal arrays in the linear-array beam domain. More recently, Tantum and Nolte (2000) discussed ways of improving MFP results by properly designing the array.

MFP and TRA retrofocusing are conceptually similar but differ in their intent and requisite environmental information. MFP can be used to efficiently locate sound sources in known environments from measurements made at known locations. A TRA can be used to acoustically illuminate an unknown sound source in an unknown but reciprocal environment using transducers at unknown locations. Although MFP with horizontal arrays has been previously studied, consideration of horizontal TRA retrofocusing performance has not. In particular, the orientation of the TRA will determine the extent of side lobes near the retrofocus location. While the findings presented here parallel those in previous horizontal-linear-array MFP studies, the interpretations and specifications of the performance limitations of horizontal arrays and the material concerning array depth are new.

The remainder of this paper is divided into three sections; the next one briefly introduces the foundations of acoustic time reversal and presents some formal analytical results highlighting the differences between horizontal and

[a]Currently at Raytheon Corporation, Tucson, AZ.
[b]Author to whom correspondence should be addressed. Electronic mail: drd@engin.umich.edu

vertical arrays. The third section presents results for the three limiting orientations (vertical, endfire, and broadside) of a TRA in two different range-independent shallow water waveguides. The fourth and final section summarizes the findings and states the conclusions of this research.

II. NARROW-BAND ACOUSTIC TIME REVERSAL

Active acoustic time-reversing arrays (TRAs) operate by listening to a remote sound source and then broadcasting what was heard back into the environment in a time reversed fashion. When the array has sufficient aperture and the environment is reciprocal, the array-broadcast sound will retrace its path(s) back to its point of origin where it will retrofocus at the location of the original sound source even when the environment, the array element locations, and the sound source location are all unknown. The quality and lifetime of this retrofocus field are determined by the characteristics of the environment, the sound signal, and the source-array geometry. Here, the emphasis is on the influence of the final factor in shallow ocean sound channels. The formal basis for acoustic time reversal is provided in Jackson and Dowling (1991) using temporal Fourier transforms. Sabra *et al.* (2002) provides an entirely time domain implementation of this material. This subsection describes the narrow-band time-reversing array response function, $G_{c\omega}$, and how it differs for vertical and horizontal arrays.

The response of a time-reversing array (TRA) is entirely described by the so-called conjugate Green's function, $G_{c\omega}$ [see Jackson and Dowling (1991), Eq. (10), or Kuperman *et al.* (1998), Eq. (7)]. For a TRA composed of discrete monopole transducers, $G_{c\omega}$ can be defined by:

$$G_{c\omega}(x,r,z,t_2-t_1) = \sum_{n=1}^{N} G_{2\omega}(x,r,z;x_n,r_n,z_n)G^*_{1\omega}(x_n,r_n,z_n;0,0,z_s) \quad (1)$$

where x, $y=r$, and z are Cartesian coordinates—cross range, range, and depth—respectively, N is the number of array elements, $G_{j\omega}(x,r,z;x',r',z')$ is the Helmholtz equation Green's function at radian frequency ω for acoustic propagation from (x',r',z') to (x,r,z) at time t_j, and the asterisk denotes complex conjugation. The time delay, t_2-t_1, is important in dynamic environments but is taken to be zero for this study; thus, the time index, j, and the time delay argument, t_2-t_1, are dropped in subsequent equations.

As illustrated in Fig. 1 a linear TRA has three limiting orientations possible within the shallow ocean: vertical, endfire, and broadside. The vertical TRA is parallel to the z axis and has elements located at $(0,R,z_n)$. The endfire TRA is parallel to the range direction (y axis) and has elements at $(0,r_n,z_a)$. The broadside TRA is parallel to the x axis and has elements at (x_n,R,z_a). Other straight-array orientations fall at some rotation angle between these three limiting orthogonal cases. All the arrays studied herein have regular element spacing. The spacing variables for vertical, endfire, and broadside TRAs are Δz, Δr, and Δx, respectively. Unless otherwise specified, the horizontal TRAs in this study are all deployed at mid-water column, and R is the distance between the source and the center of the array.

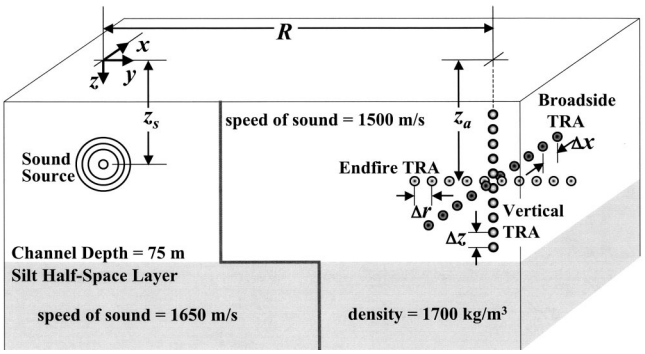


FIG. 1. Geometry of TRA orientations in a simple sound channel: vertical, horizontal endfire, and horizontal broadside. The sound channel is a Pekeris waveguide of 75-m depth. There is no attenuation present in the bottom. Parameter values are listed in the figure.

To develop some understanding about the likely performance differences between vertical and horizontal TRAs in a range-independent sound channel, consider a modal sum Green's function [see Jensen *et al.* (1994) or Kinsler *et al.* (2000)]:

$$G_\omega(x,r,z;0,0,z_s) = \sum_{m=1}^{M} \frac{i}{\rho(z_s)} \left(\frac{1}{8\pi k_m \sqrt{r^2+x^2}} \right)^{1/2} \times \psi_m(z_s)\psi_m(z) \times \exp\{i(k_m\sqrt{r^2+x^2}-\pi/4)\} \quad (2)$$

for a unit strength harmonic point source at $(0,0,z_s)$ with radian frequency ω. Here, M is the total number of propagating modes, k_m is the horizontal wave number of mode m, $\psi_m(z)$ is the vertical mode shape for mode m, and the range r is large enough for accurate use of the usual asymptotic approximation of the Hankel function.

As a further simplification for the purposes of analytical illustration, consider only the TRA produced field $G_{c\omega}(0,0,z)$ on a vertical line through the position of the original source, $(0,0,z_s)$. The combination of (1) and (2) thus produces

$$G_{c\omega}(0,0,z) = \sum_{n=1}^{N}\sum_{p=1}^{M}\sum_{m=1}^{M} \left(\frac{1}{k_p k_m (r_n^2+x_n^2)} \right)^{1/2} \times \frac{\psi_p(z)\psi_p(z_n)\psi_m(z_n)\psi_m(z_s)}{8\pi\rho(z_s)\rho(z_n)} \times \exp\{i(k_p-k_m)\sqrt{r_n^2+x_n^2}\} \quad (3)$$

for a TRA with N elements located at (x_n,r_n,z_n).

For a well-populated water-column-spanning vertical array ($r_n=$const, $x_n=0$), the sum over array elements becomes proportional to an integral over the depth. If Δz is the element spacing, the orthonormality of the vertical modes allows (3) to be simplified to a direct sum over array elements:

$$G_{c\omega}(0,0,z) = \sum_{m=1}^{M} \frac{1}{8\pi\Delta z\rho(z_s)} \left(\frac{1}{k_m r_n} \right) \psi_m(z)\psi_m(z_s), \quad (4)$$

a result presented earlier in Kuperman *et al.* (1998) with $\Delta z=1$. At the source depth, all the terms in the sum will be

positive and this leads to the TRA retrofocus. At other depths, the vertical spatial orthogonality properties of the modes will ensure a high degree of field cancellation between modes and a low side lobe level.

For a horizontal array with elements at depth $z_n = z_a$, there are no simplifications of (3) based on mode orthogonality. However, following Bogart and Yang (1994) for the conceptually similar Bartlett matched field processor, the double modal sum in (3) can be separated into diagonal ($m = p$) and off-diagonal ($m \neq p$) terms for a horizontal TRA:

$$G_{c\omega}(0,0,z) = [G_{c\omega}(0,0,z)]_{m=p} + [G_{c\omega}(0,0,z)]_{m\neq p}, \quad (5a)$$

$$[G_{c\omega}(0,0,z)]_{m=p} = \sum_{n=1}^{N} \left[\sum_{m=1}^{M} \left(\frac{1}{k_m \sqrt{r_n^2 + x_n^2}} \right) \times \frac{\psi_m(z)\psi_m^2(z_a)\psi_m(z_s)}{8\pi\rho(z_s)\rho(z_n)} \right], \quad (5b)$$

$$[G_{c\omega}(0,0,z)]_{m\neq p} = \sum_{n=1}^{N} \sum_{p\neq m}^{M} \sum_{m=1}^{M} \left(\frac{1}{k_p k_m (r_n^2 + x_n^2)} \right)^{1/2} \times \frac{\psi_p(z)\psi_p(z_a)\psi_m(z_a)\psi_m(z_s)}{8\pi\rho(z_s)\rho(z_n)} \times \exp\{i(k_p - k_m)\sqrt{r_n^2 + x_n^2}\}. \quad (5c)$$

When there are many elements and the array is long or well dispersed, the off-diagonal terms will tend to cancel with each other because the complex phase factor in (5c) will distribute real and imaginary parts over both positive and negative values. Unfortunately, even when the off-diagonal terms are self-canceling, the vertical array result (4) for $G_{c\omega}$ is not equal to $[G_{c\omega}(0,0,z)]_{m=p}$ in (5b). The difference is the additional sum over array elements, and the appearance of an extra factor of $\psi_m^2(z_a)$ in each term of the diagonal-element sum. These differences are important, a contention that will be further illustrated in the next section. However, the structure of (5) implies that any persistent horizontal TRA side lobes will be determined by the diagonal terms because the off-diagonal terms are likely to be self-canceling and may vary rapidly with small changes in element spacing and number, and array length.

To illustrate TRA performance for the different array orientations, narrow-band acoustic time reversal was studied in two waveguide environments. The first, shown in Fig. 1, is a Pekeris waveguide with a depth of 75 m. The half-space bottom has a density of 1700 kg/m^3 and a speed of sound of 1650 m/s. There is no attenuation in the bottom, the water sound speed profile is isovelocity at 1500 m/s, and the ocean surface is treated as a flat pressure-release surface. The second environment, shown in Fig. 2, has the same depth as the first and is based on conditions measured near the center of the SWARM '95 experiment (see Apel *et al.*, 1997). It has been used in prior shallow water TRA studies (Dungan and Dowling, 2000, 2001). The speed of sound profile was calculated from smoothed temperature and salinity measurements, and the bottom is a two-layer model with density and sound speed extrapolated from core and chirp measurements. Attenuation values were chosen conservatively from the tabulations in Jensen *et al.* (1994).

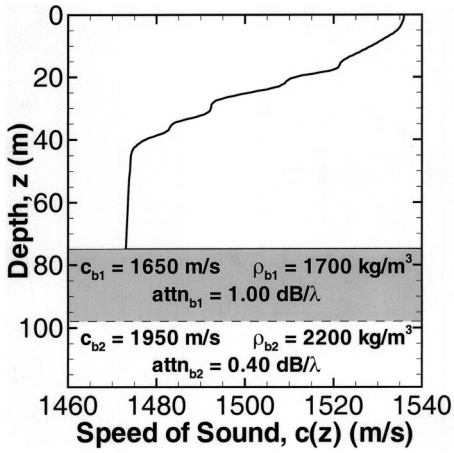


FIG. 2. More realistic sound channel also having 75-m-depth-based measurements taken during the SWARM'95 experiment. Parameter values are listed in the figure; c_{b1} and c_{b2} are the sound speeds in the upper and lower bottom layers, and ρ_{b1} and ρ_{b2} are the two bottom layer densities. The attenuation values used were chosen conservatively from tabulations in Jensen *et al.* (1994).

Although the second environment is more realistic, it still lacks range dependence, bubble scattering, bubble absorption, bottom roughness, and ocean-surface and internal-wave dynamics. These phenomena are beyond the scope of this study but some have been the subject of prior efforts. The effects of internal wave dynamics are simulated in Dungan and Dowling (2000), and assessments of rough surface scattering and TRA performance are provided in Kuperman *et al.* (1998), Khosla and Dowling (1998), and Dungan and Dowling (2001).

III. RESULTS FOR DIFFERENT ARRAY ORIENTATIONS

In range-independent sound channels, the three TRA orientations produce different results for retrofocus size in the cross-range, range, and depth directions, and for the prevalence of side lobes in the vicinity of the source. The two parameters considered here beyond orientation are the array depth and array element spacing.

Comparisons of the retrofocus fields produced by 24-element TRAs having the three orthogonal orientations are shown in Figs. 3 and 4 for the Pekeris and more-realistic shallow water waveguides, respectively. Each frame shows a range-depth slice through the retrofocus field at a frequency of 500 Hz, a source-array range of 10 km, and a source depth of 25 m. The array element spacing is varied between the three frames of Figs. 3 and 4 to produce near optimum retrofocus results for each orientation. In both environments, the water-column-spanning vertical TRA ($\Delta z = 3$ m) produces the best retrofocus [Figs. 3(a) and 4(a)] with the fewest side lobes through the depth. A somewhat longer endfire array ($\Delta r = 50$ m) produces the second best retrofocus with only a few side lobes visible in its range-depth field slices [Figs. 3(b) and 4(b)]. The one strong side lobe shown at a 50-m depth in Fig. 3(b) is related to the acoustic intensity peak that occurs at the complementary depth of a sound source in uniform-sound-velocity waveguides (Weston, 1980). This strongest side lobe will be analyzed fully later in this section. The very large aperture broadside TRA (Δx

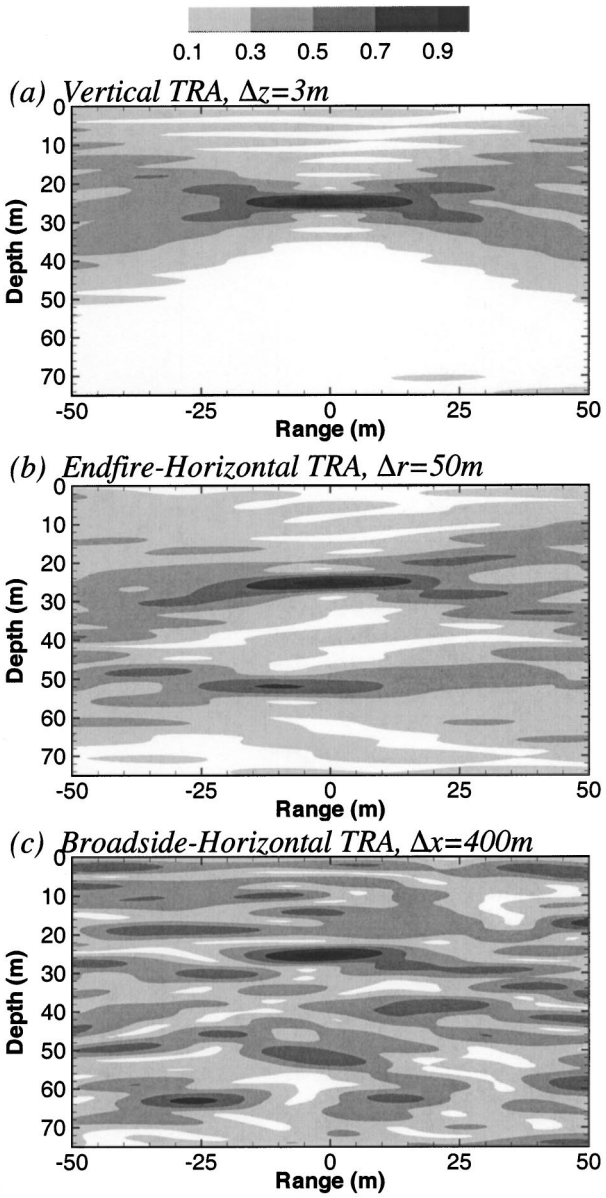


FIG. 3. Retrofocus plots through range and depth in the Pekeris waveguide of Fig. 1 for various TRA orientations and element spacing for a source depth of $z_s = 25$ m, a source-array range of $R = 10$ km, and an acoustic frequency of $f = 500$ Hz. The computed results are from a vertical TRA with 3-m element spacing is shown in (a), an endfire TRA with a 50-m element spacing in (b), and a broadside TRA with a 400-m spacing in (c).

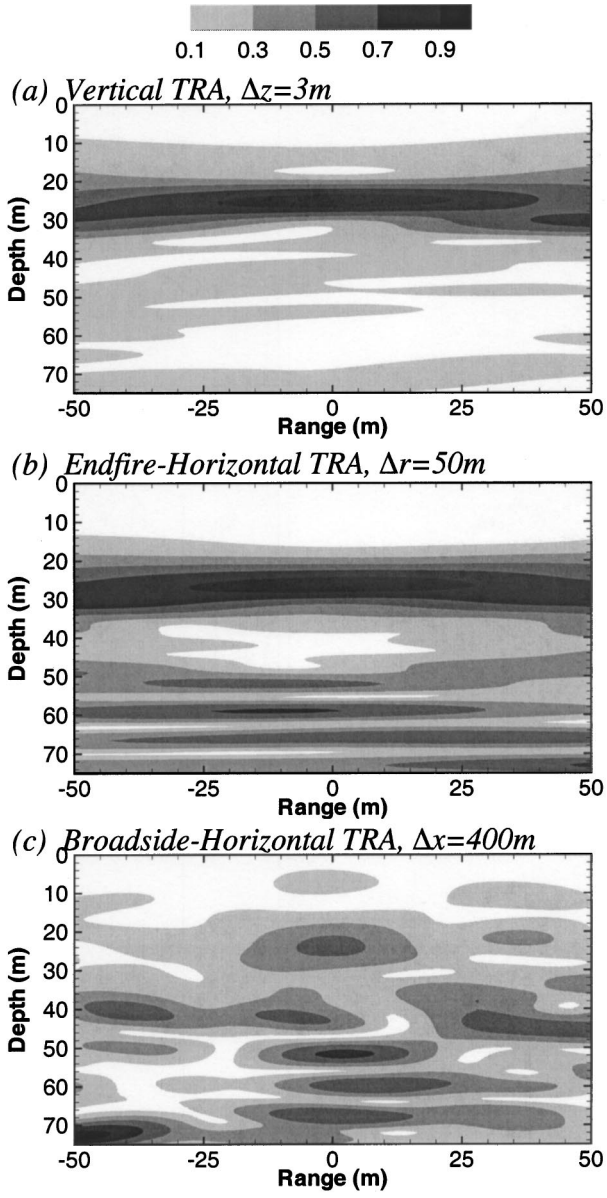


FIG. 4. Same as Fig. 3 except the results are for the more realistic sound channel of Fig. 2.

$= 400$ m) struggles to provide even minimal side lobe suppression through depth and range [Figs. 3(c) and 4(c)]. For these figures, the source-array range of 10 km lies within the Fresnel near-field of both horizontal TRAs. Therefore, when compared to a water-column spanning vertical TRA, these arrays may produce comparable or slightly smaller retrofocus sizes in the range (or y) direction, and—in horizontally uniform waveguides—very much smaller retrofocus sizes in the azimuthal (or x) direction. Additional discussion of range- and azimuthal-direction focusing is deferred to the end of this section.

The relationship between the retrofocus-field structure along a linear vertical intersection at the original source location and the array element spacing is illustrated for the two environments in Figs. 5 and 6 for the three orthogonal orientations at a frequency of 500 Hz, a source-to-center-of-array range of 10 km, and a source depth of 25 m with $N = 24$. Note that the horizontal axes of the three frames in each figure are different. The field values in each vertical intersection were normalized by their maximum value in that intersection, with black corresponding to the peak field amplitude. The vertical TRA [Figs. 5(a) and 6(a)] produces the tightest vertical retofocus with the lowest side lobes and the shortest overall length. It performs well once it spans more than half the water column. The endfire horizontal TRA [Figs. 5(b) and 6(b)] is second best. In both environments, its vertical retrofocus field structure is stable once the element spacing is 25 m or so. However, in both environments, the endfire TRA generates one or more persistent side lobes through the depth that are not produced by the vertical array in the same environment. The broadside TRA [Figs. 5(c) and 6(c)] is the least effective in suppressing side lobes on the vertical cuts through the retrofocus field. Moreover, the

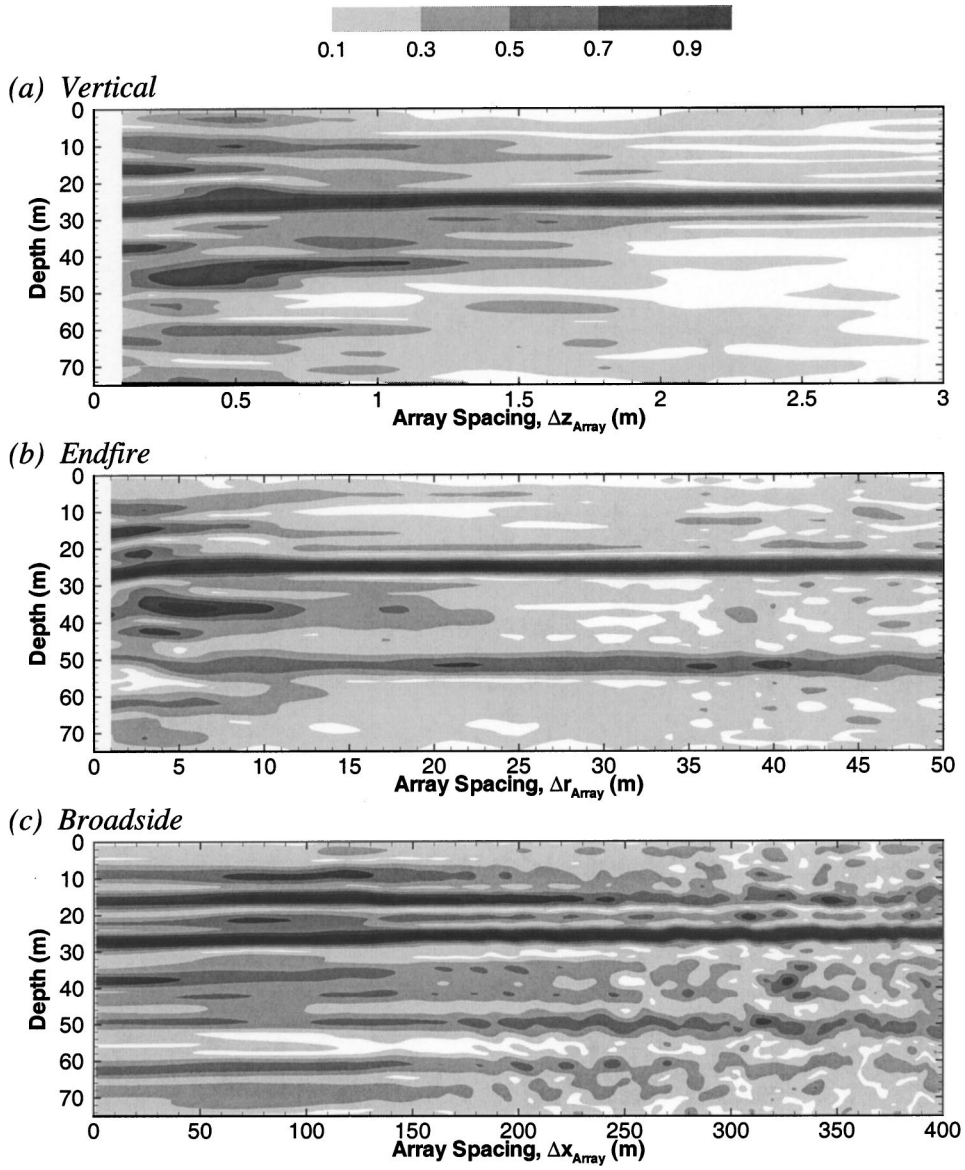


FIG. 5. Compilation of vertical intersections through the intended retrofocus location for various element spacing for a vertical TRA (a), an endfire TRA (b), and a broadside TRA (c) in the Pekeris waveguide of Fig. 1. Each array is composed of 24 elements and is centered at a depth of $z_a = 37.5$ m and a source-to-center-of-array range of $R = 10$ km. The acoustic frequency is $f = 500$ Hz with a source depth $z_s = 25$ m. Each vertical profile in (a), (b), and (c) is normalized to the maximum pressure value on that profile. Black indicates the highest field amplitudes.

broadside array must be several kilometers long ($\Delta x \geqslant 250$ m) to achieve a marginal retrofocus through the depth in the more realistic environment.

The explanation for the clear vertical-side-lobe-structure performance loss upon changing from a vertical to a horizontal TRA lies in the fundamental limitations of horizontal TRAs to resolve and properly weight the different propagating modes. These horizontal array limitations arise from nonzero element spacing (or imperfect spatial sampling), finite array aperture, and the vertical structure of the propagating modes. These causes of horizontal TRA retrofocusing imperfection are discussed in turn in the following paragraphs.

First consider the influence of element spacing. When the element spacing (Δr) of an endfire horizontal array is larger than half of the shortest horizontal wavelength (the usual Nyquist criterion), aliasing may occur because of insufficient spatial sampling by the array. For acoustic fields that are band-limited in horizontal wave number k_m ($k_M < k_m < k_1$), such as those in a shallow ocean waveguide, insufficient spatial sampling will cause problems with TRA performance when $\pi/k_1 < \Delta r < \pi/k_M$, and when $\Delta r > 2\pi/(k_1 - k_M)$ where k_1 is the horizontal wave number of the first propagating mode and k_M is the horizontal wave number of highest-order propagating mode. The first region of poor spatial sampling occurs when the element spacing is slightly larger than half an acoustic wavelength. Here, the low-order propagating modes, which have k_m values of nearly $\omega/c = k$ (where c is an average speed of sound), can be spatially aliased on TRA broadcast to higher-order modes having lower k_m values. The second region of poor spatial sampling occurs for much larger element spacing. A simple explanation of the existence of this second region is a follows. When the band-limited horizontal spatial spectrum of the acoustic field from the original source is taken to be two sided, the second region of poor spatial sampling occurs when negative spatial wave numbers in the band $-k_1 < k_m < -k_M$ are aliased on TRA broadcast onto positive wave numbers in the band $k_M < k_m < k_1$. In both regions, aliasing causes retrofocus field side lobes that depend strongly on the element spacing. For the Pekeris waveguide of Fig. 1, $k_1 = 2.094\ \mathrm{m}^{-1}$ and $k_M = k_{21} = 1.910\ \mathrm{m}^{-1}$ so the aliasing regions occur near $\Delta r \approx 1.6$ m and for $\Delta r > 34$ m on Fig. 5(b). Such aliasing effects

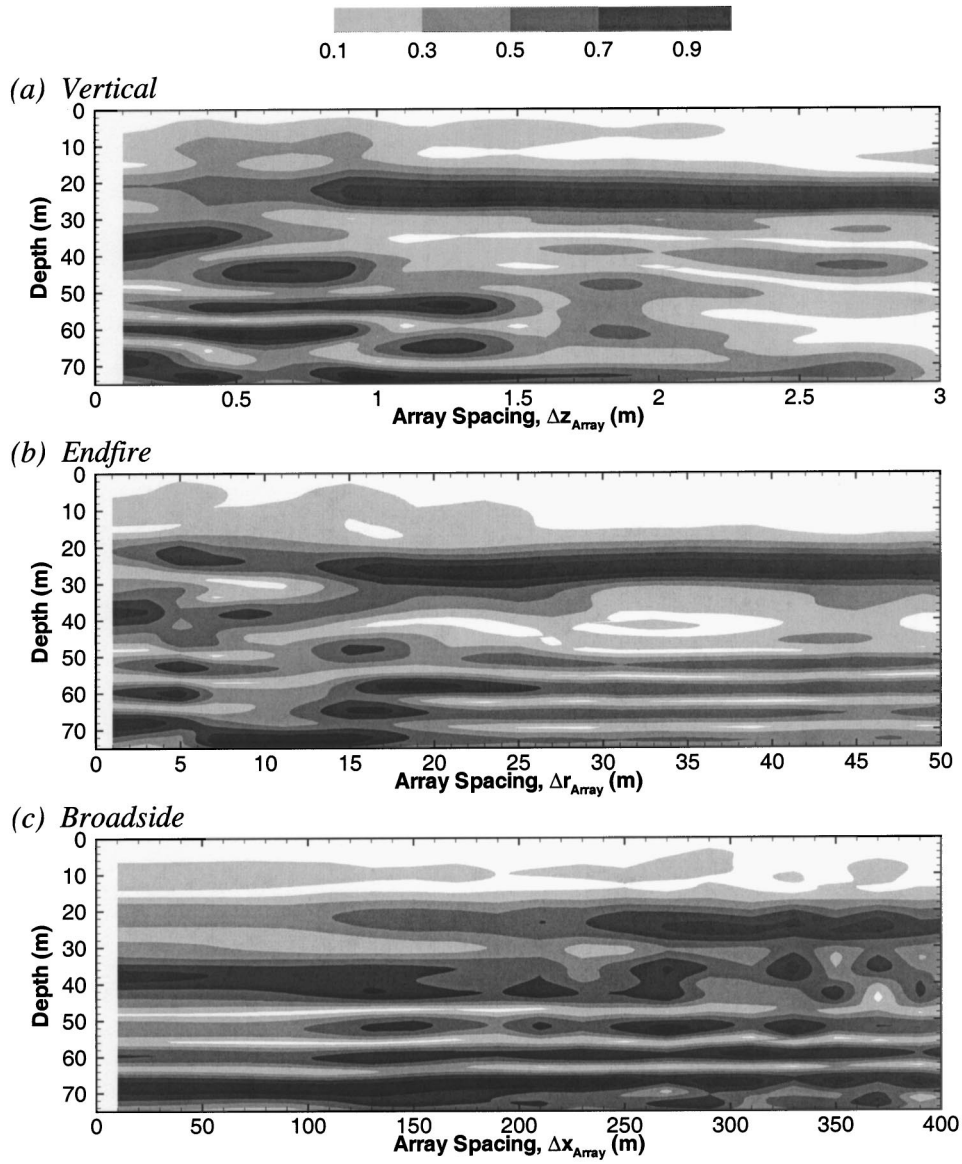


FIG. 6. Same as Fig. 5 except the results are for the more realistic sound channel of Fig. 2.

can be eliminated by setting $\Delta r < \pi/k_1$ or $\pi/k_M < \Delta r < 2\pi/(k_1 - k_M)$, or can be mitigated somewhat by using mildly uneven element spacing.

The influence of array aperture on endfire horizontal TRA retrofocusing involves the length of the array $L=(N-1)\Delta r$ and the horizontal wave numbers, k_m, of the propagating modes. Neighboring propagating modes may have k_m values that differ by less than 0.1%. Hence, a horizontal endfire array should also be long enough to resolve the two propagating modes having the closest values of k_m, i.e., L must satisfy $\min(k_{m-1}-k_m)L \geqslant 1$. For the Pekeris waveguide, Fig. 7 presents the fraction of modes that are resolved by an endfire TRA as a function of element spacing. Here, there are 21 propagating modes and the mth propagating mode is considered to be properly resolved when $(k_{m-1}-k_m)L \geqslant 1$. A comparison of Fig. 5(b) and Fig. 7 illustrates this array length effect. Figure 5(b) shows that the side lobes grow weaker when the array length increases until the array is long enough to resolve all the propagating modes. This length requirement for horizontal arrays is also covered in Tantum and Nolte (2000) in the context of MFP and modal cross-correlation matrices.

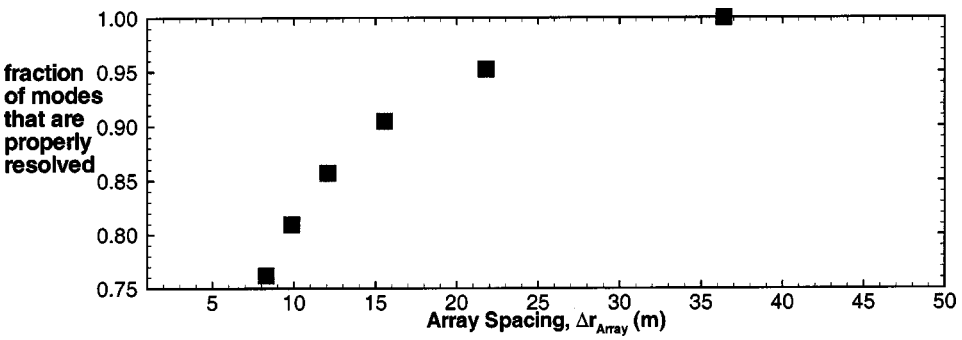


FIG. 7. Plot of the percentage of Pekeris waveguide propagating modes properly resolved by a linear horizontal endfire TRA for various element spacing. The horizontal axis is the same as in Fig. 5(b) to which this figure should be compared. Endfire TRA retrofocusing is best and nearly independent of element spacing when the fraction of resolved propagating modes is close to unity.

For a fixed number (N) of transducers, it may be impossible to satisfy both the element-spacing and array-length requirements. In fact, it might be impractical to build an array that is better than half-wavelength spaced ($\Delta r < \pi/k_1$), and long enough [$\min(k_{m-1}-k_m)L \geqslant 1$] to resolve all the propagating modes. For example, this would require $N \geqslant 560$ in the Pekeris waveguide of Fig. 1. However, aliasing problems are confined to certain element-spacing ranges when the acoustic field from the source is band limited in horizontal wavenumber. Thus, array designs based on a few tens of elements should be able to perform well when $\pi/k_M < \Delta r < 2\pi/(k_1 - k_M)$ and when $\min(k_{m-1}-k_m)L \geqslant 1$. In fact, the minimum number of elements ($N_{\min}$) can be estimated from the largest Δr and the smallest L that satisfy these two criteria:

$$N_{\min} \approx 1 + \frac{k_1 - k_M}{2\pi \cdot \min(k_{m-1}-k_m)}. \tag{6}$$

For the Pekeris waveguide of Fig. 1, this estimate yields $N_{\min} \approx 26$, and $L = (N-1)\Delta r = (25)[2\pi/(k_1 - k_M)] \approx 850\,\text{m}$. The criteria for Δr and L, and the result (6) are important because they can be used to set horizontal array element spacing, overall length, and element number based on the propagating-mode wave numbers of the sound channel in which the array will be used.

The final cause of horizontal TRA retrofocusing imperfection arises from the vertical structure of the propagating modes. Examination of the mid-water-column endfire array's strong and persistent side lobe in the Pekeris environment [see Fig. 5(b)] provides insight into this limitation of horizontal TRAs.

When an endfire TRA is located at mid-water column in a Pekeris waveguide and the source is located at depth z_s, a strong side lobe can be found at depth $D - z_s$ independently of the length of the array or the number of elements. This contention can be deduced from (5b) using approximate mode shapes or inferred from the concept of the complementary source depth (Weston, 1980). Here, we follow the former path. When the acoustic frequency is high enough so that many modes propagate in a Pekeris waveguide, the shapes of the low order modes—the ones furthest from being cut off—are approximately sinusoidal functions that satisfy pressure release boundary conditions at the channel surface $z = 0$ and at the water column depth $z = D$ (see Kinsler *et al.*, 2000):

$$\psi_m(z) \approx \sqrt{2\rho/D}\,\sin(m\pi z/D). \tag{7}$$

To set a performance gage for a horizontal array, first examine the behavior of the water-column-spanning vertical TRA. Combining (7) with (4) produces

$$G_{c\omega}(0,0,z) = \sum_{m=1}^{M} \frac{1}{8\pi\Delta z \rho(z_s)} \left(\frac{1}{k_m r_n}\right) \psi_m(z)\psi_m(z_s) = \sum_{m=1}^{M} \left(\frac{1}{4\pi\Delta z k_m r_n D}\right) \sin\left(\frac{m\pi z}{D}\right) \sin\left(\frac{m\pi z_s}{D}\right), \tag{8}$$

which, when evaluated at the retrofocus $z = z_s$, yields

$$G_{c\omega}(0,0,z_s) = \sum_{m=1}^{M} \frac{1}{8\pi\Delta z \rho(z_s)} \left(\frac{1}{k_m r_n}\right) \psi_m^2(z_s) = \sum_{m=1}^{M} \left(\frac{1}{4\pi\Delta z k_m r_n D}\right) \sin^2\left(\frac{m\pi z_s}{D}\right), \tag{9}$$

where here M can be interpreted as the number of propagating modes for which (7) is valid since (7) can not be valid for all propagating modes. All the terms in the final sum of (9) are all positive so each mode makes an in-phase contribution at the retrofocus. To see how a vertical array avoids a strong persistent side lobe, evaluate (8) at the anticipated strong side lobe location ($a.k.a.$ the complementary depth), $z = D - z_s$,

$$G_{c\omega}(0,0,D-z_s) = \sum_{m=1}^{M} \left(\frac{1}{4\pi\Delta z k_m r_n D}\right) \sin\left(\frac{m\pi(D-z_s)}{D}\right) \sin\left(\frac{m\pi z_s}{D}\right) = \sum_{m=1}^{M} \left(\frac{1}{4\pi\Delta z k_m r_n D}\right) (-1)^{m+1} \sin^2\left(\frac{m\pi z_s}{D}\right). \tag{10}$$

Here, the alternating signs of the terms in the final sum will ensure that $G_{c\omega}(0,0,D-z_s)$ is small compared to $G_{c\omega}(0,0,z_s)$. Thus, a strong side lobe will not exist at $z = D - z_s$ for the water-column-spanning vertical array.

Now compare these vertical array findings, (9) and (10), with equivalent horizontal array results. As argued earlier, the persistent features of a horizontal TRA's retrofocus field will be present in $[G_{c\omega}(0,0,z)]_{m=p}$. Thus, combining (7) with the diagonal terms, (5b), produces the horizontal-array equivalent of (8):

$$[G_{c\omega}(0,0,z)]_{m=p} \approx \sum_{n=1}^{N} \left[\sum_{m=1}^{M} \left(\frac{1}{2\pi k_m \sqrt{r_n^2 + x_n^2 D^2}} \right) \times \sin^2\left(\frac{m\pi z_a}{D}\right) \sin\left(\frac{m\pi z}{D}\right) \sin\left(\frac{m\pi z_s}{D}\right) \right], \tag{11}$$

where z_a is the depth of the array. Place the array at mid-water column, $z_a = D/2$, and evaluate (11) at the retrofocus $z = z_s$, to find

$$[G_{c\omega}(0,0,z_s)]_{m=p} \approx \sum_{n=1}^{N} \left[\sum_{m=1}^{M} \left(\frac{1}{2\pi k_m \sqrt{r_n^2 + x_n^2 D^2}} \right) \times \sin^2\left(\frac{m\pi}{2}\right) \sin^2\left(\frac{m\pi z_s}{D}\right) \right]. \tag{12}$$

As for the vertical array result (9), all the terms in the final double sum in (12) are positive so a field amplitude peak at the retrofocus depth is guaranteed. When (11) is evaluated with $z_a = D/2$ at the presumed strong side lobe location, $z = D - z_s$, then

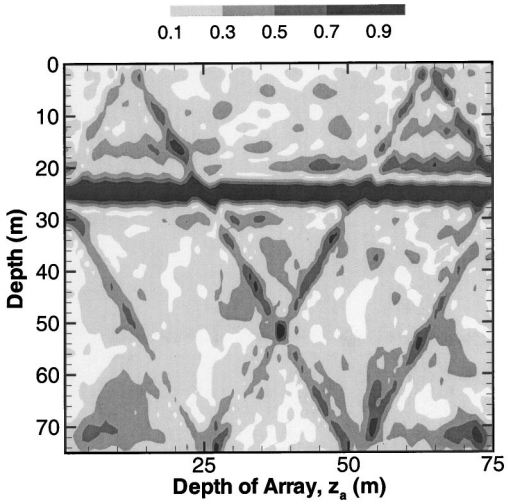


FIG. 8. Vertical slices through the retrofocus for endfire-horizontal TRAs for different array depths in the Pekeris waveguide of Fig. 1. The source depth is 25 m and the acoustic frequency is 500 Hz. The average source-array range is 10 km and the element spacing is 20 m. Two strong side lobes are possible. These overlap when the array is at mid-water-column.

$$[G_{c\omega}(0,0,D-z_s)]_{m=p} \approx \sum_{n=1}^{N}\left[\sum_{m=1}^{M}\left(\frac{1}{2\pi k_m\sqrt{r_n^2+x_n^2}D^2}\right)\sin^2\left(\frac{m\pi}{2}\right) \times \sin\left(\frac{m\pi z_s}{D}\right)\sin\left(\frac{m\pi(D-z_s)}{D}\right)\right]. \tag{13}$$

Here, the trigonometric terms appear to allow for positive and negative contributions in the sum over modes. However, manipulation of the trigonometric functions (see Appendix) yields

$$\sin^2\left(\frac{m\pi}{2}\right)\sin\left(\frac{m\pi z_s}{D}\right)\sin\left(\frac{m\pi(D-z_s)}{D}\right) = \sin^2\left(\frac{m\pi}{2}\right)\sin^2\left(\frac{m\pi z_s}{D}\right). \tag{14}$$

Thus, when $z_a=D/2$ and the propagating modes behave as if the sound channel is bounded by two pressure release surfaces, $[G_{c\omega}(0,0,z_s)]_{m=p}$ and $[G_{c\omega}(0,0,D-z_s)]_{m=p}$ are identical. Interestingly, the main retrofocus and the strong side lobe are predicted to be in-phase by this approximate analysis. Thus, for broadband TRA operations, the signal returned by the TRA to the strong side lobe depth will be an approximate replica of the intended retrofocused signal at the source depth. Moreover, the strong side lobe, illustrated in Fig. 5(b) and embodied in (13) and (14), will not be weakened or eliminated by increasing the number of array elements or lengthening the array.

When a TRA in a Pekeris waveguide is not at mid-water column, it may generate two persistent side lobes through the vertical. These side lobes happen to coincide when the array is at mid-water column. Figure 8 illustrates this phenomenon for a source depth of 25 m, an average source-array range of 10 km, and a frequency of 500 Hz for a 24-element endfire TRA having an element spacing of 20 m. Here, the TRA's field amplitude on a vertical slice through the source location is plotted for different TRA depths. The horizontal and vertical axes on Fig. 8 are the array depth, z_a, and z, the vertical coordinate, respectively. The depths of the two persistent side lobes z_1 and z_2 can be deduced by extrapolating these geometric patterns to other source depths,

$$z_1=2D-2z_a-z_s, \tag{15a}$$

$$z_2=2z_a-z_s, \tag{15b}$$

where if $-D\leqslant z_i<0$, then z_i is replaced by $|z_i|$, or if $D<z_i\leqslant 2D$, then z_i is replaced by $2D-z_i$.

Although the discussion of (7)–(15) directly applies only to TRAs in Pekeris waveguides, the same phenomenon—persistent side lobes through the depth—occurs in the other waveguide as well. For example, Fig. 6(b) shows multiple persistent side lobes as the element spacing is increased for an endfire horizontal TRA. These side lobes occur for the same reason as the one strong side lobe in the Pekeris case [Fig. 5(b)], but differ in detail because of differences in the vertical structure of the propagating modes in the two environments. The single persistent side lobe in the Pekeris waveguide case results from the mid-water-column placement of the array and the vertical symmetry of the low-order propagating modes in that environment. Such vertical symmetry does not exit in the more realistic environment, and the lone strong side lobe is replaced by several weaker ones. In either environment, changes in the sound speed profile or the depth of the horizontal TRA will alter the vertical structure of the persistent side lobes.

The fundamental reason that horizontal TRAs produce persistent side lobes when water-column-spanning vertical TRAs do not lies in the extra modal-weighting inherent in horizontal TRAs. As expected, the back-propagated modes generated by a horizontal TRA do return to the intended retrofocus location in phase, but their amplitudes are modulated by an extra factor of $\psi_m^2(z_a)$ compared to modes returned by a water-column-spanning vertical TRA. Such mode-amplitude modulation can be important. For example, if the source strongly excites the mth propagating mode [i.e., z_s is an antinode of $\psi_m(z)$] and the horizontal TRA lies at a node of the mth mode [i.e., $\psi_m^2(z_a)=0$], then the back propagated field from the horizontal TRA will lack the contribution from the mth mode necessary for vertical-TRA-quality retrofocusing. Thus, the retrofocus field from such an mth-mode-nodal-positioned horizontal TRA will have a lower amplitude at the source depth and higher side lobe levels than that obtained in the same environment from a vertical array that properly weights the mth mode. Furthermore, when there are many propagating modes it may become impossible to find a good vertical location for a horizontal TRA unless the source depth and the mode structure of the environment are known.

Fortunately, a little bit of vertical aperture can at least partially remedy this mode-structure limitation of horizontal TRAs by preventing the entire TRA from lying on or too near propagating-mode nodal depths. This contention is supported by the endfire array results provided in Fig. 9 for the Pekeris waveguide environment. The two frames of this fig-

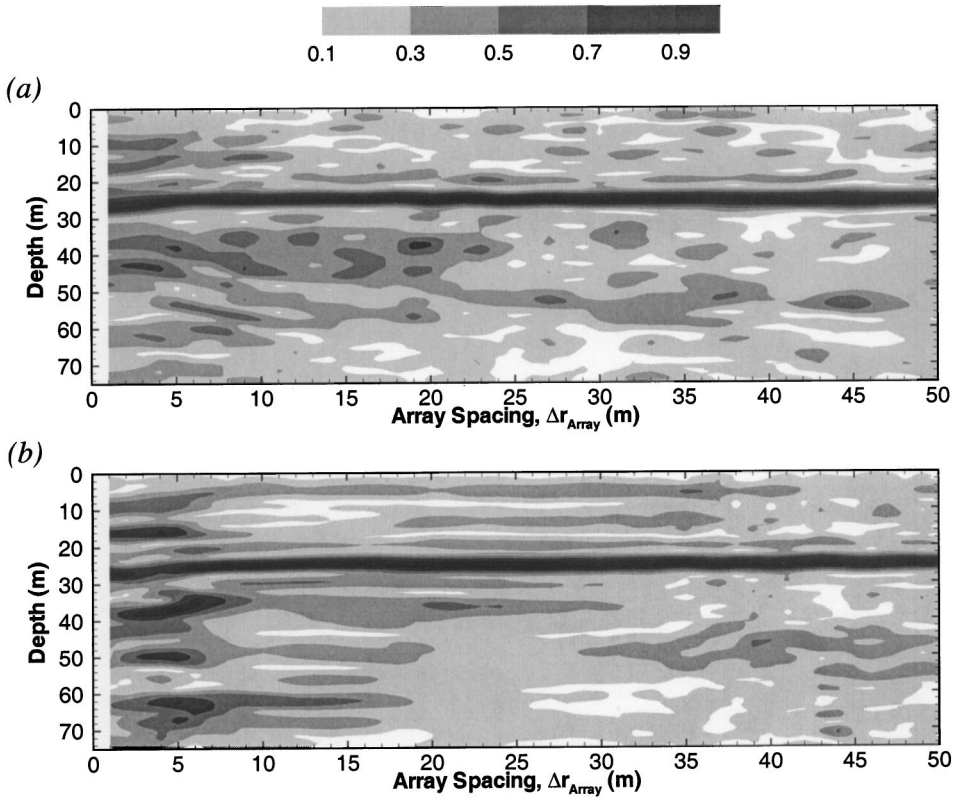


FIG. 9. Compilation of vertical intersections through the intended retrofocus location for slightly modified endfire arrays in the Pekeris waveguide of Fig. 1. In (a), the TRA elements have been randomly shifted with 3-m rms vertical displacement. In (b) the array has been tilted so that it lies on a mild diagonal and spans 6 m vertically at the center of the water column. Comparison of both frames to Fig. 3(b) shows that even a modest amount of vertical aperture helps control the persistent side lobe shown in Fig. 5(b).

ure are essentially identical to Fig. 5(b) except that in Fig. 9(a) a zero-mean random vertical displacement with root-mean-square deviation of 3 m was given to the elements of the array, and in Fig. 9(b) an end-to-end tilt from $+3$ m to -3 m was given to the array. Comparison of both frames of Fig. 9 with Fig. 5(b), shows that the strong side lobe is nearly eliminated by the addition of this small amount of vertical aperture (approximately two wavelengths). Hence, careful attention to keeping a towed endfire TRA perfectly straight and level is likely to be detrimental to the production of the best possible retrofocus.

For focusing through the water column depth, horizontal broadside TRAs face the same limitations as endfire TRAs, and examination of Figs. 5 and 6 reveals the signature of a persistent side lobe in Fig. 5(c) and several persistent side lobes in Fig. 6(c) similar to the endfire array results in Figs. 5(b) and 6(b). These side lobes form on a vertical line through the source location because the broadside array's length is effectively compressed in a horizontally uniform environment (see Tolstoy, 1993, p. 145). This effect can be explained by considering a horizontal array that forms a circular arc with all its elements equidistant from the source. For such a circular-arc TRA, the forward propagation from the source to each element will be identical. Furthermore, the backward propagation from each element to a vertical line through the source location will also be identical. Thus, the arc-TRA's retrofocus field on a vertical line through the source location will have the same vertical side lobe structure as that of a single-element TRA because the arc-TRA's field is merely the sum of N identical single-element-field vertical-line intersections. In this case, lengthening the circular arc TRA does not improve vertical retrofocusing. However, if the circular arc array is straightened into broadside array, the effective range-direction aperture of the array becomes the difference in source-element range between the closest and furthest array elements. This means that a broadside TRA with geometrical length L operating at source-to-center-of-array range of R is equivalent to a nonuniformly spaced endfire array with an effective length $\sqrt{R^2+L^2/4}-R$ ($<L/2$ when $R\neq 0$) when considering side lobes formed on vertical cuts through the source location. This effective length may even be much smaller than $L/2$ when $R>L$ so side lobes through the vertical at the source location may be abundant for broadside TRAs because the array cannot distinguish the propagating modes of the waveguide. Thus, broadside-array vertical-retrofocusing performance suffers twice because of source-array sound-channel geometry. First, a broadside TRA becomes effectively shorter than its true length so it may fail to distinguish the propagating modes of the sound channel, and, second, like the horizontal endfire array, its back propagated field improperly weights the propagating modes compared to a water-column-spanning vertical array. Fortunately, the geometrical length compression effect only influences the vertical retrofocus field of horizontal TRAs. In addition, range- and cross-range variations in natural oceanic sound channels may moderate, or even alleviate, the length compression effect.

For focusing in the range direction, vertical and the endfire TRAs produce comparable results because they primarily rely on the structure of the propagating modes to vertically confine the retrofocus, a topic well covered in Kim *et al.* (2001). For the broadside array, the range focusing is improved beyond that obtained with the other two arrays because it uses azimuthally different propagation paths to form the focal region [see Figs. 4(b) and (c)]. This effect is well

known and is equivalent to the linear-array free-space results [for example, see Ziomek (1995)] because the source-array propagation paths lie along straight lines in the $x-y$ plane in the horizontally uniform waveguide environment when viewed from above.

For focusing in the azimuthal direction, the broadside array is again the most advantageous, and, once again, free-space beamforming results provide guidance for TRA retrofocusing performance in a horizontally uniform waveguide because the propagation paths are straight when projected to the $x-y$ plane. Thus, in a horizontally uniform waveguide, the azimuthal or cross-range angular extent of a TRA retrofocus will be proportional to $1/kL$ for a broadside geometry, proportional to $1/\sqrt{kL}$ for a endfire geometry, and equal to 360 degrees when the TRA is vertical.

IV. CONCLUSIONS

This paper reports the findings of a study of how array orientation affects linear time-reversing array (TRA) retrofocusing in range-independent shallow ocean environments. Performance changes produced by variations in array orientation and element spacing have been presented. Although most of this discussion has focused on the simpler of the environments, three important conclusions can be drawn from this study.

First, the element spacing, length, and deployment depth of a horizontal TRA all play important roles in its performance in any multimode waveguide. The length of the array should be set to resolve the maximize number of propagating modes with the element spacing set to prevent spatial aliasing. When the horizontal wave numbers of the acoustic field from the original source are band limited, the acceptable element spacing may be much larger than the usual Nyquist limit. However, if spatial aliasing cannot be avoided, the element spacing might be made mildly non-uniform to mitigate adverse effects. Strategies for determining the best array deployment depth will require knowledge of the likely source depth and the propagating mode shapes in the environment. Thus, the optimum array deployment depth is likely to be set by the needs of the specific TRA application under consideration.

Second, when considering the vertical side lobe structure of the retrofocus region, lengthening a horizontal TRA or increasing its number of elements will not necessarily allow it to perform as well as a water-column-spanning vertical TRA in the same environment. Horizontal TRAs generate persistent side lobes that are not produced by vertical water-column-spanning TRAs. Such vertical TRAs sample and rebroadcast the field from the original source without additional weighting by the vertical structure of the propagating modes. Unfortunately, the additional weighting [the extra $\psi_m^2(z_a)$ factor in (5b) compared to (4)] is unavoidable for horizontal TRAs. However, when considering all three spatial dimensions, a vertical TRA does not provide any azimuthal retrofocusing in horizontally uniform environments while both types of horizontal arrays do. Interestingly, vertical arrays may overcome this problem in more complicated waveguides with range dependence since vertical TRAs can produce azimuthally confined retrofoci in inhomogeneous environments (Dowling and Jackson, 1992; Dungan and Dowling, 2001). Thus, vertical TRAs are likely to be advantageous in most shallow ocean environments. However, long horizontal TRAs can sample bottom propagated modes which might evade a purely water-column-deployed vertical array (Bogart and Yang, 1994), and horizontal arrays are more easily towed.

Third, even a small amount of vertical aperture may be beneficial in horizontal TRA side lobe level reduction. Figure 9 shows that as little as two wavelengths of vertical aperture prevents the TRA's elements from falling entirely into the nodes of the important propagating modes. Thus, horizontal TRAs should be slightly tilted or deployed with vertical offsets in the location of their elements. Furthermore, because TRAs can operate without knowledge of their element locations, unknown offsets—perhaps even those produced unintentionally by tow ship maneuvers, or ship-wake and cable turbulence—may be exploited for side lobe reduction.

And, finally, because of the analytical similarity of time reversal and matched field processing (MFP), many of the findings and conclusions of this study concerning horizontal TRAs are readily transferable to the realm of MFP with horizontal arrays.

ACKNOWLEDGMENTS

This work was sponsored by the Office of Naval Research Ocean Acoustics Program through Contract Nos. N00014-96-1-0040 and N00014-97-1-0628. In addition, the careful and thorough efforts of the reviewers of this manuscript are thoughtfully appreciated.

APPENDIX: A TRIGONOMETRIC IDENTITY

$$
\begin{aligned}
&\sin^2\left(\frac{m\pi}{2}\right)\sin\left(\frac{m\pi z_s}{D}\right)\sin\left(\frac{m\pi(D-z_s)}{D}\right)\\
&\quad=\sin^2\left(\frac{m\pi}{2}\right)\sin\left(\frac{m\pi z_s}{D}\right)\sin\left(m\pi-\frac{m\pi z_s}{D}\right)\\
&\quad=\sin^2\left(\frac{m\pi}{2}\right)\sin\left(\frac{m\pi z_s}{D}\right)\left[\sin(m\pi)\cos\left(\frac{m\pi z_s}{D}\right)\right.\\
&\qquad\left.-\cos(m\pi)\sin\left(\frac{m\pi z_s}{D}\right)\right]\\
&\quad=\sin^2\left(\frac{m\pi}{2}\right)[-\cos(m\pi)]\sin^2\left(\frac{m\pi z_s}{D}\right)\\
&\quad=\left(\frac{1-\cos(m\pi)}{2}\right)[-\cos(m\pi)]\sin^2\left(\frac{m\pi z_s}{D}\right)\\
&\quad=\left(\frac{-\cos(m\pi)+\cos^2(m\pi)}{2}\right)\sin^2\left(\frac{m\pi z_s}{D}\right)\\
&\quad=\left(\frac{-\cos(m\pi)+1}{2}\right)\sin^2\left(\frac{m\pi z_s}{D}\right)\\
&\quad=\sin^2\left(\frac{m\pi}{2}\right)\sin^2\left(\frac{m\pi z_s}{D}\right).
\end{aligned}
$$

Apel, J. R., Badiey, M., Chiu, C.-S., Finette, S., Headrick, R., Kemp, J., Lynch, J. F., Newhall, A., Orr, M. H., Pasewark, B. H., Tielbürger, D., Turgut, A., von der Heydt, K., and Wolf, S. (**1997**). ''An overview of the 1995 SWARM Shallow-water internal wave acoustic scattering experiment,'' IEEE J. Ocean. Eng. **22**, 465–500.

Bogart, C. W., and Yang, T. C. (**1994**). ''Source localization with horizontal arrays in shallow water: Spatial sampling and effective aperture,'' J. Acoust. Soc. Am. **96**, 1677–1686.

Collins, M. D. (**1993**). ''A split-step Padé solution for parabolic equation method,'' J. Acoust. Soc. Am. **93**, 1736–1742.

Collins, M. D. (**1994**). ''Generalization of the split-step Padé solution,'' J. Acoust. Soc. Am. **96**, 382–385.

Collins, M. D. (**1998**). ''New and improved parabolic equation models,'' J. Acoust. Soc. Am. **104**, 1808(A).

Dowling, D. R., and Jackson, D. R. (**1992**). ''Narrow-band performance of acoustic phase-conjugate arrays in dynamic random media,'' J. Acoust. Soc. Am. **91**, 3257–3277.

Dungan, M. R., and Dowling, D. R. (**2000**). ''Computed narrowband time-reversing array retrofocusing in a dynamic shallow ocean,'' J. Acoust. Soc. Am. **107**, 3101–3112.

Dungan, M. R., and Dowling, D. R. (**2001**). ''Computed narrowband azimuthal time-reversing array retrofocusing in shallow water,'' J. Acoust. Soc. Am. **110**, 1931–1942.

Hodgkiss, W. S., Song, H. C., Kuperman, W. A., Akal, T., Ferla, C., and Jackson, D. R. (**1999**). ''A long-range and variable focus phase-conjugation experiment in shallow water,'' J. Acoust. Soc. Am. **105**, 1597–1604.

Jackson, D. R., and Dowling, D. R. (**1991**). ''Phase-conjugation in underwater acoustics,'' J. Acoust. Soc. Am. **89**, 171–181.

Jensen, F. B., Kuperman, W. A., Porter, M. B., and Schmidt, H. (**1994**). *Computational Ocean Acoustics* (AIP, New York).

Khosla, S. R., and Dowling, D. R. (**1998**). ''Time-reversing array retrofocusing in simple dynamic underwater environments,'' J. Acoust. Soc. Am. **104**, 3339–3350.

Khosla, S. R., and Dowling, D. R. (**2001**). ''Time-reversing array retrofocusing in noisy environments,'' J. Acoust. Soc. Am. **109**, 538–546.

Kim, S., Edelmann, G. F., Kuperman, W. A., Hodgkiss, W. S., Song, H. C., and Akal, T. (**2001**). ''Spatial resolution of time-reversal arrays in shallow water,'' J. Acoust. Soc. Am. **110**, 820–829.

Kinsler, L. E., Fry, A. R., Coppens, A. B., and Sanders, J. V. (**2000**). *Fundamentals of Acoustics*, 4th ed. (Wiley, New York), pp. 259–266.

Kuperman, W. A., Hodgkiss, W. S., Song, H. C., Akal, T., Ferla, C., and Jackson, D. R. (**1998**). ''Phase-conjugation in the ocean: experimental demonstration of an acoustic time reversal mirror,'' J. Acoust. Soc. Am. **103**, 25–40.

Roux, P., and Fink, M. (**2000**). ''Time reversal in a waveguide: Study of the temporal and spatial focusing,'' J. Acoust. Soc. Am. **107**, 2418–2429.

Sabra, K. G., Khosla, S. R., and Dowling, D. R. (**2002**). ''Broadband time-reversing array retrofocusing in noisy environments,'' J. Acoust. Soc. Am. **111**, 823–830.

Song, H. C., Kuperman, W. A., and Hodgkiss, W. S. (**1998**). ''A time-reversal mirror with variable range focusing,'' J. Acoust. Soc. Am. **103**, 3234–3240.

Song, H. C., Kuperman, W. A., Hodgkiss, W. S., Akal, T., and Ferla, C. (**1999**). ''Iterative time reversal in the ocean,'' J. Acoust. Soc. Am. **105**, 3176–3184.

Tantum, S. L., and Nolte, L. W. (**2000**). ''On array design for matched-field processing,'' J. Acoust. Soc. Am. **107**, 2101–2111.

Tolstoy, A. (**1993**). *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore).

Weston, D. E. (**1980**). ''Wave-theory peaks in range-averaged channels of uniform sound velocity,'' J. Acoust. Soc. Am. **68**, 282–286.

Yang, T. C., and Yates, T. (**1998**). ''Matched-beam processing: Application to a horizontal line array in shallow water,'' J. Acoust. Soc. Am. **104**, 1316–1330.

Ziomek, L. J. (**1995**). *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing* (CRC, Boca Raton, FL).

An inversion for Biot parameters in water-saturated sand

Nicholas P. Chotiros[a)]
Applied Research Laboratories, The University of Texas at Austin, Austin, Texas 78713-8029

(Received 18 September 2000; revised 18 January 2002; accepted 4 February 2002)

The discrepancy between acoustic measurements and the theoretical predictions was investigated in the case of water-saturated sand. Two theoretical models were considered: visco-elastic and poro-elastic solid models. The visco-elastic solid model could not be reconciled with reflection loss measurements and was rejected. The poro-elastic solid model using Biot's theory [J. Acoust. Soc. Am. **103**, 2723–2729 (1998)] as formulated by Stoll [J. Acoust. Soc. Am. **70**, 149–156 (1981)] was an improvement. It was investigated using an inversion process. Operative values of grain bulk modulus and the frame bulk and shear moduli of water-saturated sand were inverted from simple measurements—reflection loss, compressional and shear wave speeds and attenuations. Although the inversion process is nonlinear, in practice, it is well behaved and converges quite rapidly to a unique solution. The issue of imprecisely known parameter values was handled in a probabilistic manner. The inversion results, using published laboratory and *in situ* measurements, showed that further improvement was needed. In an attempt to find a solution, two possible hypotheses are put forward. (1) Composite materials: The possibility that the frame may contain fluid and that the pore fluid may contain loose grains. (2) Independent coefficient of fluid content: The possibility that porosity may change with pore fluid pressure. Inversion results were encouraging for both hypotheses. It is difficult to say which of the two hypotheses is superior, and the two hypotheses are not mutually exclusive. The new hypotheses represent a significant advance because they have the potential to resolve the remaining discrepancies. At this stage, alternative interpretations of the data are possible. [DOI: 10.1121/1.1511199]

PACS numbers: 43.30.Ma, 43.30.Pc, 43.30.Vh [DLB]

I. INTRODUCTION

The elastic properties of a uniform elastic medium may be inverted from simple nondestructive measurements. For a fluid, the sound speed is simply related to the bulk modulus and the density. Given any two, it is possible to compute the third. The reflection coefficient at the boundary between two fluids is simply related to the ratios of the densities and sound speeds. For a uniform elastic solid, it is a little more complicated. In addition to the compressional wave, there is also a shear wave, and a corresponding shear modulus. In the visco-elastic medium model, wave attenuation is taken into account by using a complex phase speed. Measurements of compressional and shear wave speeds and attenuations and the normal incidence reflection loss would be sufficient to invert for the density and the elastic properties of any uniform elastic medium.

In the case of water-saturated sand, visco-elastic medium theory is not consistent with measurements. This result was found both in the laboratory[1,2] and *in situ*,[3–6] as shown in Fig. 1. The experimental results are shown as boxes that encompass the maximum and minimum values reported, and the limits of the frequency band. The laboratory results were obtained in well-controlled conditions. In Ref. 1, the sand was vibrated and achieved a porosity of 36%, while in Ref. 2 the sand was allowed to settle naturally and the porosity obtained was 41%. In both cases, great care was taken to remove gas from the water-saturated sand. The *in situ* results were obtained with varying degrees of precision. While there are numerous publications of subbottom profiling echograms, few were calibrated. References 3 and 4 are among the very few pre-1980 sources that contained absolute measures of reflection loss. Reference 5 contains measurements from a sand bed off Florida, and Ref. 6 was from a site near Elba, Italy. The visco-elastic wave theory prediction of reflection loss is similarly presented. A single box is used to represent the upper and lower bounds for all cases. The reflection loss is controlled by just two parameters: the density and sound speed ratios. The effects of shear wave and attenuation are negligible at normal incidence. The density ratio is a function of the specific density of quartz,[7] which ranges from 2.6 to 2.7, and the porosity, which ranges from 36% to 50%, depending on sorting.[8] At 100 kHz, measured *in situ* sound speed ratios[9] in water-saturated sand and water range from 1.10 to 1.16, for the corresponding range of porosity values. For simplicity, the upper and lower bounds of the visco-elastic theory prediction were computed, using these upper and lower bounds, assuming that the sound speed ratios at 100 kHz are applicable across the frequency band indicated. Although more precise theoretical predictions for individual experiments will reduce the uncertainty interval, that level of detail is not necessary to show that the experimental data are inconsistent with the model.

The limitations of the visco-elastic medium model, where water-saturated sand and other granular ocean sediments are concerned, generated renewed interest in the theory of sound propagation in poro-elastic media, originally put forward by Biot.[10] A poro-elastic medium is modeled as a porous and elastic solid with fluid-filled pore spaces. In the

a)Electronic mail: chotiros@arlut.utexas.edu

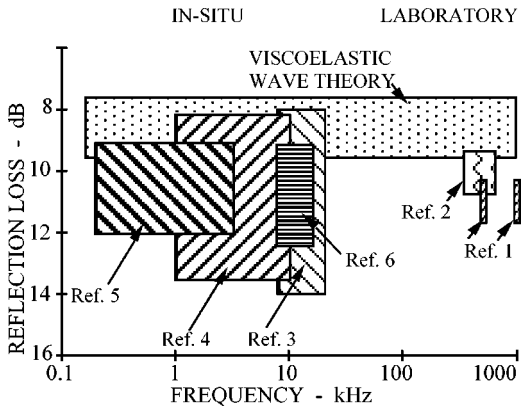


FIG. 1. Comparison of visco-elastic wave theory with published measurements of reflection loss at normal incidence.

case of water-saturated sand, the structure formed by the sand grains is collectively considered to be a contiguous porous elastic solid whose pore spaces are filled with water. Its acoustic response is represented by two coupled equations of motion, one for the solid frame and the other for the pore fluid. As formulated by Stoll,[11] the model requires thirteen input parameter values. The thirteen parameters may be divided into three groups: (1) properties of the component materials, (2) parameters related to fluid flow in the pore spaces, and (3) parameters describing the dynamic response of the frame. The parameters in the first group are well-known physical constants, such as the densities and elastic properties of the component materials. Parameters in the second group include the fluid viscosity, which can be accurately measured and is well tabulated for many fluids. The remaining parameters in this group are permeability, pore size, and the virtual mass constant, which may be measured or estimated from fluid dynamic principles, but with less precision. The parameters in the last group are the most elusive because they cannot be directly measured and difficult to estimate with any precision because the physics involved is not understood. Their values have to be adjusted to fit the measured compressional and shear wave speeds and attenuations. In the case of water-saturated sand, there were a number of studies[12–15] in the 1980s which converged on parameter values similar to those of Stoll and Kan,[16] as shown in the first column of Table I. For brevity, this parameter set will be referred to as the Stoll parameter set.

Although the Stoll parameter set gives a value of reflection loss that is similar to that obtained by direct measurement, they do so with a porosity value that is significantly higher than that of typical water-saturated sands, i.e., 0.47 against typical values that are closer to 0.4. For example, with a porosity of 0.39 measured from a sand bed near Panama City, FL, the model predictions are not consistent with the experimentally measured combination of wave speeds and reflection loss values. In 1995, Chotiros[17] argued that the value of the grain bulk modulus should be adjusted for two possible reasons: (1) Micrographs of sand grains[18] revealed numerous cracks and imperfections which suggest that sand grains may be weaker and hence more compressible than perfect quartz crystals. (2) The actual parameter called for in Biot's equations is the ''unjacketed'' bulk modulus of the frame,[19] which may be equated to the grain bulk modulus only under the assumption of an elastic and isotropic frame structure made entirely of the grain material. Typical parameter values are displayed in the second column of Table I. Using a value of only 7 GPa, instead of the tabulated bulk modulus of quartz crystals of 36 GPa, it is possible to obtain agreement with the measured combination of wave speeds and reflection loss values.

TABLE I. Biot parameter values for water-saturated sand in the literature.

	Units	Stoll, Kan (1981) Sand	Chotiros (1995) PCE
Bulk parameters			
1. Porosity	···	0.47	0.39
2. Grain density	kg/m^3	2650	2650
3. Liquid density	kg/m^3	1000	1000
4. Grain bulk modulus	GPa	36	7
5. Liquid bulk modulus	GPa	2.00	2.25
Flow parameters			
6. Viscosity	kg/m-s	0.001	0.001
7. Permeability	μm^2	100	50
8. Pore size	μm	10	50
9. Virtual mass constant	···	1.25	1.78
Frame response			
10. Frame shear modulus	GPa	0.0261	0.018
11. Shear log decrement	···	0.15	0.15
12. Frame bulk modulus	GPa	0.0436	5.230
13. Bulk log decrement	···	0.15	0.15
Biot model at 100 kHz			
Sed. Fast wave speed (c_1)	m/s	1600	1710
Sed. shear wave speed (c_s)	m/s	132	100
Sed. fast wave att. (α_1)	dB/m	4.1	70
Sed. shear wave att. (α_s)	dB/m	1040	1460
Reflection loss (R_n)	dB	11.00	11.00
Equivalent elastic model			
Density ratio	···	1.87	2.01
Velocity ratio	-	1.14	1.14
Reflection loss	dB	8.82	8.17

The Chotiros parameter set has led to serious discussions[20,21] for a number of reasons. One reason is that they predict a slow wave speed that is much faster than that of the Stoll values, and which could have explained certain experimental observations, in particular, the penetration of sound into water-saturated sand at angles beyond the critical angle for the fast wave. It has also attracted more extensive measurements of the physical and acoustic properties of sand. A recent acoustic measurement of the bulk modulus of sand grains[22] has found its value to be very close to 36 GPa, which removes the first possible justification of the Chotiros approach.

In this study, it will be shown that the procedures and the subsequent parameter values of neither Stoll nor Chotiros are correct for water-saturated sand. The adopted approach is based on a new procedure to invert acoustic measurements for the Biot model parameters, particularly the grain and frame elastic moduli. In Sec. II, the inversion procedure will be described. Its convergence will be demonstrated with both the Stoll and Chotiros sets of parameter values. Then, it will be applied to published experimental measurements, both

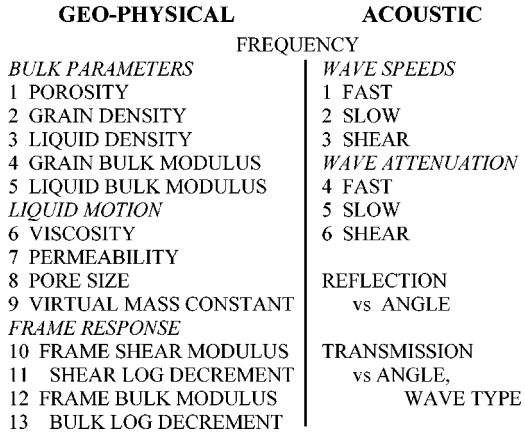


FIG. 2. The connection between geophysical and acoustic parameters.

from the laboratory and *in situ*, from which it will be concluded that the approach of neither Stoll nor Chotiros was entirely correct. A couple of hypotheses will be put forward in an effort to explain the discrepancy between theory and experiment. The results for both hypotheses will be described.

II. CONNECTION BETWEEN ACOUSTIC AND PHYSICAL PROPERTIES

In Stoll's formulation, the solutions for simple harmonic waves in a poro-elastic medium are completely determined by the acoustic frequency and thirteen input parameters. They are explicitly listed in Table I. For want of a better term, they will be collectively referred to as the geophysical parameters. All acoustic properties of the medium, such as wave propagation speed, attenuation, and reflection coefficient, are completely determined by these parameters. The term ''acoustic properties'' in this case is taken to include properties of the shear wave as well. While there are thirteen independent parameters on the geophysical side of the acoustic-geophysical boundary, it is not clear how many there are on the acoustic side. The readily measurable acoustic parameters include speeds and attenuations of the fast and slow compressional and shear waves, and the reflection and transmission coefficients at the water-sand boundary, as illustrated in Fig. 2. The Biot–Stoll equations define the relationships between the parameters across the acoustic-geophysical boundary. The equations may be found in Stoll and Kan.[16]

Due to the complexity of the equations, it is not always clear how any particular parameter on one side of the boundary affects another parameter on the opposite side. A numerical approach will be taken to circumvent some of the complexities of the analytical expressions. As an illustration, the sensitivities of certain readily measurable acoustic parameters to the values of the geophysical parameters were computed, at a frequency of 100 kHz, using values from the Stoll parameter set. The results are shown in Table II, in the form of a matrix of dimensionless partial derivatives. Each element is defined as the fractional change of an acoustic parameter, y, per unit fractional change of a geophysical parameter, x. A zero value indicates the absence of any connection. Conversely, a nonzero value is indicative of a connection, and the magnitude is a measure of the sensitivity. The sign is indicative of the sense of the correlation, i.e., positive or negative correlation. The acoustic parameters were chosen for their simplicity. They are the shear, fast and slow wave speeds and attenuations and the magnitude of the normal incidence reflection coefficient. Reflection loss is defined as the negative of the reflection coefficient magnitude in decibels.

For the shear wave, it is evident that the wave speed c_s is dependent mainly on the grain density, ρ_r, the real part of frame shear modulus μ_0, and porosity β, for which the dimensionless partial derivatives all have magnitudes greater than 0.1. For example, a negative 0.41% change in the shear wave speed, c_s, is indicated as a response to a 1% increase in the porosity β. One can say that it is slightly dependent on

TABLE II. Partial derivatives of acoustic parameters with respect to changes in each geophysical parameter.

$(\partial y/\partial x)(x/y)$	y Shear wave speed	Fast wave speed	Slow wave speed	Shear wave atten.	Fast wave atten.	Slow wave atten.	Normal reflect. coeff.
x Parameter	c_s	c_1	c_2	α_s	α_1	α_2	R_n
Bulk properties							
1. β, porosity	0.41	−0.22	0.69	−0.23	1.39	−0.39	−1.04
2. ρ_r, grain density	−0.44	−0.18	−0.26	0.41	1.56	0.17	0.29
3. ρ_f, fluid density	−0.03	−0.28	−0.22	0.05	−1.44	0.20	−0.30
4. K_r, grain bulk mod.	0	0.03	$-5.0E-4$	0.00	−0.01	$-2.6E-4$	0.04
5. K_f, fluid bulk mod.	0	0.45	$3.7E-3$	$1.2E-6$	−0.54	$9.3E-4$	0.72
Fluid motion							
6. η, fluid viscosity	$-4.5E-4$	$-3.8E-4$	$-1.8E-3$	0.03	0.52	0.12	$6.1E-4$
7. κ, permeability	$9.1E-4$	$7.1E-4$	$3.8E-3$	−0.05	−0.77	−0.18	$-1.3E-3$
8. a, pore size	$-1.1E-3$	$-8.6E-4$	$-4.5E-3$	0.04	0.63	0.15	$1.5E-3$
9. c, virtual mass	−0.11	−0.08	−0.49	0.01	−1.47	0.31	0.15
Frame response							
10. μ_0, frame shear mod	0.49	$1.9E-3$	0.22	−0.47	0.05	−0.22	$2.3E-3$
11. δ_μ, shear log dec.	$1.8E-3$	$-1.8E-5$	$8.3E-4$	0.94	0.07	0.36	$-2.6E-5$
12. K_b, frame bulk mod.	0	$1.6E-3$	0.27	0	0.05	−0.27	$1.9E-3$
13. δ, bulk log dec.	0	$-1.8E-5$	$1.1E-3$	0	0.06	0.45	$-2.7E-5$

pore size a, fluid viscosity η, permeability κ, and shear log decrement δ_μ, since the magnitudes of these partial derivatives are less than 0.1 but greater than zero. Finally, it appears to be independent of the remaining parameters. On the basis of these values, it may be concluded that a reasonably accurate estimate of the frame shear modulus may be made from measurements of just the shear wave speed, porosity and grain density. The shear wave attenuation shows similar dependencies, but in addition, it shows strong sensitivities to the fluid viscosity η, permeability κ, and shear log decrement δ_μ.

The fast compressional wave speed, c_1, appears to be mainly sensitive to fluid density, ρ_f, porosity, β, grain density, ρ_r, and fluid bulk modulus, K_f, for which the dimensionless partial derivatives all have magnitudes greater than 0.1. It is slightly sensitive to the grain bulk modulus, K_r, and the virtual mass constant, c, and faintly sensitive to all of the remaining parameters. The attenuation shows similar tendencies, in addition to strong sensitivities to the virtual mass c, fluid viscosity η, permeability κ, and bulk log decrement δ.

The magnitude of the reflection coefficient appears to be sensitive to the same parameters as the compressional wave speed, and this is not unexpected since the reflection coefficient of an elastic medium is directly related to the acoustic impedance, which is the product of density and sound speed. However, the sense and magnitude of the dependencies are different. There is a negative correlation between porosity and reflection loss, while between porosity and fast wave speed the correlation is positive.

The slow wave speed, in contrast, is mainly sensitive to the real parts of the bulk K_b and shear μ_0 moduli of the frame, and the virtual mass term c, in addition to the porosity, β, and grain density, ρ_r. It is faintly dependent on all of the remaining parameters. The slow wave attenuation has similar tendencies in addition to strong sensitivities to fluid viscosity η, permeability κ, and both log decrements, δ and δ_μ.

III. INVERSION

A process is sought by which bulk and shear moduli of a poro-elastic medium may be inverted from acoustic measurements. It is assumed that the medium is isotropic and homogeneous. The target parameters are the grain and frame elastic moduli and the log decrements. Before the inversion can be accomplished, values of the remaining geophysical parameters need to be assigned, or at least bounded within manageable limits. With reference to the parameter list, the densities of the solid and fluid components are easily measured. The bulk modulus and viscosity of the fluid, water in this case, are well tabulated. The porosity is also measurable using standard laboratory methods. If the fast and shear wave speeds and attenuations, and the reflection loss are chosen as the acoustic measurements, then only approximate values of the virtual mass term, permeability, and pore size are needed, since the selected acoustic properties are only slightly sensitive to them, with a few exceptions. The exceptions concern the attenuation coefficients, which have significant dependencies on the virtual mass and the permeability. However, in the case of a medium that is not very lossy, this would only affect the values of the log decrements.

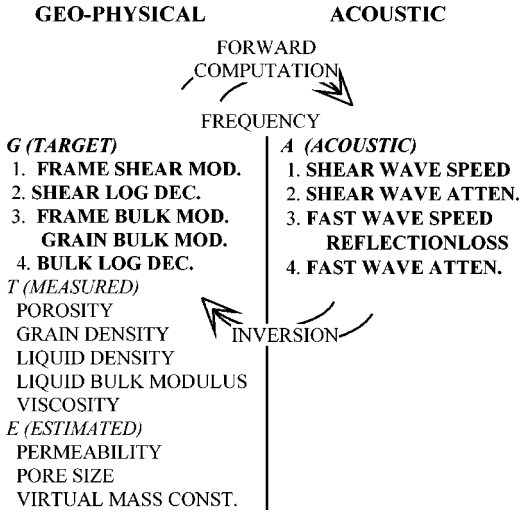


FIG. 3. Inversion of five geophysical parameters from five acoustic parameters.

The process is essentially one of connecting a vector, A, consisting of the five acoustic measurements, to a vector, G, of the five target geophysical unknowns, given a vector of measurable physical constants, T, and a fourth vector, E, of the parameters that are not precisely quantifiable, but could be estimated with a certain level of confidence. The five target geophysical unknowns are, specifically, the grain bulk modulus K_r, the frame bulk K_b and shear μ_0 moduli, and the frame log decrements, δ and δ_μ. The five acoustic measurements in A are the reflection loss at normal incidence, and the speeds and attenuation coefficients of the fast and shear waves. In the forward direction, the relationship may be expressed as

$$A = F(G,T,E). \tag{1}$$

In the inverse direction, the expression is

$$G = F^{-1}(A)\big|_{T,E}. \tag{2}$$

The vectors and their component parameters are illustrated in Fig. 3. In practice, it is not necessary to invert for all five parameters in G simultaneously. The shear wave properties are somewhat separable from those of the fast wave and reflection loss at normal incidence, and the attenuation coefficients are only weakly coupled to the wave speed values. Therefore, the process was divided into four steps. (1) A value of the frame shear modulus was inverted from the shear wave speed, and then (2) a value of the shear log decrement from the shear wave attenuation. (3) The grain and frame bulk moduli were simultaneously inverted from the fast compressional wave speed and the reflection loss. This is the only part that truly required a simultaneous solution. (4) Finally, the frame bulk log decrement was inverted from the fast wave attenuation. At each step, Newton's method was used to invert for the relevant parameter value in an iterative manner. This involved running the forward process to compute one or more partial derivatives, followed by a simple linear interpolation to compute the next best esti-

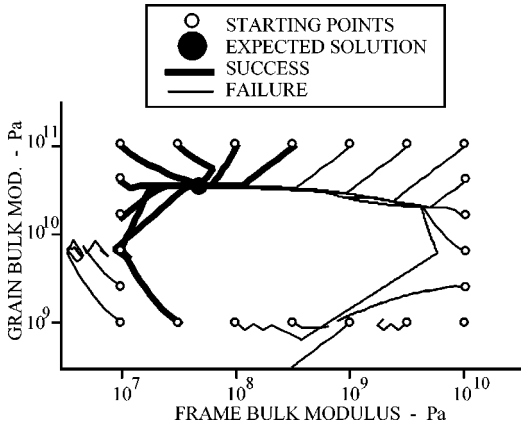


FIG. 4. Convergence test result using the Stoll and Kan parameter set.

mate. The four steps were repeated iteratively until a stable solution was achieved, to allow the weak connections between all the parameters involved to work themselves out.

IV. CONVERGENCE

The parameter values of Stoll and Chotiros, as shown in Table I, were used to verify the convergence of the inversion algorithm. As a test, computed values of reflection loss, fast and shear wave speeds and attenuations were fed into the inversion algorithm. The starting values of the shear modulus and the log decrements did not have much effect on the convergence of the process. Of particular interest was the convergence performance as a function of the starting values of the grain and frame bulk moduli since this was the most complicated part of the process. The solution space was expected to be bounded between 0.01 and 10 GPa for the frame bulk modulus, and between 1 and 100 GPa for the grain bulk modulus. To test convergence, a set of discrete starting points were evenly distributed around the periphery of the search space. The inversion process was started from each of the starting points. The results for the Stoll and Kan parameter set are shown in Fig. 4. About half of the inversions converged on the correct solution. The others either failed the convergence test or exceeded the number of iterations allowed. No other convergence points were found. The results for the Chotiros parameter set are very similar, as shown in Fig. 5. It appears that starting points above and left of the solution are most likely to converge. The inversion strategy adopted was to start at the top left-hand corner of the search space. If convergence was not achieved, then each of the remaining three corners were tried. This guaranteed convergence if a solution existed at all.

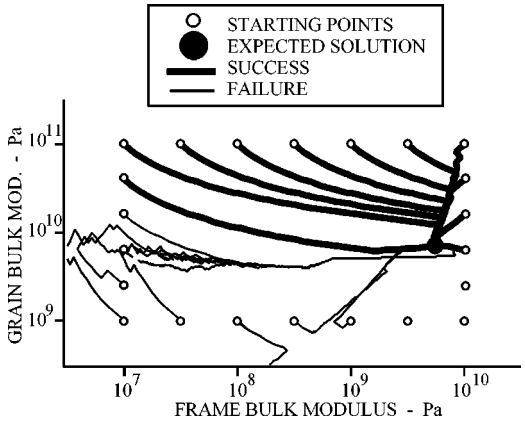


FIG. 5. Convergence test result using the Chotiros parameter set.

V. SENSITIVITY TO UNCERTAINTIES IN FLOW PARAMETERS

Sensitivity of the inversion result to the assumed values of the flow parameters, particularly permeability κ, pore size a, and virtual mass c, were of concern because these values are difficult to know accurately. Although there are methods to measure permeability, in many cases the accuracy of the measurements is still being debated. It is also questionable if a value of permeability, measured in steady flow experiment, could be accurately applied to a high-frequency oscillating flow situation. There are similar uncertainties associated with the pore size parameter and the virtual mass constant. The approach adopted is to treat these parameters as random variables. Each one is assumed to follow a log-normal probability distribution function (PDF), parametrized in terms of a mean and a standard deviation. In the case of the virtual mass constant, $(c-1)$ was chosen to be the log-normal random process, rather than c, because the value of c is not permitted to be less than one.

The log-normal PDFs were centered about the best estimates of permeability κ_0, pore size a_0, and virtual mass excess c_0. Regarding permeability, Taylor-Smith[23] has compared a number of measurement methods in order to investigate their relative accuracies and found discrepancies approaching two orders of magnitude. Therefore, the permeability was given a log-normal PDF with a range of two orders of magnitude between the ± 1 s.d. limits. Since the Kozeny–Carman equation[24] indicates that the pore size parameter is proportional to the square root of the permeability, it would follow that its standard deviation should be half that of the permeability. As for the virtual mass constant, its minimum value is unity. Stoll[11] indicated that its maximum value may be as high as 3. Therefore, the standard deviation was set such that the value of c would be 3 at $+1$ s.d. At -1 s.d., the value of c would be just slightly greater than one. The PDF curves are shown in Fig. 6.

Convolving the inversion process with the relevant PDFs would give an expected value of the target vector, G,

$$\langle G\rangle = \int_{-\infty}^{\infty}\int_{\infty}^{\infty}\int_{\infty}^{\infty} G(\kappa,a,c)p_{\kappa}(x_{\kappa})p_a(x_a)p_c(x_c) \times dx_{\kappa}dx_{\alpha}dx_c, \tag{3}$$

where $x_{\kappa}=\ln(\kappa)$, $x_a=\ln(a)$, and $x_c=\ln(c-1)$, and the functions $p_{\kappa}(x_{\kappa})$, $p_a(x_{\alpha})$, and $p_c(x_c)$ are the respective PDFs.

Since the inversion, G, is the result of a numerical process that cannot be analytically integrated, it is necessary to numerically compute the integral. One approach is to discretize the PDF into a number of equal probability intervals,

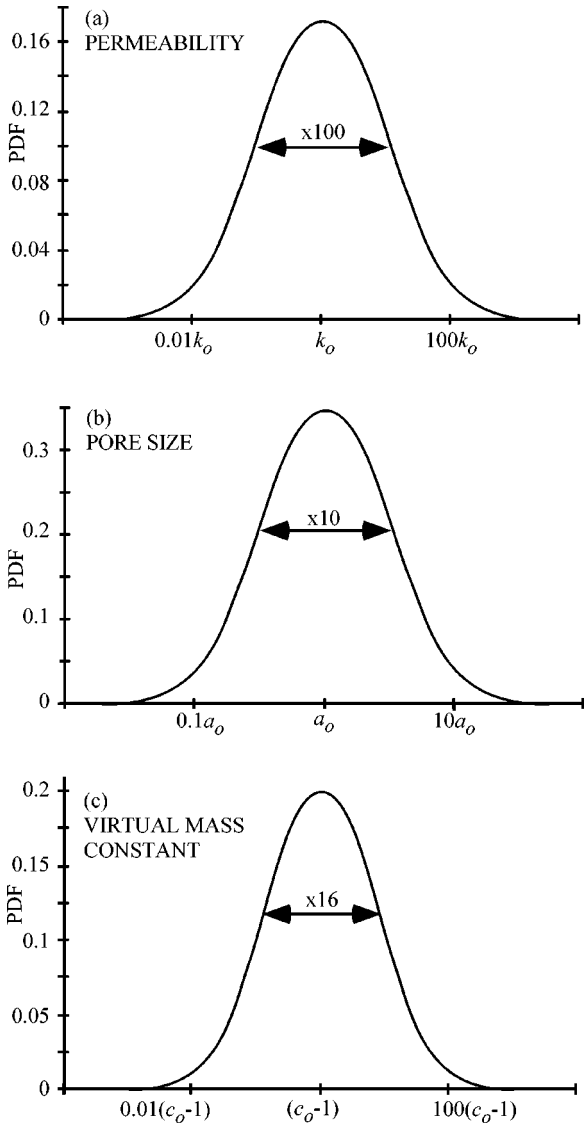


FIG. 6. Log-normal probability density functions of (a) permeability, (b) pore size, and (c) virtual mass constant.

as illustrated in Fig. 7. Thus, the convolution integrals may be computed as nested summations to obtain an approximate numerical solution,

$$\langle G\rangle \approx N^{-3}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} G(\kappa_i, a_j, c_k), \tag{4}$$

where $x_{\kappa i}=\ln(\kappa_i)$, $x_{aj}=\ln(a_j)$, and $x_{ck}=\ln(c_k-1)$, in which the discrete samples $x_{\kappa i}$, x_{aj}, and x_{ck} were chosen at equal probability intervals.

One disadvantage of numerical sampling is that it is impossible to cover all possible combinations of values of the random processes. For $N=9$, the distance between the outermost samples of each random process covers 0.889 of the total probability. For a combination of three random processes, there are 729 ($=N^3$) samples to compute for a coverage of 0.70 of the total probability.

Rather than simply computing the expected value, it is instructive to look at the individual inversion results, G, within the summation. First of all, not all of the inversions were successful. Using the Stoll values from Table I and for

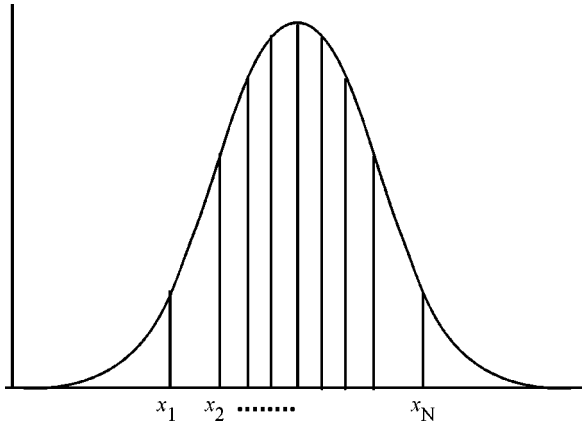


FIG. 7. Discrete values at equal intervals of probability.

$N=9$, 40% of the inversions were successful. The values of grain and frame bulk moduli of the successful inversions are plotted in Fig. 8(a). The target values, as listed in Table I, are at the point labeled ''Expected value.'' Ideally, the effects of the random variations in permeability, pore size, and virtual mass should have averaged out, leaving an unbiased set of average values in close proximity to the target values. As it turned out, the average values were quite far removed from the target values. It was observed that, for several cases where the virtual mass, c, was less than 1.25, or where the pore size, a, was greater than 20 μm, the inversions were unsuccessful. This was not due to a failure in the inversion process, but because certain parameter sets containing these values of a and c were incompatible with the desired acoustic reflection loss and wave speeds. In most cases, the inversion process failed to converge at step (3), the simultaneous inversion of the grain and frame bulk moduli. The result is a

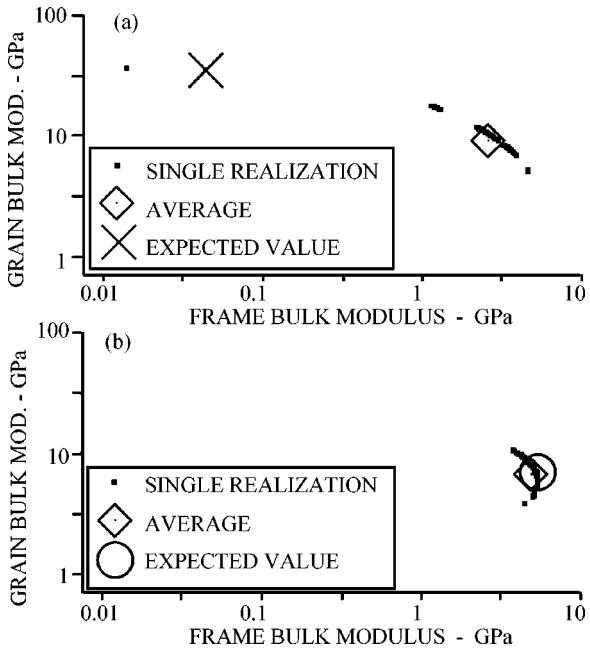


FIG. 8. Inversion of grain and frame bulk moduli with probabilistic variations in permeability, pore size, and virtual mass constant, using (a) Stoll and Kan parameters and (b) Chotiros parameters.

TABLE III. Inversion results using model test data.

		Stoll, Kan (1981)		Chotiros (1995)	
Invertable fraction		0.40		0.79	
Max. failure point		3		3	
	Units	Avg. value	s.d. Np	Avg. value	s.d. Np
Grain bulk mod. (K_r)	GPa	10.30	0.39	6.99	0.24
Frame bulk mod. (K_b)	GPa	2.96	0.35	4.92	0.09
Frame shear mod. (μ)	GPa	0.029	0.05	0.018	0.06
Bulk log dec. (δ)	⋯	0.03	1.68	0.18	0.61
Shear log dec. (δ_μ)	⋯	0.15	0.15	0.14	0.21

biased average. The successful inversions were spread along a parabolic curve that extended over an order of magnitude in both directions. A list of the inverted parameter values and their standard errors are shown in the first two columns of Table III.

Using the Chotiros parameter set, the inversion results were more robust. A larger proportion of the inversions, 79%, were successful. The average results were very close to the target values. Most of the unsuccessful inversions involved virtual mass values greater than 10, which are extremely unlikely values. The values of grain and frame bulk moduli of the successful inversions are shown in Fig. 8(b). The points were well clustered in the cusp of a parabolic curve, and the average values fell very close to the target values. In this case, the uncertainties in the values of the flow parameters did not prevent accurate inversion of the target parameters. A list of the inverted parameter values and their standard errors are shown in the last two columns of Table III.

VI. INVERSION OF EXPERIMENTAL DATA USING THE STOLL FORMULATION

Two laboratory data sets will be considered. The first is that of Nolle *et al.*,[1] which contains measurements of permeability and acoustic properties of graded sands at four different grain sizes, 640, 400, 170, and 120 μm, and at two frequencies, 200 and 500 kHz. For brevity, only the two extreme cases of wavelength to grain size ratio will be considered, i.e., 120 μm at 200 kHz and 640 μm at 500 kHz. The second is that of Drevet, Brussieux, and Sessarego,[2] using 250 μm sand at 500 kHz, and containing measurements of permeability and tortuosity, in addition to the acoustic properties. Drevet did not actually measure the normal incidence reflection loss, but the reflection loss over a range of incidence angles from 20° upward. From what is known of reflection loss trends, it is safe to say that the value at 20° is identical to the value at normal incidence to within a small fraction of a decibel. None of the data sets included shear speed measurements.

In addition, one *in situ* data set will be considered. It was collected as part of the Coastal Benthic Boundary Layer Program (CBBL),[25] in which a set of measurements were made in August 1993 on the West Florida sand sheet, approximately 43 km (23 nmi) southeast of Panama City, FL. The sand was mostly deposited during lower sea stands, with more recent deposits of silt and clay left by severe storms. The data set includes *in situ* compressional and shear wave speed measurements in addition to reflection loss measurements at normal incidence. For practical reasons, these measurements were made at different frequencies. The reflection loss was measured at 12 kHz, the compressional wave at 58 kHz, and the shear wave at 2 kHz. The inversion process was modified to take these differences into account. The parameter values are shown in Table IV.

Since these are experimental measurements, the error associated with each measurement needed to be quantified. An effort was made to estimate the standard error of each parameter value, assuming a log-normal distribution. The basic bulk properties, including densities, the bulk modulus and viscosity of fresh water, and the porosity of the sand have very small standard errors, i.e., less than 0.01 Np (less than 1%). With regard to the flow-related parameters, i.e., permeability, pore size, and virtual mass constant, it was more difficult. The flow parameters, which were not measured, but estimated from the grain size statistics according to the procedure in Chotiros,[17] were assigned a standard error of 0.7 Np, which corresponds to a multiplication or division by 2. A direct measurement was considered to be superior to an estimate based on theoretical expressions, and these were assigned a standard deviation of 0.4 Np, which corresponds to a multiplication or division by a factor of 1.5. This standard error was assigned to the measured values of permeability in all the laboratory data, and the measured virtual mass constant of the Drevet sand.

The grain properties are fundamental. For quartz sand, they are assumed to be equal to those of quartz crystals. There are a few sources from which this information may be obtained. The *CRC Handbook*[26] gave the value of bulk modulus as 36.6 GPa (5.3×10^6 p.s.i.) with a Poisson's ratio of 0.16 and a density of only 2200 kg/m^3. From the *Manual of Applied Geology for Engineers*,[7] the density was given as 2650 kg/m^3 with a pressure wave speed range between 5600 and 6300 m/s. Assuming the above-mentioned value of Poisson's ratio, this translates to a bulk modulus of between 38 and 48 GPa. In the *Reference Data for Radio Engineers*[27] the density was stated as 2650 kg/m^3, and two sound speeds were given, 5440 and 5720 m/s. The resulting values of bulk modulus are 36 and 40 GPa. From these references, it appears that the density is 2650 kg/m^3; the value in the CRC Handbook is clearly erroneous. The bulk modulus is between 36 and 48 GPa. It is not necessary to know the bulk modulus very accurately, because the acoustic properties are only weakly dependent on it when its value lies in this range, as indicated by the values of the partial derivatives shown in Table II. Following Stoll's lead, the bulk modulus will be assigned a value of 36 GPa, except in the case of the Drevet data where the authors have chosen a higher value of 42.7 GPa.

The accuracy of the acoustic measurements was also quantified. The fast wave speed was measured in all cases. Wave speeds measured in the laboratory are very accurate, and typically have a standard error of less than 0.01 Np. In the *in situ* data set, the standard error was computed from the spread of the measured values and it came to almost exactly

TABLE IV. Experimental data sets for inversion.

Data set		Nolle	Nolle		Drevet		CBBL	
		120	640		250		559	
Grain diameter	Units	Avg. value	Avg. value	s.d. Np	Avg. value	s.d. Np	Avg. value	s.d. Np
Mean	μm	120	640		250		559	
	ϕ	3.06	0.62		2.00		0.84	
Standard deviation	Np	0.26	0.13		0.49		0.44	
Bulk parameters								
Porosity	⋯	0.37	0.36	<0.01	0.41	<0.01	0.40	<0.01
Grain density	kg/m^3	2650	2650	<0.01	2650	<0.01	2680	<0.01
Grain bulk modulus	GPa	36.00	36.00		42.70		36.00	
Liquid density	kg/m^3	1000	1000	<0.01	1000	<0.01	1020	<0.01
Liquid bulk modulus	GPa	2.25	2.25	<0.01	2.25	<0.01	2.38	<0.01
Liquid motion								
Viscosity	kg/m s	0.001	0.001	<0.01	0.001	<0.01	0.001	<0.01
Permeability	μm^2	7	270	0.40	16	0.40	451	0.70
Pore size	μm	25	130	0.70	58	0.70	150	0.70
Virtual mass constant	⋯	1.25	1.85	0.70	1.31	0.40	1.75	0.70
Frequency	kHz	200	500		500		58	
Water sound speed (c_w)	m/s	1500	1500	<0.01	1500	<0.01	1527	<0.01
Sed. fast wave speed (c_1)	m/s	1740	1740	<0.01	1690	<0.01	1711	0.01
Sed. fast wave att. (α_1)	dB/m	60	18	0.40	180	0.40	30	0.24
Frequency	kHz	200	500		500		2	
Sed. shear wave speed (c_s)	m/s	100	100	0.70	100	0.70	120	0.14
Sed. shear wave att. (α_s)	dB/m	3000	7500	0.70	7500	0.70	30	0.70
Frequency	kHz	200	500		500		12	
|Reflection coeff.| (R_n)	⋯	0.282	0.282	0.07	0.318	0.07	0.269	0.18
Reflection loss	dB	11.00	11.00		9.94		11.39	

0.01 Np. The measurement of attenuation was considered to be less accurate. The standard error of the laboratory measurements of the compressional wave attenuation was assigned a rather conservative value of 0.4 Np. The standard error of the *in situ* measurement, 0.24 Np, was computed from the spread in the measured values. The highest and lowest measured values of *in situ* shear wave speed were reported by Richardson,[25] but the standard error was not provided. As a conservative estimate, the standard error was made to be equal to half the spread between the lowest and highest values. Since there were no measurements of shear wave speed and attenuation in the laboratory data, a typical value of 100 m/s and an attenuation of 15 dB/m/kHz were assigned, along with a conservative standard error of 0.7 Np. The standard error of the reflection loss was estimated from the spread of the measured values. The parameter values and the corresponding standard errors are listed in Table IV.

The results of the inversion are very sensitive to the accuracy of the reflection measurement. Significant scattering effects are expected to be present in both the laboratory and *in situ* data. In the laboratory data, at a rms surface roughness of one grain-radius, the Rayleigh coherent reflection coefficient is expected to be approximately 1 dB lower than that of a perfectly flat surface. For the *in situ* data collected at lower frequencies, a similar scattering process due to sand ripples occurs. The Rayleigh coherent reflection coefficient is applicable to a continuous sinusoidal signal. A pulse was used in the experiments for practical reasons, i.e., to avoid unwanted signal paths. No details were given by Nolle *et al.* of the exact method used to compute reflection loss. In the Drevet *et al.* and in the CBBL experiments, the reflection loss computation was based on the ratio of the fast Fourier transform of reflected and transmitted pulses at each resolvable frequency. This is equivalent to taking the ratio of the time integrals of the signal intensities at each frequency. Due to scattering, the reflected signal amplitude would be reduced, but its pulse duration would be increased. For a weakly scattered signal, the reduction in amplitude is roughly balanced by the increase in duration. The energy scattered out of the receiver beam is approximately balanced by energy scattered into the beam from other regions. The cancellation of the scattering loss has been studied and it is expected to work well near normal incidence.[28]

All parameters with standard errors $\leqslant$0.01 Np were considered to be exact for practical purposes. The remaining seven parameters were treated as random processes. They included permeability, pore size, virtual mass constant, reflection loss at normal incidence, fast wave attenuation, and shear wave speed and attenuation. The average inversion result was formulated as nested convolution integrals involving the corresponding PDFs,

$$\langle G\rangle = \int_{-\infty}^{\infty}\int_{\infty}^{\infty}\int_{\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\infty}^{\infty}\int_{\infty}^{\infty}\int_{\infty}^{\infty} G(\kappa,a,c,R_n,\alpha_1,c_s,\alpha_s) \times p_\kappa(x_\kappa)p_a(x_a)p_c(x_c)p_R(x_R)p_{\alpha 1}(x_{\alpha 1})p_{cs}(x_{cs}) \times p_{\alpha s}(x_{\alpha s})dx_\kappa dx_\alpha dx_c dx_R dx_{\alpha 1} dx_{cs} dx_{\alpha s}, \quad (5)$$

where $x_{Rn}=\ln(R_n)$, $x_{\alpha 1}=\ln(\alpha_1)$, $x_{cs}=\ln(c_s)$, and $x_{\alpha s}=\ln(\alpha_s)$; $p_{Rn}(x_{Rn})$, $p_{c1}(x_{c1})$, $p_{cs}(x_{cs})$, and $p_{\alpha s}(x_{\alpha s})$ are the corresponding PDFs.

As before, the integrals were discretized for numerical computation. In this case the chosen value of N was 5, simply to keep the computation load within reasonable limits.

TABLE V. Maximum likelihood inversion results from experimental data: variable grain bulk modulus.

Data set		Nolle 120	Nolle 640	Drevet 250	CBBL 559
Invertable fraction	···	0.34	0.31	0.65	0.53
Max. failure point	···	4	4	3	3
Grain bulk mod. (K_r)	GPa	8.51	8.63	16.78	7.52
Frame bulk mod. (K_b)	GPa	5.06	5.05	2.52	4.47
Frame shear mod. (μ)	GPa	0.037	0.019	0.037	0.027
Bulk log dec. (δ)	···	0.09	0.01	0.40	0.22
Shear log dec. (δ_μ)	···	0.21	0.16	0.22	0.23
Frequency	kHz	200	500	500	58
Slow wave speed	m/s	732	1328	695	1073
Slow wave atten.	dB/m	1522	168	4637	53

Since the number of samples goes as N,[7] the number of samples came to 78 125. For each random process, the coverage was 80% of the total probability space between the outermost samples. For the combination of all seven random processes, the total coverage was just 21% ($0.8^7 \times 100\%$), despite the large number of samples.

The statistics of the inversion results are shown in Table V. The average parameter values from all the successful inversions, such as those in Table III, when applied to the forward model, do not actually produce the correct acoustic parameter results. Due to the complicated relationship between the geophysical and acoustic properties, the average values of successful inversions are not themselves the product of a successful inversion. For example, the average point shown in Fig. 8, does not actually lie upon the curve of successful inversions, and if applied to the forward model, will not produce the correct acoustic parameter values. Therefore, the average values were rejected in favor of the maximum likelihood (ML) values, as shown in Table V and subsequent tables. The ML values are defined as the inversion results obtained at the highest value of the likelihood ratio, which was taken to be the product of the PDFs in Eq. (5).

With the Nolle data, a relatively small proportion of the inversions was successful. With the Drevet and the CBBL data sets, the proportion of successful inversions was higher. In all cases, most of the successful inversions appeared to correspond with lower values of reflection loss and higher values of fast and shear wave attenuations. The plots of grain and frame bulk moduli are shown in Fig. 9. The Nolle data gave results that clustered very tightly about the Chotiros values. The Drevet and CBBL data gave results that were spread over a larger range of values. The values were closer to the Chotiros values than those of Stoll.

From these results, it would appear that the Stoll values are inconsistent with the experimental data, while the Chotiros values, which feature a value of grain bulk modulus that is considerably lower than that of quartz crystals, are more appropriate. Chotiros[21] had argued that the grain bulk modulus of sand might differ from that of quartz crystals due to fractures and other imperfections. However, a recent measurement[22] of the grain bulk modulus of quartz sand grains at ultrasonic frequencies showed that the value of grain bulk modulus was considerably higher than the value used by Chotiros and very close to that of quartz crystals as originally stated by Stoll.[16] The situation may be summed up very simply. On the one hand, it is found that the proper value of grain bulk modulus is incompatible with experimental data, but on the other hand, a compatible value of grain bulk modulus can be found but it is unacceptable. A similar result was found in the inversion of grain and frame bulk

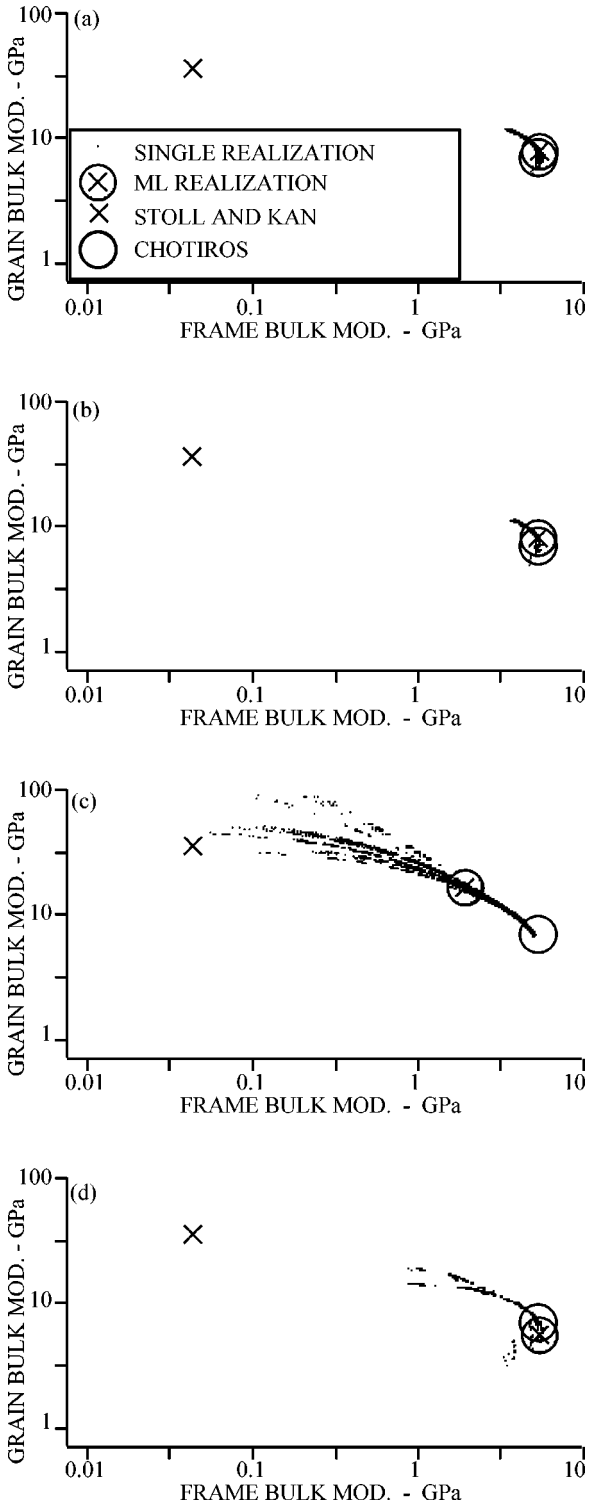


FIG. 9. Inversion results showing grain bulk modulus vs frame bulk modulus: (a) Nolle 120 μm, compacted sand, (b) Nolle 640 μm, compacted sand, (c) Drevet 250 μm, freshly settled sand, and (d) CBBL 559 μm, *in situ* West Florida sand sheet.

TABLE VI. Maximum likelihood inversion results with experimental data: composite material hypothesis.

Data set		Nolle 120	Nolle 640	Drevet 250	CBBL 559
Invertable fraction		0.69	0.87	0.91	0.64
Max. failure point		2	3	2	3
Grain bulk mod. (K_r)	GPa	11.08	36.00	42.70	36.00
Frame bulk mod. (K_b)	GPa	2.46	0.75	0.66	0.51
Frame shear mod. (μ)	GPa	0.016	0.014	0.016	0.023
Bulk log dec. (δ)	···	0.48	0.08	0.61	1.98
Shear log dec. (δ_μ)	···	0.03	0.14	0.09	0.09
Composite porosity (β_c)	···	0.44	0.57	0.44	0.58
Frame fluid fract. (β_f)	···	0.15	0.00	0.00	0.00
Solid fract. of fluid (β_s)	···	0.36	0.36	0.07	0.31
Frequency	kHz	200	500	500	58
Slow wave speed	m/s	796	545	433	415
Slow wave atten.	dB/m	2364	1281	9991	2112

moduli of water-saturated sand stones by Gazanhes[29] using an entirely different approach. Clearly there is an inconsistency between theory and experiment.

The Biot theory is labeled phenomenological, because it has just enough physics to embody the basic properties of a porous medium at a macroscopic level, without straying too far from the standard equations for elastic media. Based on an elegant pair of coupled wave equations, without being encumbered by too much physics at the microscopic level, the Biot theory should be able to match all macroscopic scale experimental measurements within experimental errors, i.e., all wave speed, attenuation, and reflection measurements. Any significant deviation from experimental data should be explicable in terms of either inadequacies of the basic assumptions or limitations of the theoretical model.

In order to solve the inconsistency between theory and experiment, two hypotheses will be considered. They will be called (1) composite material (CM) and (2) independent coefficient of fluid content (ICFC) hypotheses.

VII. COMPOSITE MATERIAL HYPOTHESIS

In the Biot equations, the unjacketed frame bulk modulus is a fundamental parameter. It may be equated to the grain bulk modulus under the assumption that the frame is made up exclusively of connected sand grains. It is possible that a proportion of the pore fluid may be trapped between grains or be otherwise attached to the frame. If so, this part of the fluid should really be counted as an integral part of the frame. The effective porosity would be reduced and the frame would be a composite of fluid and solid. This hypothesis was tested by modifying the inversion process to adjust the fluid fraction of the frame, leaving the grain bulk modulus constant at the expected value for quartz. The results were very disappointing. The inversion was unsuccessful for most of the realizations. In the course of the review process, an attempt at using the complementary hypothesis was mentioned, in which a proportion of the grains are unattached to the frame and should be considered as a part of a composite fluid.[30] It is even possible that both frame and fluid may be composites. For want of a better title, this hypothesis will be called the ''composite material'' (CM) hypothesis. Under the CM hypothesis, the density, ρ_{sc}, and bulk modulus, K_{rc}, of the composite frame are given in terms of the fluid fraction of the frame, β_f, by

$$\rho_{sc}=(1-\beta_f)\rho_s+\beta_f\rho_f,$$
$$K_{rc}=K_rK_f/[(1-\beta_f)K_f+\beta_fK_r]. \tag{6}$$

Similarly, the density, ρ_{fc}, and bulk modulus, K_{fc}, of the composite pore fluid are given in terms of the solid fraction of the fluid, β_s, by

$$\rho_{fc}=(1-\beta_s)\rho_f+\beta_s\rho_f,$$
$$K_{fc}=K_rK_f/[(1-\beta_s)K_r+\beta_sK_f]. \tag{7}$$

The parameters β_f and β_s are related to the composite porosity, β_c, and the geophysical porosity, β_g, by

$$\beta_g=\beta_c(1-\beta_s)+(1-\beta_c)\beta_f. \tag{8}$$

The geophysical porosity, β_g, is simply the measured ratio of fluid volume to total volume. The composite porosity, β_c, is the volume of the composite fluid per unit volume. To test the CM hypothesis, the bulk moduli and densities of grain and pore fluid, and the porosity term, in the Biot–Stoll equations (K_r, K_f, ρ_s, ρ_f, and β) were replaced by the composite equivalents (K_{rc}, K_{fc}, ρ_{sc}, ρ_{fc}, and β_c). The inversion algorithm was modified to adjust the values of β_f and β_s instead of K_r, which was kept at the stated value for quartz crystals. There was one problem: one variable has been replaced by two, which made the problem underdetermined. The approach taken was a heuristic one. Since it had been determined that β_f by itself rarely led to a realizable inversion, its value was set to zero and the value of β_s was adjusted. In the event that adjusting β_s failed to produce a successful inversion, the value of β_f was increased and the inversion process repeated until a successful inversion was achieved. The value of β_f was increased in steps of 0.05 from 0 to 0.2. This approach would give the lowest value of β_f to the nearest 0.05, consistent with a successful inversion. The inversion results are summarized in Table VI. The fraction of successful inversions ranged from 0.64 for the CBBL data to 0.91 for the Drevet data. This is a significant increase over the results in Table V in which the value of the grain bulk modulus was adjusted. It would appear that the com-

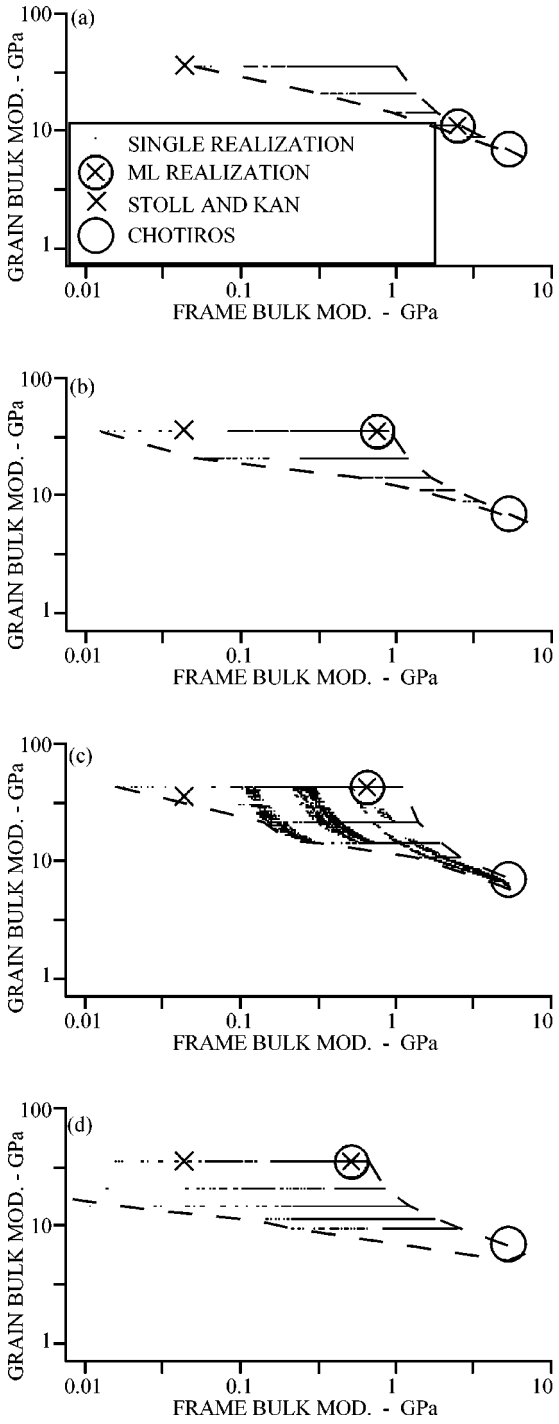


FIG. 10. Inversion results under the composite materials hypothesis: (a) Nolle 120 μm, compacted sand, (b) Nolle 640 μm, compacted sand, (c) Drevet 250 μm, freshly settled sand, and (d) CBBL 559 μm, *in situ* West Florida sand sheet.

posite material hypothesis is likely to be more consistent with both acoustic and geophysical measurements than a reduction in the value of the grain bulk modulus. Of the maximum likelihood results, the fluid fraction of the frame, β_f, was zero in all cases but one. The solid fraction of fluid, β_s, was between 0.3 and 0.4 in all cases but one. The composite porosity, β_c, was significantly higher than the geophysical porosity in all cases. The scatter plots of the inverted values of the composite grain and frame bulk moduli are shown in Fig. 10. The results are dominated by horizontal lines at the five starting values of the frame fluid fraction, i.e., 0, 0.05, 0.1, 0.15, and 0.2. Each line is the aggregate of thousands of successful inversion points. It is possible to end up with values other than these initial values, if the inversion process determined that an additional amount of pore fluid belonged to the frame, but this appears to have happened only in the case of the Drevet data. Most of the inverted values lie between the Stoll and Chotiros values. In spite of the bunching about the five starting values of the frame fluid fraction, the trends are discernible. Successful inversions appear to be possible within the region indicated by the extremities of the straight lines, as indicated by the dashed boundaries. The most likely solutions tend to have a frame bulk modulus in the region of 1 GPa. This is less than that of the Chotiros 1995 model, but greater than that of the Stoll and Kan 1981 model.

VIII. INDEPENDENT COEFFICIENT OF FLUID CONTENT (ICFC) HYPOTHESIS

The foregoing results were based on a version of Biot's theory that was formulated by Stoll, but now it is necessary to go back to Biot's original equations, in which the grain bulk modulus was not a primary input parameter. In its place were two other terms, the ''unjacketed bulk modulus'' and the ''coefficient of fluid content.'' The ''unjacketed'' bulk modulus is defined as the change in pore fluid pressure per unit change in the volume contained within the outer envelope of the frame, as measured in an ''unjacketed'' experiment. The coefficient of fluid content is defined as the volume of fluid absorbed by a porous solid, per unit volume, per unit pore pressure increase, the inverse of which also has the dimensions of bulk modulus. Thus, including the grain bulk modulus, there are three types of bulk moduli associated with the frame response to pore pressure. Biot and Willis[31] noted that under the isotropic and elastic frame assumption, the porosity would remain constant and independent of pore pressure, and the following relationship would be applicable:

$$\gamma_{\mathrm{BW}} = f_{\mathrm{BW}}(c_{\mathrm{BW}} - \delta_{\mathrm{BW}}), \tag{9}$$

where γ_{BW} is the coefficient of fluid content, f_{BW} porosity, c_{BW} compressibility of pore fluid, and δ_{BW} compressibility of the solid material. The subscript ''BW'' was added to avoid confusion with the symbols used in the Biot–Stoll equations. In Stoll's formulation, this relationship is implicitly assumed, allowing the coefficient of fluid content to be eliminated entirely as an independent variable and the grain bulk modulus to be substituted for the unjacketed bulk modulus. In the ICFC hypothesis, the coefficient of fluid content is treated as an independent parameter. If its value were greater than the right-hand side of the above-given equation, then we have a situation where the porosity of the frame increases with pore pressure, and vice versa.

Hickey and Sabatier[32] discussed the importance of the coefficient of fluid content. Using their approach, only one change in Stoll's formulation is necessary. That change is in the definition of the parameter, D, in Eq. (5) of Ref. 16, originally defined as

$$D = K_r\{1 + \beta[(K_r/K_f) - 1]\}.$$

The new definition is

$$D = K_r[1 + (K_r/K_{rf})], \tag{10}$$

where the term K_{rf} is the inverse of the coefficient of fluid content as defined by Biot and Willis,

$$K_{rf} = 1/\gamma_{\mathrm{BW}}. \tag{11}$$

Since the constitutive coefficients that affect sound propagation, i.e., M, C, and H, are dependent on D, the change will ripple through and affect all aspects of sound wave propagation modeling. The forward algorithm was modified by replacing the definition of D with Eq. (10), and the inversion process was modified to adjust the value of the coefficient of fluid content instead of the grain bulk modulus, which was kept constant at the value of quartz.

The results of the inversion on all four data sets were very encouraging, as shown in Fig. 11 and Table VII. The proportion of samples that were successfully inverted is considerably higher than obtained under the original Stoll formulation, as shown in Table V and comparable to the CM hypothesis in Table VI.

The average values of K_{rf} were significantly smaller than the value predicted for an isotropic-elastic frame, as defined in Eq. (9). The difference is shown in Table VII in terms of a ratio in the row labeled "K_{rf}/K_{rfo} ratio," where K_{rfo} is the isotropic-elastic frame value. That this ratio is less than unity suggests that the local porosity of the frame must increase with pore pressure. The Drevet data, which used freshly settled sand in the laboratory without any compacting, has a ratio that is closest to unity, while that of the CBBL data, which was taken *in situ*, is furthest from unity. This suggests that the porosity increase in the Drevet data is proportionally smaller than in the CBBL data. The Nolle data is somewhere in between. The frame bulk modulus values are comparable to the K_{rf} values.

IX. DISCUSSIONS

The composite materials (CM) hypothesis contains two elements, neither of which is new. The first element is that a proportion of the fluid may be included within the frame. This was anticipated by Biot[33] in the form of "fluids in microscopic spaces in regions of quasi contact and cracks... . These microscopic gaps must be distinguished from the pores proper which are of much larger size and where the main fluid motion occurs." The second element is that of including a proportion of the grains within the pore fluid. Direct support may be found in recent theoretical and experimental studies of the distribution of stress in granular media.[34] They indicate that the frame stress is not distributed uniformly among all the grains within the medium, but concentrated within a subset of grains that form a network of "force chains." Since grains within the force chains provide the main support for the skeletal frame, the remainder of the grains must be relatively unattached and free to move as a suspension within the pore fluid. They cannot move over large distances because they will collide into the force chains but small oscillatory movement compatible to acoustic wave motion would be feasible.

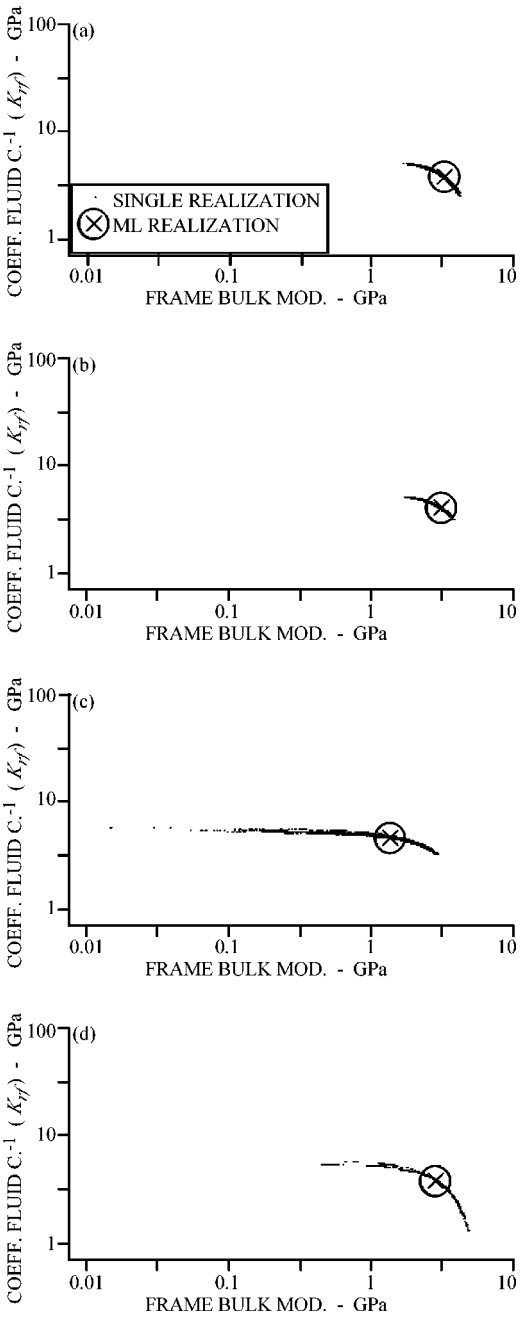


FIG. 11. Inversion results under the independent coefficient of fluid content hypothesis: (a) Nolle 120 μm, compacted sand, (b) Nolle 640 μm, compacted sand, (c) Drevet 250 μm, freshly settled sand, and (d) CBBL 559 μm, *in situ* West Florida sand sheet.

This leads to an interesting situation for a water-saturated sand bed in which the properties of the water above the bed are different from that of the effective pore fluid within the bed. It is not known if the transition is abrupt or gradual. These issues directly affect the reflection coefficient. The present model assumes that the transition is abrupt and that the boundary conditions of continuity of fluid momentum and pressure are upheld. In the case of the laboratory measurements, where the sand samples are homogeneous and the interface between water and sand is well defined, these assumptions are likely to be valid. With *in situ* measurements, the situation is unclear.

The assumption of an isotropic, elastic, and homoge-

TABLE VII. Maximum likelihood inversion results with experimental data: independent coefficient of fluid content hypothesis.

Data set		Nolle 120	Nolle 640	Drevet 250	CBBL 559
Invertable fraction		0.83	0.79	0.64	0.69
Max. failure point		1	1	1	1
Coeff. fluid C^{-1} (K_{rf})	GPa	3.88	4.10	4.71	3.87
K_{rf}/K_{rfo} ratio		0.60	0.62	0.88	0.61
Frame bulk mod. (K_b)	GPa	3.26	3.12	1.37	2.88
Frame shear mod. (μ)	GPa	0.02	0.02	0.02	0.03
Bulk log dec. (δ)	···	0.14	0.02	0.32	0.26
Shear log dec. (δ_μ)	···	0.11	0.17	0.11	0.16
Frequency	kHz	200	500	500	58
Slow wave speed	m/s	732	788	581	657
Slow wave atten.	dB/m	1522	336	5167	182
K_{rfz}/K_{rfo} ratio		0.85	0.84	0.89	0.85

neous frame is an attractive assumption. One consequence of this assumption is that porosity must be independent of pore pressure. A porosity that changes as a function of pore pressure does not necessarily indicate a departure from isotropy or elasticity. It does indicate that pore fluid pressure is capable of changing the frame in a way that is not compatible with the assumption of a monolithic structure. There are no measurements of porosity as a function pore pressure in water-saturated sand or glass beads that the author is aware of. According to the critical porosity concept, porosity greater than the critical value is an indication of a fluid supported structure. The critical porosity[35] for clastic sediments is estimated to be between 36% and 40%, which overlaps the range of porosity values for water-saturated sand. This would suggest that water-saturated sand is a borderline fluid supported structure. Small variations in porosity as a function of pore pressure may be approximated by a simple linear relationship, the slope of which may be directly related to the coefficient of fluid content.

The values of frame bulk modulus obtained by inversion, as shown in Tables VI and VII, are one or two orders of magnitude greater than that of the shear modulus, depending on the hypothesis chosen. This would indicate a Poisson's ratio that is very close to 0.5, and quite different from the value of dry sand which is in the range 0.1–0.3, contradicting the assumption that the frame moduli of water-saturated sand must be the same as that of the dry sand. This assumption is inconsistent with several observations. At the simplest level, it is evident to anyone who has walked bare foot on a beach that dry and wet sand are very different. In the laboratory, when water is introduced into dry sand, the volume of the sand is observed to increase slightly. Conversely, when water-saturated sand is allowed to dry by slow evaporation, its volume is found to decrease. These are indications that the dry and water-saturated skeletal frames are different. Quantitative evidence is found in measurements of wave speeds in air and water saturated sand[36] and glass beads.[37] In both cases, the frame bulk and shear moduli of the air-saturated sample may be inverted from the corresponding sound and shear wave speed measurements. When the values of the dry frame moduli are applied to compute the sound speed in the water-saturated sample, the value obtained is always lower than the measured value by a few percent. One may be tempted to ignore the discrepancy, but it is greater than the experimental error associated with sound speed measurement, which is typically less than 1%. A large increase in the frame bulk modulus is required to account for the discrepancy. The result is a Poisson's ratio that is very close to 0.5. An investigation of the physical cause is beyond the scope of the present study. It may involve the pore fluid in a form of local flow or squirt process, which has the potential to increase the frame bulk modulus without significantly changing the shear modulus.

The frame response to fluid pressure marks the major difference between the two hypotheses. In CM, the isotropic and elastic frame assumption is retained. The frame may dilate or contract but its porosity is independent of fluid pressure. This is illustrated in Fig. 12(a). The frame is represented by a lattice structure of spherical grains. For simplicity, the spheres are regularly arranged and physical contacts between spheres are not shown to emphasize that the pores are interconnected. The spacing is designed to give a porosity of 0.41. The initial boundary of a unit volume of the porous material is marked by the black circular boundary. The corresponding frame boundary is drawn in red, and the boundary of the corresponding body of pore fluid is marked by the edge of the blue region. Initially, all three boundaries are circular and coincide exactly. Consider a pore fluid pressure that varies sinusoidally in the x direction as illustrated. The isotropic and elastic frame would respond by being compressed in the high-pressure region and dilated in the low-pressure region, as illustrated by the movement of the red boundary relative to the black and the bunching of the grains in the high pressure region. The fluid responds similarly. Since the fluid is more compressible, there would be a net flow into the high-pressure region and out of the low-pressure region, as indicated by the change in boundary of the blue region relative to the red frame boundary. The porosity remains constant. In the CM hypothesis, the small amount of fluid within the frame could serve to redistribute the frame stress and lessen the peak stress at the grain–grain contacts. The fluid supported grains within the composite fluid would tend to move with the fluid. Under the ICFC hypothesis, the frame response is expected to be different. Two scenarios are illustrated. The case where the reciprocal of the coefficient of fluid content is very small is illustrated in Fig. 12(b). Although the grains are still compressed in the high-pressure region, they do not bunch together as they do in Fig. 12(a). On the contrary, they move apart and bunch together in the low-pressure region instead. This does not necessarily signify a departure for isotropy or elasticity, but it acknowledges that pore fluid pressure has the ability to modify the frame. One effect is an increase in fluid flow into the high-pressure region and away from the low-pressure region. There is, however, a middle ground. There is a value of coefficient of fluid content at which there is no net grain movement as a function of fluid pressure. The zero grain displacement value of K_{rf}, which is named K_{rfz}, is given by

$$K_{rfz}=(\beta/K_f)+(1-\beta)/K_r. \tag{12}$$

The ratio K_{rfz}/K_{rfo} is

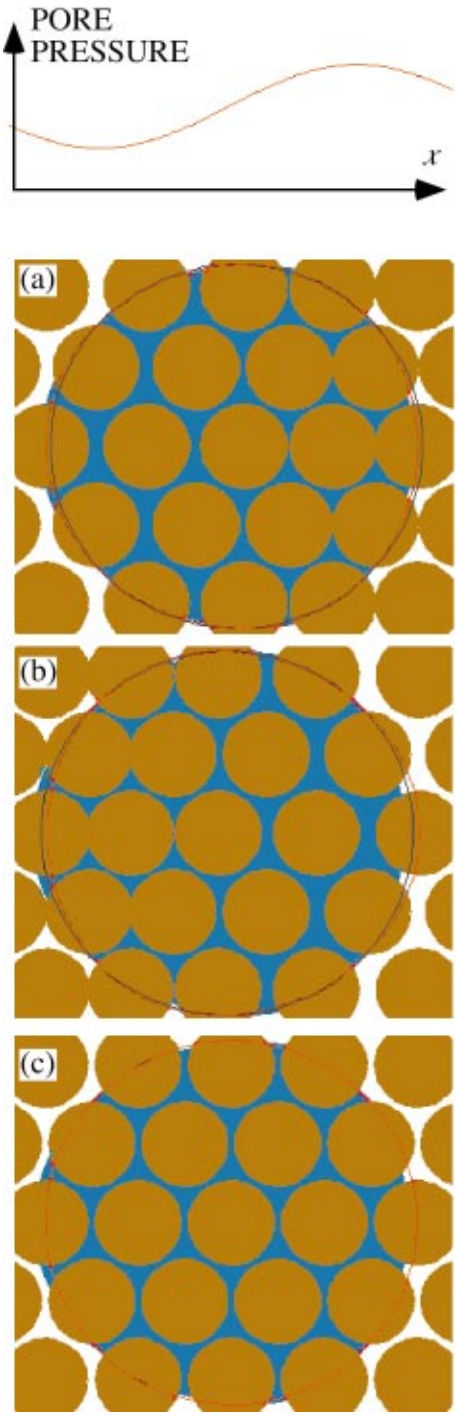


FIG. 12. Illustration of frame response to pore pressure: (a) isotropic and elastic, (b) independent coefficient of fluid content, $K_{rf}/K_{rfo} \ll 1$, and (c) $K_{rf}=K_{rfz}$.

$$K_{rfz}/K_{rfo}=[(\beta/K_f)-(\beta/K_r)]/[(\beta/K_f)+(1-\beta)/K_r]. \tag{13}$$

The value of this ratio is shown at the bottom of Table VII for each experimental data set. Three of the four inverted K_{rf}/K_{rfo} ratios are less than K_{rfz}/K_{rfo}, indicating that the frame response is of the type illustrated in Fig. 12(b). The one exception is the Drevet data set, where the two ratios are almost equal, indicating that the frame response may be closer to that of Fig. 12(c). These conclusions are contingent on acceptance of the ICFC hypothesis. If the fraction of successful inversions may be used as a guide, then it would seem that the first and last data sets favor ICFC, but the middle two data sets would seem to favor CM. There is also the possibility that both hypotheses may coexist but that is impossible to determine with the data available.

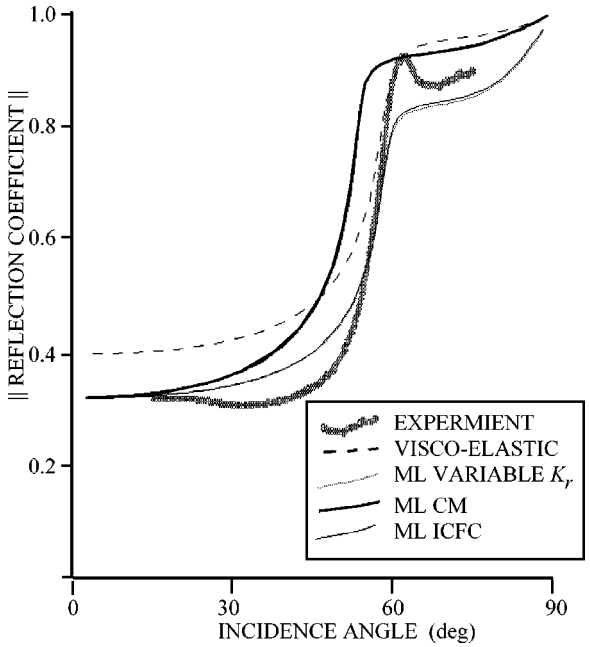


FIG. 13. Comparing experimental measurements by Drevet *et al.* and theoretical predictions based on inversion results: (a) visco-elastic, and maximum likelihood inversion results of (b) reduced grain bulk modulus, (c) composite material, and (d) independent coefficient of fluid content hypotheses.

Finally, the two hypotheses may be compared with additional acoustic data from the Drevet measurements, specifically reflection loss measurements from 20° to 70°. The physical processes that control the reflection loss at near normal incidence (20°) are quite different from those at shallow grazing angles (70°). In the former case the incident energy is mainly refracted into the fast wave but in the latter the fast wave is evanescent and the energy that is not reflected has to go elsewhere. The Stoll model was unable to fit the reflection loss curve, although by adjusting several parameters, including porosity, virtual mass constant, and fast wave speed, Drevet was able to obtain some agreement. The reflection coefficient curve computed from the visco-elastic model and the maximum likelihood inversion results are compared to the measured curve in Fig. 13. The visco-elastic model is unable to fit the reflection loss at normal incidence. The reduced grain bulk modulus inversion result of Table V is also shown for completeness; its curve is very similar to the ICFC curve. It is interesting to observe that the apparent critical angle of the ICFC and CM curves are different, even though the sediment fast wave speed and total density are identical in both models. The difference is likely connected to the composite pore fluid in the CM model, which contains a significant proportion of the grains, resulting in a reduced frame density. The ICFC curve is a better fit of the critical angle, but there are significant differences between both hypotheses and the measured curve.

The frequency dependence of the inverted parameters is

beyond the scope of this study. Each inversion was carried out for a single frequency. This is certainly true of the three laboratory data sets. In the case of the at-sea data set, all the acoustic measurements were not carried out at precisely the same frequencies due to practical limitations, but within a fairly narrow band of frequencies, and it has been implicitly assumed that the relevant parameter values remained constant within the band. In each case, the inversion result is unlikely to be applicable at any other significantly different frequency. Frequency dependence is a very difficult problem, which will have to be addressed in a later study. Biot[33] stated that the visco-elastic parameters obtained from static measurements are expected to be valid at low frequencies, where the inertial forces dominate. Frequency dependence will become significant when viscous forces overshadow the inertial forces. At that point, the problem should be handled using operators in lieu of constants.

A Poisson's ratio approaching a half is unusual if one starts with the assumption that the elastic properties of a porous frame should be related to those of the solid material by a simple scaling factor related to its porosity. The Poisson's ratio of quartz is between 0.1 and 0.2. A porous frame made of quartz would be expected to have a similar value of Poisson's ratio. This is true of the skeletal frame formed by an aggregate of dry sand grains. However, in the case of water-saturated sand, the physics is modified by the forces at the surface between sand and water. Biot[33] stated that ''the model of a purely elastic matrix saturated with a viscous fluid will only be applicable in exceptional cases.... More generally, the physical chemistry of microscopic domains of fluid in cracks and integranular gaps should be an important factor in determining the over-all rigidity and dissipation in the viscoelastic porous matrix.'' It would appear that the fluid film between granular surfaces is as important as the solid contact between grains. In this way, the skeletal frame is at least partially supported by fluid, and it is not surprising that the Poisson's ratio should tend toward a value of one half, which is fully consistent with the inversion results obtained thus far.

X. CONCLUSIONS

The visco-elastic model predictions were found to be inconsistent with measured acoustic parameters of water-saturated sand. The Biot model fared better but there were issues related to parameter values. As formulated by Stoll in terms of 13 input parameters, it was compared to four experimental data sets. Critical parameters of the model were inverted from acoustic measurements, in a process that treated the uncertainties in the values of the flow-related parameters in a probabilistic manner. Inversion processing was applied across a search space of equal probability sample points. The model was unable to match the experimentally measured values of simple acoustic properties of water saturated sand, i.e., wave speeds attenuations and reflection loss at the water-sediment interface. This conclusion is based mainly on two results: (a) the low value of grain bulk modulus obtained by inversion compared to the known value for quartz, and (b) the relatively small fraction of successful inversions. Two hypothetical modifications of the model were considered: (1) The possibility that the frame may contain fluid and that the pore fluid may contain loose grains were considered. This was called the composite material (CM) hypothesis. Both frame and fluid are allowed to be composites containing both fluid and solids. (2) The possibility that porosity may change as a function of fluid pressure was considered. This does not necessarily indicate a departure from isotropy or elasticity, but it does indicate that pore fluid pressure is capable of changing the frame in a way that is not compatible with the assumption of a monolithic frame structure. The rate of change of porosity is related to the coefficient of fluid content, which was not an independent parameter in the original model. This was called the independent coefficient of fluid content (ICFC) hypothesis. Under this hypothesis, the number of input parameters is increased by one, from 13 to 14.

Inversion computations were made under each hypothesis, with the grain bulk modulus set at a value that was appropriate for quartz grains. The results were encouraging because the fraction of successful inversions improved significantly without having to compromise the value of any of the measured geophysical and acoustic parameters. It is difficult to say which of the two hypotheses is superior. Of the four experimental data sets, two produced a higher fraction of successful inversions with one hypothesis, and the other two did the opposite. Both hypotheses are feasible. It is not unreasonable to have a fraction of the pore fluid attached to the frame, and a proportion of the grain suspended in the pore fluid, as contemplated under the CM hypothesis. It is also not unreasonable to contemplate a departure from a porosity that is independent of pore pressure. Furthermore, the two hypotheses are not mutually exclusive. At this stage, alternative interpretations of the data are possible.

Future work will explore the properties of the new hypotheses, with further comparisons with *in situ* data. Regarding the ICFC hypothesis, the porosity changes in response to pore pressure suggest that the wave propagation process may be significantly nonlinear, and it should be possible to define and predict a coefficient of nonlinearity and then devise a test to measure it. There may be applications in other types of granular media.

ACKNOWLEDGMENTS

I would like to acknowledge my colleagues in the CBBL program, particularly chief scientist, Michael Richardson, Naval Research Laboratory (NRL), Robert Stoll, Lamont-Doherty Geological Observatory, Columbia University, and the scientists and students from all the participating institutions, for their cooperation and the many discussions that we have had on this and related topics. I would like to express my appreciation to Marc Brussieux, Jean-Pierre Sessarego, and Claire Drevet, Groupe d'Etude Sous-Marine d'Atlantique (GESMA), for numerous discussions and for sharing their data. I would like to acknowledge the contribution of Craig Hickey and James Sabatier with regard to the coefficient of fluid content. I would like to express my appreciation to Morris Stern, Eric Smith, Frank Boyle, and all my colleagues at the Applied Research Laboratories, University of Texas at Austin (ARL:UT) for their contributions in

the data collection and theoretical analysis. This work was sponsored by the Office of Naval Research, Code 321 OA, under the management of Jeff Simmen.

[1] A. W. Nolle, W. A. Hoyer, J. F. Mifsud, W. R. Runyan, and M. B. Ward, ''Acoustical properties of water-filled sands,'' J. Acoust. Soc. Am. **35**, 1394–1408 (1963).
[2] C. Drevet, M. Brussieux, and J. P. Sessarego, ''High frequency acoustic wave reflection on the surf zone seafloor,'' Acustica **5**, 701–706 (1999).
[3] N. P. Chotiros, ''Inversion and sandy ocean sediments,'' in *Full Field Inversion Methods in Ocean and Seismic Acoustics, Lerici, Italy, June 1994*, edited by O. Diachok, A. Caiti, P. Gerstoft, and H. Schmidt (Kluwer Academic, Dordrecht, 1995), pp. 353–358.
[4] J. L. Jones, C. B. Leslie, and L. E. Barton, ''Acoustic characteristics of underwater bottoms,'' J. Acoust. Soc. Am. **36**, 154–157 (1964).
[5] D. J. Dodds, ''Attenuation estimates from high resolution subbottom profiler echoes,'' in *Bottom-Interacting Ocean Acoustics*, edited by W. A. Kuperman and F. B. Jensen, NATO Conference Series, Vol. IV; Marine Sciences (Plenum, New York, 1980), pp. 525–540.
[6] N. P. Chotiros, A. P. Lyons, J. Osler, and N. G. Pace, ''Normal incidence reflection loss from a sandy sediment,'' J. Acoust. Soc. Am. (to be published).
[7] *Manual of Applied Geology for Engineers* (Institution of Civil Engineers, London, 1976).
[8] D. C. Beard and P. K. Weyl, ''Influence of texture on porosity and permeability of unconsolidated sand,'' Am. Assoc. Petr. Geol. Bull. **57**, 349–369 (1973).
[9] E. L. Hamilton, G. Shumway, H. W. Menard, and C. J. Shipek, ''Acoustic and other physical properties of shallow-water sediments off San Diego,'' J. Acoust. Soc. Am. **28**, 1–15 (1956).
[10] 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).
[11] R. D. Stoll, *Sediment Acoustics* (Springer, New York, 1989).
[12] T. Yamamoto, ''Acoustic propagation in the ocean with a poro-elastic bottom,'' J. Acoust. Soc. Am. **73**, 1587–1596 (1983).
[13] M. Stern, A. Bedford, and H. R. Millwater, ''Wave reflection from a sediment layer with depth dependent properties,'' J. Acoust. Soc. Am. **77**, 1781–1788 (1985).
[14] P. R. Ogushwitz, ''Applicability of the Biot theory. III. Wave speeds versus depth in marine sediments,'' J. Acoust. Soc. Am. **77**, 453–464 (1985).
[15] J. M. Hovem and G. D. Ingram, ''Viscous attenuation of sound in saturated sand,'' J. Acoust. Soc. Am. **66**, 1807–1812 (1979).
[16] R. D. Stoll and T. K. Kan, ''Reflection of acoustic waves at a water-sediment interface,'' J. Acoust. Soc. Am. **70**, 149–156 (1981).
[17] N. P. Chotiros, ''Biot model of sound propagation in water-saturated sand,'' J. Acoust. Soc. Am. **97**, 199–214 (1995).
[18] W. N. Mack and E. A. Leistikow, ''Sands of the World,'' Sci. Am. 62–67 (1996).
[19] M. A. Biot and D. G. Willis, ''The elastic coefficients of the theory of consolidation,'' J. Appl. Mech. **24**, 594–601 (1957).
[20] R. D. Stoll, ''Comments on 'Biot model of sound propagation in water-saturated sand [J. Acoust. Soc. Am. **97**, 199–214 (1995)],' '' J. Acoust. Soc. Am. **103**, 2723–2725 (1998).
[21] N. P. Chotiros, ''Reply to 'Biot slow waves in sands near the seafloor, by R. D. Stoll,' '' J. Acoust. Soc. Am. **103**, 2726–2729 (1998).
[22] K. B. Briggs, M. D. Richardson, K. Williams, and E. I. Thorsos, ''Measurement of grain bulk modulus using sound speed measurements through liquid/grain suspensions,'' J. Acoust. Soc. Am. **104**, 1788 (1999).
[23] D. Taylor-Smith, ''Geophysical-geotechnical predictions,'' Proc. Inst. Acoust. **15**, 1–14 (1993).
[24] P. C. Carman, *Flow of Gases through Porous Media* (Academic, New York, 1956).
[25] M. D. Richardson, Coastal Benthic Boundary Layer Special Research Program: First Year, Report No. NRL/MR/7431-94-7099, Naval Research Laboratory, Stennis Space Center, MS 39529-5004, 1994.
[26] *CRC Handbook of Chemistry and Physics*, 59th ed. (CRC Press, Boca Raton, FL, 1978-9), p. F-80.
[27] *Reference Data for Radio Engineers*, 5th ed. (Sams, Indianapolis, 1970).
[28] N. P. Chotiros, ''Reflection and reverberation in normal incidence echo-sounding,'' J. Acoust. Soc. Am. **96**, 2921–2929 (1994).
[29] C. Gazanhes and J.-P. Hérault, ''Dispersion acoustique dans des roches poreuses,'' J. Phys. III **3**, 2071–2086 (1993).
[30] J. G. Berryman, ''Elastic wave propagation in fluid-saturated porous media,'' J. Acoust. Soc. Am. **69**, 416–424 (1981).
[31] M. A. Biot and D. G. Willis, ''The elastic coefficients of the theory of consolidation,'' J. Appl. Mech. **24**, 594–601 (1957), Eq. (24).
[32] C. J. Hickey and J. M. Sabatier, ''Choosing Biot parameters for modeling water-saturated sand,'' J. Acoust. Soc. Am. **102**, 1480–1484 (1997).
[33] M. A. Biot, ''Generalized theory of acoustic propagation in porous dissipative media,'' J. Acoust. Soc. Am. **34**, 1254–1264 (1962).
[34] C.-H Liu, S. R. Nagel, D. A. Schecter, S. N. Coppersmith, S. Majumdar, O. Narayan, and T. A. Witten, ''Force fluctuations in bead packs,'' Science **269**, 513–515 (1995).
[35] A. Nur, D. Marion, and H. Yin, ''Wave velocities in sediments,'' *Shear Waves in Marine Sediments, La Spezia, Italy, October 1990*, edited by J. M. Hovem, M. D. Richardson, and R. D. Stoll (Kluwer Academic, Dordrecht, 1991), pp. 131–140.
[36] E. L. Hamilton, ''Elastic properties of marine sediments,'' J. Geophys. Res. **76**, 579–605 (1971).
[37] M. R. J. Wyllie, A. R. Gregory, and L. W. Gardner, ''Elastic wave velocities in heterogeneous and porous media,'' Geophysics **21**(1), 41–70 (1956).

Performance of an underwater acoustic volume array using time-reversal focusing

Joseph A. Root and Peter H. Rogers[a)]
G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 2 October 2000; revised 28 February 2002; accepted 29 July 2002)

Time reversal permits acoustic focusing and beam forming in inhomogeneous and/or high-scattering environments. A volumetric array geometry can suppress back lobes and can fit a large, powerful array of elements into small spaces, like the free-water spaces on submarines. This research investigates applying the time-reversal method to an underwater acoustic volume array. The experiments evaluate the focusing performance of a 27-element volume array when different scattering structures are present *within* the volume of the array. The array is arranged in a $3\times3\times3$ cubic matrix configuration with 18.75-cm vertical and horizontal element spacing. The system utilizes second-derivative Gaussian pulses to focus on a point 30 cm from the array. Results include a comparison between time-reversal focusing and standard focusing, an evaluation of the volume array's ability to suppress back lobes, and an analysis of how different scattering environments affect focal region size. Potential underwater applications for a volume array using time reversal include acoustic imaging, naval mine hunting, sonar, and underwater communications. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1509073]

PACS numbers: 43.30.Vh, 43.20.Fn, 43.30.Yj [DLB]

I. INTRODUCTION

In this paper, we present a group of underwater acoustic time-reversal focusing experiments. In these experiments, a volumetric array of piezoelectric elements utilizes time reversal to focus sound back to a source point located near the array. Different scattering structures are constructed within the array volume to test array performance under different scattering conditions. The experiments address three main issues: (1) the ability of time reversal to focus in complex scattering environments versus the ability of a standard, ''time-delay'' method to focus in the same scattering environments; (2) the ability of the volumetric array to suppress a back lobe; and (3) the effect of various scattering environments on the characteristics of time-reversal focal regions.

A volumetric array/time-reversal combination offers several benefits, which point to some potential uses for such a system. First, time reversal enables focusing in high-scattering (i.e., cluttered) environments. Second, a volume array can suppress back lobes, which a plane array, by itself, cannot do. Third, provided a fixed-array element spacing, a volume array can fit more elements into certain spaces than can a planar array. This third feature becomes important when the application requires that the array fit into a given amount of space, but it also requires the highest signal levels and/or directivity possible. The drawback of the volume array is that for the same number of elements, the directivity of the volume array (excluding the back lobe) is generally worse than the directivity of the plane array. The beneficial characteristics of a volume array using time reversal suggest several naval applications: mine detection, echo ranging, underwater communications, and acoustic imaging.

Several articles published over the last decade indicate a resurgent interest in time reversal. A number of previously published papers discuss the basic theoretical aspects of time reversal.[1–4] Some research explores time reversal's ability to focus or form ultrasonic beams in the presence of various scattering environments.[5–10] Other research investigates applying time reversal to areas like medical ultrasound, ultrasonic kidney- and gallbladder stone destruction, and ultrasonic nondestructive evaluation.[1,8,11,12] One particularly significant group of experiments explores using line arrays to conduct time-reversal experiments in an ocean waveguide.[13,14]

The experiments we shall discuss here differ from previous work. Most of these previous experiments involve single transducers, line arrays, or plane arrays. Although theoretical formulations for time reversal easily incorporate volume arrays, other researchers have yet to publish actual experiments involving volume arrays. In addition, earlier research generally involves scattering objects placed *between* the array and the focusing point. With the experiment described here, the scattering material is located among the elements, *within* the volume of the array. This aspect is important because it shows how an array may be colocated with a significant amount of other equipment (as in the free-flooded spaces on a submarine) and still function well by using time reversal. Finally, the scatterers in these experiments are pressure release. That is, they provide a $p=0$ type boundary condition. Small pressure-release scatterers alter the sound field much more than do small scatterers which present any other (nonresonant) boundary condition. This is evident from the fact that, by definition, for a pressure-release scatterer the scattered field must be equal in size to the incident field near the boundary.

Fink, Jackson, and Dowling[1–3] describe time reversal

[a)]Electronic mail: peter.rogers@me.gatech.edu

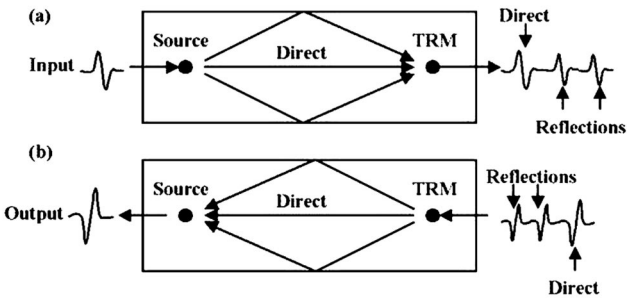


FIG. 1. An example of time reversal within a chamber containing a source and a one-element time-reversal mirror (TRM). (a) The top diagram displays the transmission of an initial signal from the source. The TRM records the direct signal as well as all the echoes created by the chamber walls. (b) The bottom diagram shows the result when the TRM transmits the time record in reverse. The echoes depart first, followed eventually by the direct-path signal. The signals retrace the paths back to the source point, resulting in all the signals arriving at the source point simultaneously.

quite well in their publications, so here we shall review the fundamentals only briefly. Acoustic time-reversal processing takes advantage of the time-reversal invariance of the lossless wave equations. Time-reversal invariance describes how, for example, a spherical wave propagating from a point source could theoretically be halted and then propagated all the way back to its source, with the inward traveling field looking exactly like the outward-going field, moving in reverse. Researchers cannot perform perfect time reversal, meaning that they cannot perfectly reconstruct an inward traveling spherical wave, but they can accomplish a practical, though imperfect, form of time reversal that can be useful for many applications. [''Perfect'' time reversal fails to be achieved for two reasons: first (as is widely appreciated), the array which is to radiate the inward-going wave does not usually constitute a closed surface in the sense required by the Helmholtz equation and second (as is not so widely appreciated), true time reversal requires that the source which created the original outgoing wave also be time reversed and thus become a sink. A perfect focus cannot be created in the absence of such a sink. The sink prevents the incoming wave from passing through the origin and interfering with itself to produce a focal region whose size is consistent with what would be expected from diffraction theory.]

Figure 1 illustrates a simple, practical time-reversal focusing example. Time reversal requires two main components: a source that generates an initial signal and a source/receiver array commonly referred to as either a time-reversal mirror (TRM) or phase-conjugate (PC) array. For this example, the array contains only one source/receiver element. The top diagram displays the source transmitting an initial signal. The TRM records the direct signal, which arrives first, followed by the later arrivals of echoes created by the chamber walls. Processing equipment then reverses the time record such that $p(t)$ becomes $p(-t)$. The bottom diagram shows what happens when the TRM transmits the reversed time record. The echoes depart first, followed eventually by the direct path's signal. The signals retrace their paths back to the source point, resulting in all the signals arriving at the source point simultaneously. It is important to note that, for time reversal to focus effectively, the element positions and

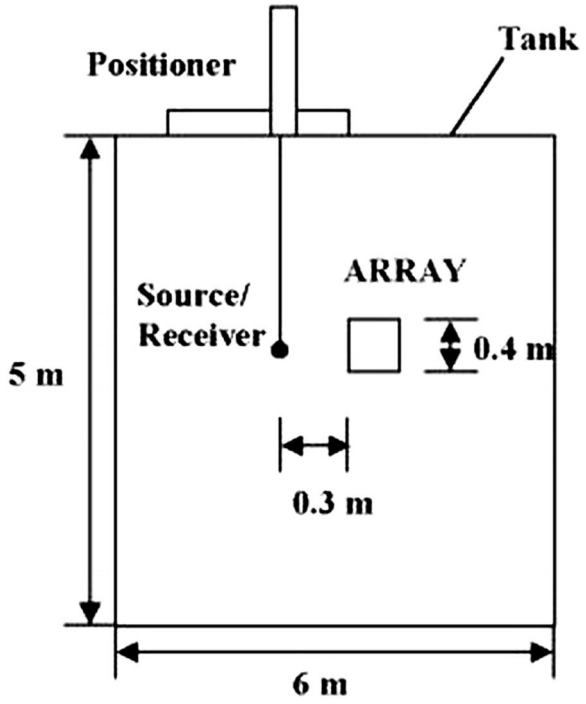


FIG. 2. The basic setup used to collect data. The time-reversal mirror and the source/receiver sit roughly in the center of the tank, as far from the tank surfaces as possible. The positioner moves the receiver through data-collection locations in the field.

the medium must remain static to ensure reciprocity.[15] In this example, the chamber provides additional echoes that supply the TRM with an aperture. More often, researchers employ multielement TRM arrays to produce an aperture for time-reversal focusing or beam forming.

II. TIME-REVERSAL FOCUSING EXPERIMENT

A. Basic setup

Figure 2 depicts the basic experimental apparatus. The main components are the volumetric array, a single source, a single receiver, a cylindrical water-filled tank, and supporting electronics. The array contains 27 individual 1.875-cm-diameter spherical piezoelectric transducers arranged in a $3\times3\times3$ cubic matrix configuration. These transducers possess a 93-kHz resonance frequency and an approximately flat receiving frequency response over the frequency band used in the experiment. The TRM elements are horizontally and vertically spaced at 18.75-cm increments. The array structure itself is constructed of 1.25-cm-diameter, solid PVC rod. The acoustic impedance of PVC is roughly twice that of water, making PVC reasonably transparent. A diagram of the array is shown in Fig. 3.

A transducer of the same type as those used in the array serves as the source that transmits the initial pulse which will be time reversed. For all experiments this source is positioned 30 cm from the front, center element of the array. The focal point for all experiments, both standard focusing and time-reversal focusing, is colocated with this source position. A hydrophone suspended from an XYZ positioner collects field data from the focusing experiments. The supporting electronics include the D/A and A/D, amplifiers, preamplifi-

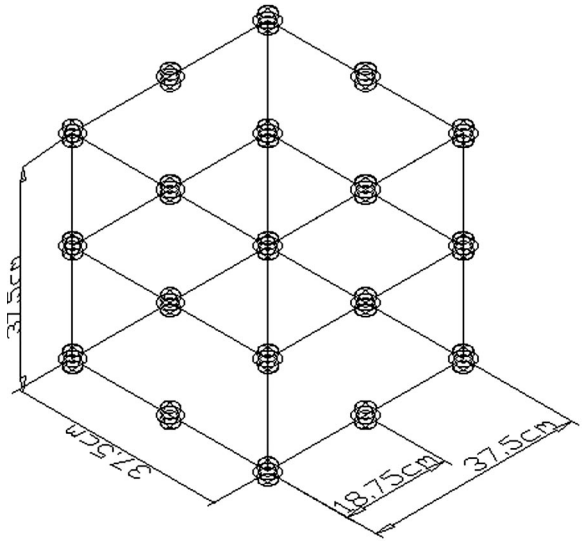


FIG. 3. A diagram of the basic, low-scattering array used in the experiments. The array contains 27 spherical piezoelectric elements arranged in a 3×3×3 cubic matrix configuration. Solid PVC rods, 1.25 cm in diameter, make up the array framework. Vertical and horizontal element spacing is set at 18.75 cm.

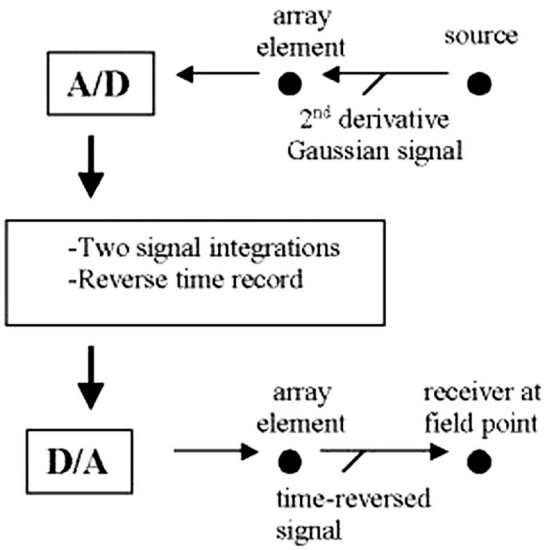


FIG. 4. Simple diagram of the experiment's time-reversal process. First, a source transmits a Gaussian signal that is then received by an array element. The array element does an A/D conversion, after which time a PC digitally integrates the signal twice and reverses the time record. After a D/A conversion, the element transmits the time-reversed signal, which is then sampled at a field point.

ers, and multiplexers. The basic acoustic waveform used for all experiments is a second-derivative Gaussian pulse centered at 25 kHz, generated by driving the piezoelectric transducer with a Gaussian voltage waveform. Choices for the focal point location, signal frequency, and array parameters reflect the simple goal of creating a reasonably sized, distinct focal region near the array.

B. Procedure

The experiments involve two types of acoustic focusing: (1) standard focusing, which serves as a comparison to time reversal and (2) time-reversal focusing. With the standard method, the array focuses by transmitting second-derivative Gaussian pulses with time delays. Those time delays are calculated using the known position of the focal point relative to the array elements. When a voltage waveform is applied to a piezoelectric transducer, below resonance, the transducer produces an acoustic pressure signal with a shape equivalent to two derivatives of the applied voltage. Thus, for the transducers to transmit second-derivative Gaussian pressure signals, the driving voltage waveforms must be Gaussian pulses.

Figure 4 illustrates the basic steps involved in time-reversal focusing. In time reversal the first step is always recording the initial signal. For the initial signal, the source transmits a single second-derivative Gaussian pulse centered at 25 kHz (generated by applying a Gaussian voltage waveform). Each element in the array acquires a time record containing a direct pulse and the associated scattered signals. Processing includes two steps. In the first step, the signal is integrated twice. If we do not perform these integrations, retransmitting the raw signal causes the frequency content to shift away from 25 kHz when the transducers take two derivatives of the voltage waveform. In the second processing step, the time record is reversed, turning it into a time-reversed waveform. After the processing is complete, the array elements transmit the time-reversed waveforms. A hydrophone moves through the field and measures the results of the time-reversal operation by recording a time record at each field point. The time-reversal operations in the experiment are synthetic, meaning that the transducers are not parallel processed for the time-reversal transmission. Each transducer transmits individually, and the resulting time records taken at a specific field point are summed, simulating what would happen if all the transducers transmitted simultaneously.

C. The arrays

One of the objectives of these experiments is to compare focusing performance with different configurations of scattering objects, positioned within the volume of the array. To satisfy this objective, the experiments include data taken with three different array types: a low-scattering array, a high-scattering array, and a high-scattering, vertical-tube array. The differences between these three arrays lie in the extent and type of scattering material within the array.

The low-scattering array consists only of the PVC frame and array elements. The intent of this array is to compare a theoretical model that includes no scattering with an experiment that includes only a small amount of scattering. The high-scattering array consists of the low-scattering array, modified by the wrapping thin sheets of closed-cell, rubber foam around each of the array framework members. The foam attached to the high-scattering array greatly increases scattering from the array structure. The main reason for the selection of a foam as the scattering material is its close approximation to a pressure release boundary. However, as we will discuss later, due to its low sound speed, a resonance occurred which had to be taken into account. The high-scattering, vertical-tube array includes the high-scattering ar-

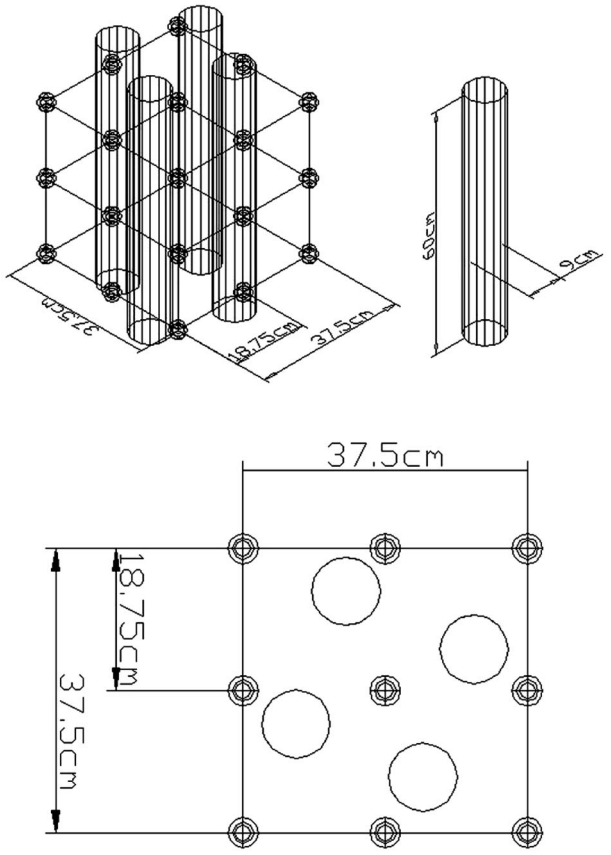


FIG. 5. Diagram of the high-scattering, vertical-tube array. Four vertical tubes are positioned vertically at regular intervals within the array volume. The hollow, open-ended PVC tubes extend 60 cm in length and have a 9-cm diameter. Closed-cell foam covers the outsides of the tubes, providing them with a high reflection coefficient.

ray with the further addition of four 9-cm-diameter, foam-covered, hollow PVC tubes. Figure 5 displays the location of the tubes inside the array. The vertical tubes further increase the amount of scattering material within the array volume, and allow us to gauge time-reversal focusing performance in the presence of large scattering objects such as might be found in the free-flooded volumes of a submarine.

D. Data presentation

Much of the data presented in this paper appear in three-dimensional pseudocolor plots. Each data point on the plots represents the normalized rms pressure value, in dB, of a time record acquired at that point in the field.

III. THEORY

In this section we present a theoretical analysis which shows how a volumetric TRM functions in the presence of multiple scatterers. The derivation is not conceptually different than others published elsewhere,[1,4–10] but we include it to provide a background for the discussion of our results.

The system is probably better understood in terms of reciprocity rather than time reversal. As has been pointed out by Tanter *et al.*,[9] this approach has the advantage of being valid even in the presence of absorption. Consider two point transducers S_A and S_F located at $\mathbf{r_A}$ and $\mathbf{r_F}$, respectively, in

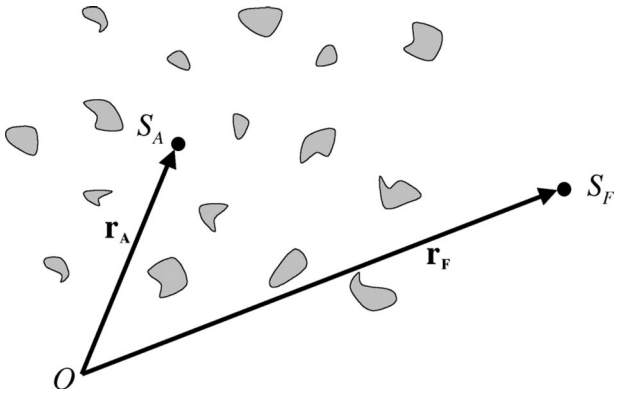


FIG. 6. Transducers S_A and S_F and nearby scatterrers.

the midst of a number of acoustic scattering objects as shown in Fig. 6. (Transducer S_A is one of the array elements and S_F is the transducer at the focal point.)

The time-dependent Green's function for the problem $G(\mathbf{r_1}|\mathbf{r_2},t-t_0)$ is defined as the acoustic pressure at time t and point $\mathbf{r_1}$ due to a point source at $\mathbf{r_2}$ whose volume velocity is a delta function centered at time t_0. Since the environment is time invariant, the Green's function can only depend on $t-t_0$ and not the individual times t and t_0. It may not be possible to calculate G, but such a function must always exist since, in principle, it could be measured experimentally. In particular, the acoustic pressure at point transducer S_A due to an impulse emitted by S_F at time t_0 is given by

$$p(\mathbf{r_A},t)=G(\mathbf{r_A}|\mathbf{r_F},t-t_0). \tag{1}$$

If the receiving sensitivity of S_A is M_A, then the voltage measured at S_A is

$$V_A(t)=M_A G(\mathbf{r_A}|\mathbf{r_F},t-t_0). \tag{2}$$

If, as in Fig. 6, there are many scatterers, $V_A(t)$ will persist over an extended time. For example, if we assume for simplicity that the scattered arrivals at S_A are simply delayed replicas of the direct signal with a positive or negative amplitude of varying sizes, $V_A(t)$ might look something like the simulation shown in Fig. 7. The direct signal is shown in black, and the total solution obtained by adding in 200 additional arrivals is shown in gray. In the simulation these arrivals are simply replicas of the direct signal with random delays and amplitudes with a slow overall exponential decay in time.

If S_A were driven with a volume velocity given by $F(t')$, by Eq. (1) the pressure measured at S_F would be given by

$$p(\mathbf{r_F},t)=\int_0^\infty G(\mathbf{r_F}|\mathbf{r_A},t-t')F(t')dt'. \tag{3}$$

Let $V_A(t)$ be recorded starting at t_{start} (an arbitrary time before the arrival of the initial signal at any sensor) and continuing until some time t_{end} when the signal is essentially over everywhere in the array. For the time-reversal mirror, $F(t')$ is chosen to be proportional to this recording played backward in time starting at some later time t_{go}. That is, $F(t')$ is given by

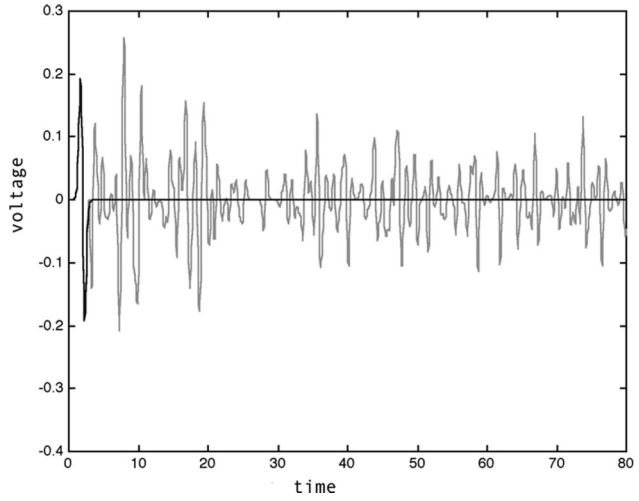


FIG. 7. Simulation of $V_A(t)$ for a high-scattering environment. The direct signal is the black line; the total solution obtained by adding in 200 additional arrivals is shown in gray.

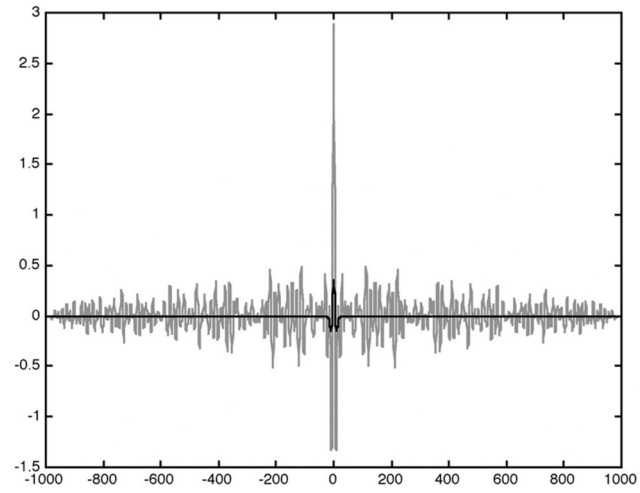


FIG. 8. Autocorrelation of the function shown in Fig. 7 (gray curve). The pressure at $\mathbf{r_F}$ would be proportional to this. The black curve is the autocorrelation of the direct part of the signal.

$$F(t') = K_F V_A(t_{\text{end}} - t' + t_{\text{go}}), \tag{4}$$

where K_F is the proportionality constant. Using Eqs. (2) and (4) in Eq. (3), we thus obtain

$$p(\mathbf{r_F},t) = M_A K_F \int_0^\infty G(\mathbf{r_F}|\mathbf{r_A},t-t') \times G(\mathbf{r_A}|\mathbf{r_F},t_{\text{go}}+t_{\text{end}}-t_0-t')dt'. \tag{5}$$

But, by reciprocity

$$G(\mathbf{r_F}|\mathbf{r_A},\tau) = G(\mathbf{r_A}|\mathbf{r_F},\tau), \tag{6}$$

so

$$p(\mathbf{r_F},t) = M_A K_F \int_0^\infty G(\mathbf{r_F}|\mathbf{r_A},t-t') \times G(\mathbf{r_F}|\mathbf{r_A},t_{\text{go}}+t_{\text{end}}-t_0-t')dt'. \tag{7}$$

Thus, the pressure measured at S_F is related to the autocorrelation of $G(\mathbf{r_F}|\mathbf{r_A},\tau)$. It will have a pronounced (positive) maximum equal to

$$p_{FA} = M_A K_F \int_0^\infty |G(\mathbf{r_F}|\mathbf{r_A},\tau)|^2 d\tau, \tag{8}$$

when

$$t = t^* = t_{\text{go}} + t_{\text{end}} - t_0. \tag{9}$$

For the example of Fig. 7, the autocorrelation of the plotted function, which would be proportional to the pressure at $\mathbf{r_F}$, is shown in Fig. 8. The pressure near $\mathbf{r_F}$ would also be concentrated spatially, since the signals transmitted to $\mathbf{r_F}$ are coming from all of the scatterers as well as directly from the source, and hence there is a large effective aperture. (The nature relationship between the temporal and spatial focusing in a TRM is discussed in Ref. 9.) The important thing to notice about Eq. (9) is that t^*, the time that the large pressure is observed at $\mathbf{r_F}$, depends only on t_0, the time that the initial signal was transmitted from S_F, t_{go}, the starting time for the retransmission, and t_{end}, the time the recording of the received signal at $\mathbf{r_A}$ is ended. It does not depend at all on $\mathbf{r_A}$ itself. If there were additional point transducers $(S_B, S_C, S_D, \ldots)$ within the volume and each received and recorded the same initial signal, stopped recording at the same time (t_{end}), and retransmitted it backward starting at the same time (t_{go}), the large signal would arrive at $\mathbf{r_F}$ from each of the sources at the exact same time, t^*. Thus, the processing we described in the previous section array produces a powerful signal at and in the vicinity of the focal point. The analysis also shows that, as has been pointed out by many others,[5,8,9] the presence of the scatterers actually tends to improve the focusing performance of the array.

Note from Eq. (8) that the contribution of S_A to the maximum value of the signal at the focal point is proportional to the total signal energy it receives from the initial pulse. The total maximum pressure at the focal point due to all sources is proportional to the total signal energy received from the initial pulse by all the array source/receivers. In general, the signal energy received by a given array transducer will be larger when there are scatterers present than it would be in the free field. However, we note for later reference that adding scatterers does not always continue to increase the received signal energy. Indeed, if too many strong scatterers were present there might be no received signal energy at all.

IV. RESULTS AND DISCUSSION

In this section we shall use experimental results to address the following issues: (1) the ability of time reversal to focus with the three arrays versus the ability of a standard, time-delay method to focus using the same arrays; (2) the ability of the volumetric array to suppress back lobes; and (3) the effect of interior scattering structures on the characteristics of time-reversal focal regions. Only limited emphasis is placed on exact numerical benefits of time reversal, considering that time reversal is greatly influenced by the exact nature of the scattering environment and that our scattering environment is admittedly quite arbitrary.

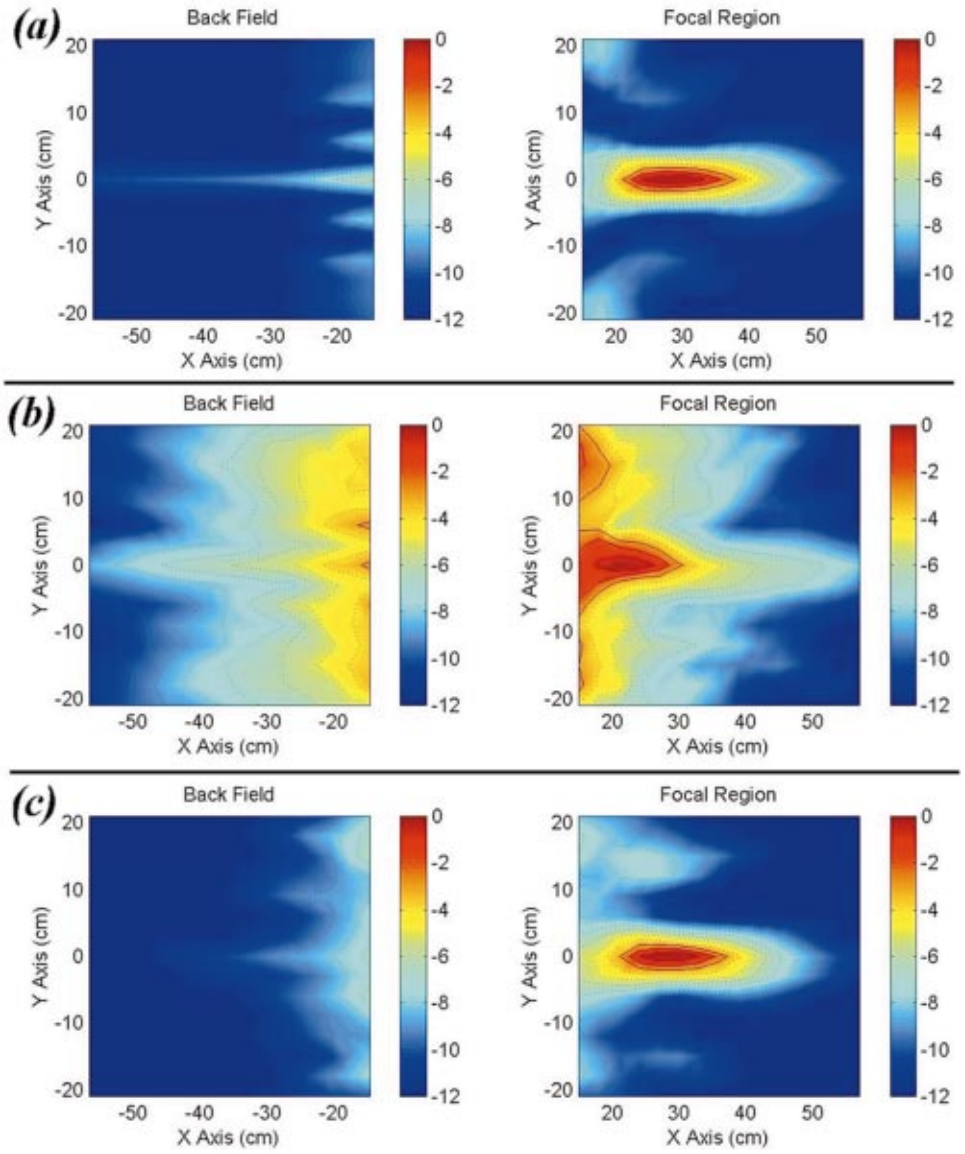


FIG. 9. Experimental results comparing the performance of time-delay focusing to time-reversal focusing in different scattering environments. The plots on the left-hand side show time-delay focusing results, while the plots on the right-hand side show time-reversal focusing results. Each horizontal pair of plots displays values in dB normalized to the highest value from either plot. (a) Focusing using the low-scattering array. (b) Focusing using the high-scattering array. (c) Focusing using the high-scattering, vertical tube array. The initiating signal is a broadband second-derivative Gaussian pulse centered at 25 kHz.

A. Time-reversal focusing versus time-delay focusing

In this section, we compare time-reversal focusing to time-delay focusing using the different arrays. An issue of interest is the sidelobe energy in the time-reversal data. For the purposes of this discussion, sidelobe energy refers to any significant energy outside of the focal region. Figure 9 displays six separate plots. The left-hand side of Fig. 9 depicts standard focusing results. The right-hand side of Fig. 9 shows time-reversal focusing results. The horizontal pairs of plots are shown in dB normalized to the highest value from *either plot* in the pair.

Figure 9(a) displays two plots of data gathered using the low-scattering array. The two plots in Fig. 9(a) appear very similar, indicating that the PVC array structure is relatively transparent acoustically and hence does not scatter a lot. However, the plots reveal two small but significant differences. First, the time-reversal plot contains less sidelobe energy than the standard plot. Second, the peak strength of the standard focal region is weaker than the peak strength of the time-reversal focal region by roughly 1 dB. These two observations indicate that even a small amount of scattering can contribute additional energy to focal regions created by time reversal.

The ability of time reversal to harness scattered energy to construct a focal region is again apparent in Fig. 9(b), which displays the results of the high-scattering array experiments. The left-hand plot shows that additional scattering seriously degrades standard focusing performance. The right-hand plot shows that, despite the additional scattering, time reversal continues to form a clear, distinct focal region.

Figure 9(c) displays the results of increasing the scattering even further with the high-scattering, vertical-tube array. While the high-scattering material places a large number of what amount to small (relative to the acoustic wavelengths involved) scattering bodies in the medium, the vertical tubes provide scattering bodies that are large relative to the acoustic wavelengths. Figure 9(c) shows that while standard focusing proves incapable of focusing in this environment, time-reversal processing once again forms a distinct focal region. This result is significant because it shows that even large scattering bodies within the volume of the array do not inhibit the ability of time-reversal processing to focus.

TABLE I. Absolute peak pressures from focal regions created with standard focusing and time-reversal focusing for the three array types. These values display the relative peak pressures present in the focal regions better than Fig. 7.

Array configuration	Absolute peak pressure from standard focal region (Pa)	Absolute peak pressure from time-reversal focal region (Pa)
Low-scattering	1850	2150
High-scattering	600	1950
High-scattering, vertical-tube	500	2200

Table I displays some information not contained in Fig. 9 due the way data for that figure were normalized. For each of the cases treated in Fig. 9, Table I compares absolute values of the peak pressures present in the focal regions. The data show that for standard focusing, the peak pressure at the focal point decreases drastically as scattering increases but it always remains high for time-reversal focusing.

Next, let us compare just the time-reversal plots (which are on the right-hand side of Fig. 9) to reveal the effect of the scattering environments on sidelobe levels. One can see a trend in the time-reversal data: As scattering increases, the amount of sidelobe energy increases. With the low-scattering array, shown in Fig. 9(a), the sidelobes are small. After the high-scattering material is added, the amount of diffuse sidelobe energy increases, as seen in Fig. 9(b). After the vertical tubes are added, the diffuse sidelobe energy increases even further, as shown in Fig. 9(c). This trend suggests that, although good focusing may be achieved in a high-scattering environment by using time-reversal processing, there may be some deterioration in sidelobe performance.

An increase in sidelobe levels seems to contradict results published by Derode,[5] which describe a significant decrease in sidelobe levels in the presence of many scattering objects. We must keep in mind that this experiment and Derode's differ greatly in geometry and, as discussed earlier, in the nature of the scatterers, and differences in the scattering environment can greatly affect the performance of a time-reversal mirror. In this case, the scattering environment is tailored to resemble something like the free-water spaces of a submarine, rather than an environment designed to be specifically well suited to time reversal. For the scattering environment used in these experiments, the results show that the scattering creates significant sidelobe levels. This is not difficult to explain. As discussed in the Theory section, scatterers which decrease rather than increase the energy received at the array elements will decrease rather than increase the energy received *at* the focus. Any scatterers, however, increase the energy produced *away* from the focus. Hence, the relative level of the sidelobes increases.

Another factor, peculiar to this experiment, significantly affects the high-scattering vertical-tube sidelobe results. Although the material coating the vertical tubes was very thin, the sound speed in the material was also very low, and what appears to be a thickness resonance in the material was found to occur around 23 kHz. The resonance affects the sidelobe levels, as shown in Fig. 9(c), by increasing the amount of diffuse energy in the field in the vicinity of 23 kHz. (Recall that our initiating signal is a broadband pulse centered at 25 kHz and that the results plotted in Fig. 9 include all of the received energy. The resonance, although fairly narrow band, can contaminate the broadband beam pattern.) In the absence of this resonance we would expect the sidelobe levels in Fig. 9(c) to be much lower. There are two possible ways in which the resonance could result in higher sidelobe energy. First, the effective aperture could be decreased near the resonance frequency because the footprint of the (resonant) vertical-tube array is smaller than the footprint of the (nonresonant) transducer array. Second, a narrow-band signal has a periodic rather than a localized autocorrelation function. The energy is thus less localized in both time and space. In addition to its effect on the sidelobes, the 23-kHz resonance also adversely affects focal region lengths and widths, as we will see later.

B. Back lobe suppression

Figure 10 characterizes the volume array's ability to suppress back lobes when focusing. [Note that a planar array would produce a back-side focus equivalent in size and strength to the front-side focus since the time delays to x and $-x$ are the same.] Each horizontal pair of plots in the figure represents a different experiment. The right-hand plots show what happens on the focusing side (front side) of the array, and the left-hand plots show what happens on the side of the array opposite the focusing (the back side). The x-axis distances shown on the plots represent position relative to the front face of the array for the front-side plots and position relative to the back face of the array for the back-side plots. To get a better idea of how the data are displayed, just imagine the volume array sitting between the two plots. Each horizontal pair of plots shows values in dB normalized to the maximum value taken from either plot.

Figure 10(a) illustrates results for a *theoretical model* containing no scattering. The model represents both time-reversal focusing and standard focusing because, in the absence of scattering, the two focusing techniques would be the same. The most important feature of Fig. 10(a) is a small center back lobe that would theoretically appear even in a nonscattering environment. The experimental data in Fig. 10(b) show what happens when the high-scattering array is used with standard focusing. As seen previously, the additional scattering significantly degrades focusing. The back lobe energy appears much higher because scattering reduces the focal region strength, which raises the relative magnitude of the lobes in the back region.

The time-reversal focusing experiment using the high-scattering array, seen in Fig. 10(c), illustrates how time-reversal processing creates a strong focal region, while at the same time blurring and suppressing the back lobes. The acoustic pressure in the back region for the time-reversal experiment is somewhat stronger but more diffuse than for the model. The scattering prevents the appearance of a distinct back lobe, like the one shown for the model. All in all, the results show that the volume array does, indeed, suppress back lobes. They show that scattering from the medium can act to blur any remaining back lobes, preventing the appearance of any particularly strong ones.

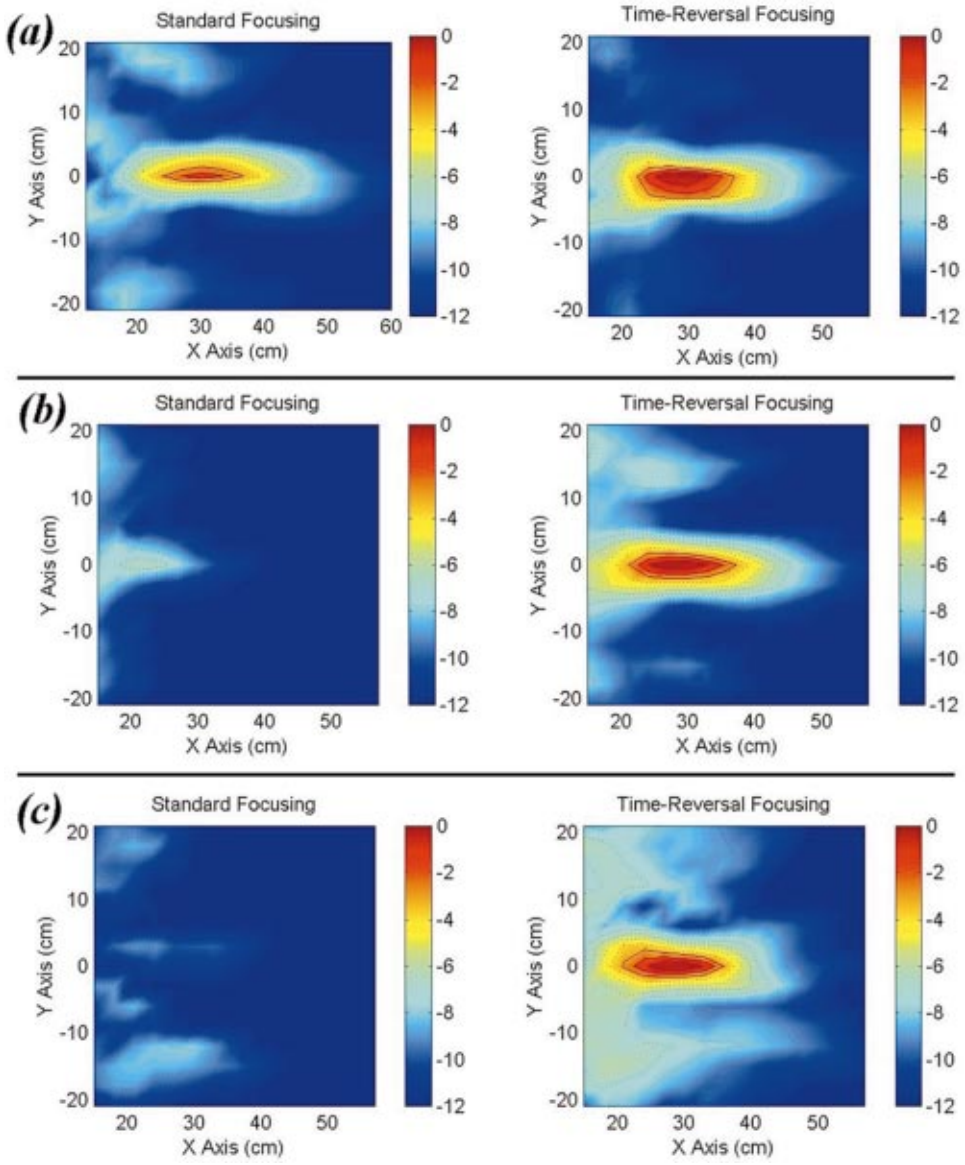


FIG. 10. Back lobe and focal region results corresponding to the results of Fig. 9. The left-hand plots show data from a horizontal plane extending through the center of the array's back lobe. The right-hand plots show data from a horizontal plane extending through the center of the focal region. Both plots display values in dB normalized to the highest value from either plot. (a) Theoretical focusing model containing no scattering. (b) Time-delay focusing data acquired using the high-scattering array. (c) Time-reversal focusing data acquired using the high-scattering array.

C. Characteristics of focal regions

In this section we shall compare the characteristics of the time-reversal focal regions created using the different arrays. The characteristics of greatest interest are the focal region length and focal region width. We use these characteristics to determine how the different scattering environments affect time-reversal focusing.

Before we proceed to the data analysis, let us first discuss some aspects of time reversal that can help us interpret the results. Time-reversal processing depends on signals from the initial source, both direct and scattered, reaching elements of the array. Time-reversed signals contain any signal energy that reaches the array elements, and increased energy in the time-reversed signals leads to smaller, stronger focal regions. Given this, one may be tempted to think that simply increasing the amount of scattering automatically leads to improved focusing. Derode *et al.*,[9] however, have shown that increased scattering improves time-reversal focusing at first, but that continued addition of scatterers will eventually degrade the focus. They attribute this to increased correlation of the scattered signal due to crossing paths. We believe that in our scattering environment, which is made up largely of pressure-release scattering elements, focusing performance can be explained entirely in terms of differences in total energy received by the array elements for the different environments [as discussed following Eq. (9)].

To see if that is true, let us compare the model and time-reversal focal region lengths and widths. Figure 11 shows the focal region lengths as a function of frequency. For calculations, the edges of the focal region are defined as being 3 dB down from the peak magnitude. The most important result of Fig. 11 is that all three of the arrays, for the most part, create focal regions as short as or shorter than the model, which includes no scattering. This result means that not only does time reversal continue to focus in complex media, but it can also focus better than it would with no scattering at all. Figure 12 shows the focal region widths, and its trends appear similar to the trends of Fig. 11.

But how, one might ask, is it possible for the time-reversal focal regions to be tighter than the conventional focusing model predicts they can be? The answer to this question lies in something not immediately obvious. For focal

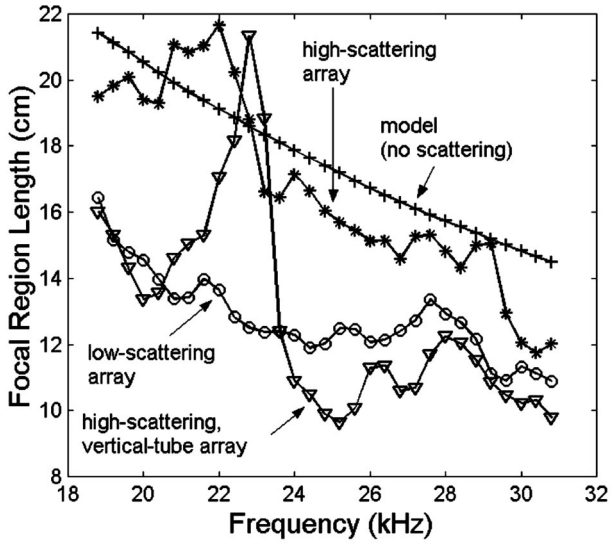


FIG. 11. Lengths of focal regions created by time reversal, as a function of frequency. Model data are also included for comparison. The edges of the focal region are defined at the points where the pressure drops 3 dB from the peak value in the focal region.

regions to be tighter than the model predicts, the aperture of the array must be somehow increased. (This is not the case for the geometries used in other experiments,[1,5–10] where the scatterers were located between the array and the focal point. For those cases, the focus can be tighter because the presence of the scatterers effectively moves the aperture closer to the focal point. The result is a shorter focal length and hence a tighter focus for a given aperture.) The diagram of the array (shown in Fig. 3) indicates that the elements are on the extreme edges of the array, meaning that they ought to mark the limits of the array aperture. Moreover, image sources from all tank-reflecting surfaces are time windowed out of the data, so they cannot increase the array aperture. What is not shown in Fig. 3, however, is the extension of the array support structure outward approximately 5 centimeters from the outermost transducer locations. This additional structure provides scattered images with a greater aperture than that of the aperture implied by the transducer positions. The scattering from external array structure was significant because it too was covered with the compliant foam for the high-scattering, vertical-tube arrays.

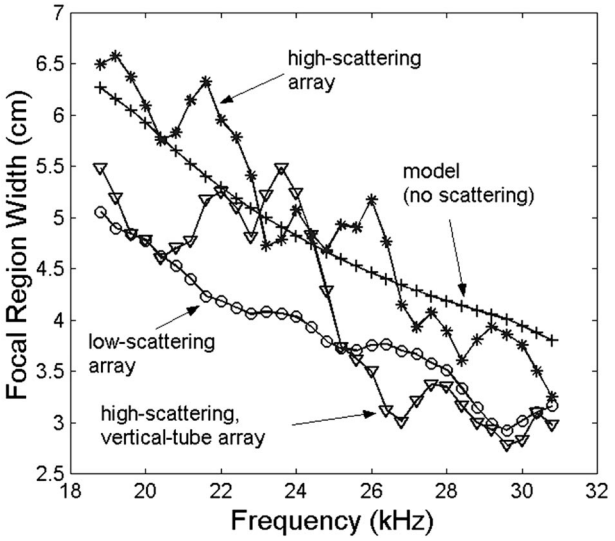


FIG. 12. Widths of focal regions created by time reversal, as a function of frequency. A model is also included for comparison. The edges of the focal region are defined at the points where the pressure drops 3 dB from the peak value in the focal region.

The data show that time reversal can focus well in the presence of scattering, but does increasing the amount of scattering always improve focusing? To answer that question, let us look at each of the different arrays and the model individually. Based on the initial discussion relating focusing to scattering, the high-scattering, vertical-tube array ought to have the smallest focal region because it has the most scattering. Then, logically, the high-scattering array should have a broader focus. Finally, the low-scattering array ought to have the largest focal region of the three arrays because it most closely resembles the model with no scattering. However, Fig. 11 shows that, on average, the low-scattering and high-scattering vertical-tube arrays create tighter focal regions than either the model or the high-scattering array. The high-scattering array and model create roughly the same size focal regions. The low-scattering array and the high-scattering, vertical-tube array also have roughly similar focal regions. These results indicate that the sheer amount of scattering in the medium does not necessarily dictate how well time-reversal focusing performs. The large spike in focal region size for the high-scattering, vertical-tube array, apparent in Figs. 11 and 12, is due to the resonance mentioned earlier and will be discussed later.

The scattering environment can improve time-reversal focusing performance by increasing the number of scattered images that it can use to focus. Time reversal is a robust phenomenon, but sometimes the scattering environment can hamper it. It is quite possible for a scattering environment to prevent some direct or scattered signals from ever reaching the array elements. When this happens, these signals are not time reversed by our time-reversal mirror, and focusing is degraded. In the language of our theoretical development, scattering can increase or decrease the energy received by the array elements and accordingly increase or decrease the focusing performance. The data in Figs. 11 and 12 show that the different array configurations scatter differently, such that the high-scattering array focus (which obviously must scatter less than the high-scattering vertical-tube array) is the worst of the three arrays. It is interesting to note that the same trend is observed in the total signal level produced at the focal point, as can be seen from Table I, where the lowest signal level occurs for the high-scattering array.

In the case of the low-scattering array, there is not much scattering to begin with and the resulting focusing is good. This result makes sense because the array structure, made of PVC, prevents very little signal energy from reaching the array elements. Although the least scattering occurs for this array, time reversal takes advantage of what scattering there is to create a tight focal region. In the case of the high-scattering array, the additional scattering material does not particularly help because it establishes a focal region larger than the low-scattering array and roughly the same size as

that of the model which includes no scattering. This broadening of the focal region after the addition of the scattering material is probably due to a decrease in the energy received by the elements in the back rows of the array. If that is the case, how does focusing improve again after the addition of the vertical tubes? (The notion that the vertical tubes would serve to screen rear-array elements from direct and scattered signals even more seems quite compelling.) The vertical tubes must, indeed, obstruct some rear elements of the array, but they also create a scattering ''wall'' directly behind the front-array elements, greatly multiplying the useful number of scattered images. Moreover, by blocking the signals to the rear rows and enhancing the signals to the front rows, the net effect of the tubes is to move the center of gravity of the array forward, which reduces the average focal length and tightens the focus. This would explain why the difference in signal level at the focus is not commensurate with the difference in the size of the focal region.

The increase in focal region length and width around 23 kHz for the high-scattering vertical-tube array is an easily explainable result of the resonance noted earlier. The 23-kHz resonance occurs only in the foam-covered vertical tubes, increases the amount of 23-kHz diffuse energy in the field, and as mentioned earlier, decreases the effective aperture. As shown earlier in Fig. 9, this increased sidelobe levels. Figures 11 and 12 show that the resonance can also broaden the focal regions. With the exception of the region around the resonance, the remaining data for the high-scattering vertical-tube array are consistent.

V. CONCLUSIONS

In this paper we describe a group of experiments that assesses the performance of an underwater acoustic volumetric array that uses time-reversal focusing. Time reversal proves to be capable of focusing as well as or, in some cases, better than a theoretical model of a volume array with no scattering, despite the presence of large amounts of scattering material within the volume of the array. The ability of the time reversal to focus can be affected (positively or negatively) by not just the amount of scattering in the medium, but also on the details of the size, shape, and location of the scatterers. While time reversal permits tight focusing in highly scattering media, sidelobes can still be present, although time reversal can also lower those sidelobe levels considerably. Both sidelobes and focusing can be seriously affected by resonance in structures near the array. When using time reversal, the volume array effectively suppresses back lobes even in a high-scattering environment.

Additional research could investigate several other aspects of a volume array using time reversal. Our experiments utilized a regularly spaced cubic matrix array. Other array geometries could prove more useful, particularly a spherical geometry. Further research should also analyze the far-field beam patterns associated with near-field time-reversal focusing.

Neither the volume array nor time reversal represents particularly new ideas; however, fusing these two concepts can offer solutions to several underwater applications, particularly ones facing space limitations, strong signal requirements, complex scattering environments, and the need for good directivity.

ACKNOWLEDGMENTS

The authors would like to thank Mr. James Martin, Mr. David Trivett, and Dr. Gary Caille for their advice and assistance.

[1] M. A. Fink, ''Time reversal in acoustics,'' Contemp. Phys. **37**, 95–109 (1996).
[2] D. R. Jackson and D. R. Dowling, ''Phase conjugation in underwater acoustics,'' J. Acoust. Soc. Am. **89**, 171–181 (1991).
[3] D. R. Dowling and D. R. Jackson, ''Narrow-band performance of phase-conjugate arrays in dynamic random media,'' J. Acoust. Soc. Am. **91**, 3257–3277 (1992).
[4] D. R. Dowling, ''Acoustic pulse compression using passive phase-conjugate processing.'' J. Acoust. Soc. Am. **95**, 1450–1458 (1994).
[5] A. Derode, P. Roux, and M. Fink, ''Robust acoustic time reversal with high-order multiple scattering,'' Phys. Rev. Lett. **75**, 4206–4209 (1995).
[6] C. Dorme and M. A. Fink, ''Ultrasonic beam steering through inhomogeneous layers with a time-reversal mirror,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 167–175 (1996).
[7] C. Draeger, J.-C. Aime, and M. Fink, ''One-channel time-reversal in chaotic cavities: Experimental results,'' J. Acoust. Soc. Am. **105**, 618–625 (1999).
[8] P. Roux, A. Derode, A. Peyre, A. Tourin, and M. Fink, ''Acoustical imaging through a multiple scattering medium using time reversal,'' J. Acoust. Soc. Am. **107**, L7–12 (2000).
[9] A. Derode, A. Tourin, and M. Fink, ''Limits of time-reversal focusing through multiple scattering: Long-range correlation,'' J. Acoust. Soc. Am. **107**, 2987–2998 (2000).
[10] M. Tanter, J. Thomas, and M. Fink, ''Time reversal and the inverse filter,'' J. Acoust. Soc. Am. **108**, 223–234 (2000).
[11] J.-L. Thomas and M. A. Fink, ''Ultrasonic beam focusing through tissue inhomogeneities with a time-reversal mirror: Applications to transskull therapy,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 1122–1129 (1996).
[12] C. Prada, J.-L. Thomas, and M. Fink, ''The iterative time-reversal process: Analysis of the convergence,'' J. Acoust. Soc. Am. **97**, 62–71 (1995).
[13] W. A. Kuperman, W. S. Hodgkiss, and H. C. Song, ''Phase conjugation in the ocean: Experimental demonstration of an acoustic time-reversal mirror,'' J. Acoust. Soc. Am. **103**, 25–40 (1998).
[14] H. C. Song, W. A. Kuperman, and W. S. Hodgkiss, ''A time-reversal mirror with variable range focusing,'' J. Acoust. Soc. Am. **103**, 3234–3240 (1998).
[15] D. R. Dowling, ''Phase-conjugate array focusing in a moving medium,'' J. Acoust. Soc. Am. **94**, 1716–1718 (1993).

Source localization in a time-varying ocean waveguide

Cristiano Soares[a)]
Faculdade de Ciências e Tecnologia, Universidade do Algarve, Campus de Gambelas, 8000-Faro, Portugal

Martin Siderius[b)]
SACLANT Undersea Research Centre, Viale S. Bartolomeo 400, 19138 La Spiezia, Italy

Sérgio M. Jesus[c)]
Faculdade de Ciências e Tecnologia, Universidade do Algarve, Campus de Gambelas, 8000-Faro, Portugal

(Received 4 September 2001; revised 15 July 2002; accepted 29 July 2002)

One of the most stringent impairments in matched-field processing is the impact of missing or erroneous environmental information on the final source location estimate. This problem is known in the literature as model mismatch and is strongly frequency dependent. Another unavoidable factor that contributes to model mismatch is the natural time and spatial variability of the ocean waveguide. As a consequence, most of the experimental results obtained to date focus on short source-receiver ranges (usually <5 km), stationary sources, reduced time windows and frequencies generally below 600 Hz. This paper shows that MFP source localization can be made robust to time–space environmental mismatch if the parameters responsible for the mismatch are clearly identified, properly modeled and (time-)adaptively estimated by a focalization procedure prior to MFP source localization. The data acquired during the ADVENT'99 sea trial at 2, 5, and 10 km source-receiver ranges and in two frequency bands, below and above 600 Hz, provided an excellent opportunity to test the proposed techniques. The results indicate that an adequate parametrization of the waveguide is effective up to 10 km range in both frequency bands achieving a precise localization during the whole recording of the 5 km track, and most of the 10 km track. It is shown that the increasing MFP dependence on erroneous environmental information in the higher frequency and at longer ranges can only be accounted for by including a time dependent modeling of the water column sound speed profile.
[DOI: 10.1121/1.1508786]

PACS numbers: 43.30.Wi, 43.60.Pt, 43.30.Pc [DLB]

I. INTRODUCTION

Matched-field processing (MFP) is an inversion method that allows a source to be located from receptions on an array and has mostly been applied to sound sources in the ocean[1,2] (see also Ref. 3 and references therein). The MFP technique compares the received field with replica fields generated for all possible source locations using an acoustic propagation model. Localization results degrade when inaccurate or insufficient data is used as input to the propagation model. This is often called the model mismatch problem and can occur when environmental information, such as the ocean sound speed, is not known in sufficient detail. Mismatch also occurs when there is uncertainty in the measurement geometry, such as receiver array position.[4,5] In classical MFP the environment and measurement geometry are assumed known and the inversion search space includes only parameters relating to the source location.

To mitigate model mismatch, a class of matched-field processors, including the so-called uncertain processors (OFUP)[6] and the focalization processor,[7] emerged in the past decade. These processors include both environmental and geometric parameters in the search space, hence reducing potential model mismatch problems. [Throughout this paper geometric parameters are those related to source and receiver positions, as well as propagation channel dimensions (e.g., water depth). Environmental parameters are those that characterize acoustic properties of the propagation channel (e.g., seabed sound speed).] It has been shown that these processors can successfully locate acoustic sources in the ocean even if the environmental knowledge is limited.[8–10]

Focalization and OFUP should perform well if the most important environmental and geometric parameters are included in the search space. This becomes more difficult as source frequency increases and therefore the relevant time and length scales needed to describe the environment decrease. This is one reason why most experimental studies on both classical MFP and focalization have been applied to data at frequencies below 500 Hz. However, sound sources in the ocean often occur at frequencies above 500 Hz. The sources may be natural (e.g., from marine life) or man-made (e.g., engines or sonar sound projectors). The main issue being addressed in this paper is to experimentally test whether MFP can be applied in a shallow water scenario at frequencies up to 1500 Hz. It will be shown that at these higher frequencies, the variability of the water column sound speed plays an important role that needs to be included in the processing.

The data considered here was collected in May, 1999

[a)]Electronic mail: csoares@ualg.pt
[b)]Electronic mail: thomas.martin.siderius@saic.com; Current address: Science Applications International Corporation, 10260 Campus Point Dr., San Diego, CA 92121.
[c)]Electronic mail: sjesus@ualg.pt

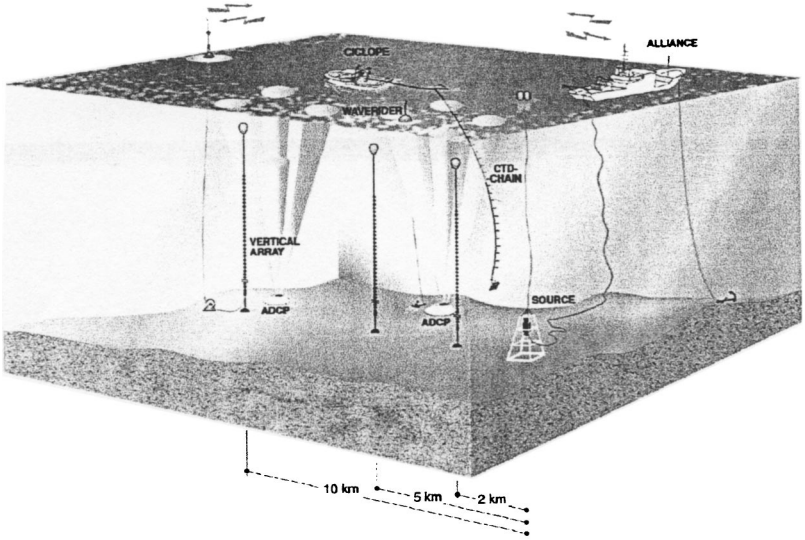


FIG. 1. ADVENT'99 sea trial setup with three vertical line array at positions 2, 5, and 10 km from a bottom mounted acoustic source. R/V Alliance is transmitting acoustic signals and collecting vertical line array data through an RF link. ITNS Ciclope is towing a CTD chain along the acoustic transmission track.

during the ADVENT'99 sea trial. One of the goals of this experiment was to test the performance of field inversion methods in shallow water under controlled conditions. The signals emitted were multitones (MT) and linear frequency modulated (LFM) sweeps in the bands 200–700 Hz and 800–1600 Hz. A vertical array was deployed on three separate days at ranges of 2, 5, and 10 km from the source.

In Ref. 11, Siderius *et al.* performed MFP inversion (focalization) on the ADVENT'99 low frequency data. The data from 2 and 10 km ranges were inverted for both the source location and seabed properties. Time- and range-independent water column sound speed profiles were used for each of the two tracks considered. Their results indicate that for the 2 km track a consistent high correlation between measured and modeled broadband data can be obtained and this resulted in reliable estimates for the source location and seabed properties. Inversion of data taken at 10 km showed increased variability in estimates for both source position and seabed properties. The conclusion is that the environmental variability can destroy coherent processing and propagation prediction of acoustic data leading to erroneous estimates for source location and seabed properties. In the case of the ADVENT'99 data this was true for the low frequency data at 10 km source receiver separation.

This paper reports successful source localization results using the ADVENT'99 data taken at ranges of 2, 5, and 10 km and up to frequencies of 1500 Hz. This is achieved through a focalization process that accounts for the time and space variability of the environment (including ocean sound speed structure). The focalization search space is directed using a genetic algorithm (GA).[12] As a further demonstration of the utility and robustness of the method, localization was tested using data from short arrays (subapertures from the ADVENT'99 vertical array). Successful application of MFP to short arrays with few sensors is important for practical applications since long, vertical arrays may not always be available. Results will be presented in Sec. III showing successful source localization in the 800–1600 Hz band at 5 km range using four sensors with a total vertical aperture of just 8 m. Vertical apertures of this size may be possible even from the droop of a horizontally towed array.

II. THE ADVENT'99 EXPERIMENT

During the first three days of May of 1999 the ADVENT'99 experiments were conducted by the SACLANT Undersea Research Center and TNO–FEL on the Adventure Bank off the southwest coast of Sicily (Italy). Figure 1 shows a drawing of the experimental setup. The bathymetry in the area has slight variations but the average water depth is 80 m. The acoustic sources were located at 76 m depth, mounted on steel framed tower that was sitting on the seabed. The signals were received on a 62 m-32 hydrophone vertical array that was deployed at ranges of 2, 5, and 10 km. Only data from 31 hydrophones are considered in this paper (element 25 is missing). Broadband LFM and MT signals were transmitted using two sound projectors, one for lower frequencies (200–700 Hz), and another for higher frequencies (800–1600 Hz). The transmission time was around 5 hours for the 2 and 5 km tracks, and 18 hours for the 10 km track. The data used here has an estimated signal-to-noise ratio greater than 10 dB.

For sound speed measurements a 49-element conductivity-temperature-depth (CTD) chain was towed by ITNS Ciclope. The CTD chain spanned around 80% of the water column reaching a maximum depth of 67 m, and was continuously towed between the acoustic source and the vertical array (10 km track). The data was sampled every 2 s which corresponds to a vertical sound speed profile measurement approximately every 4 m in range. See Refs. 13 and 14 for a more detailed description of the experiment.

III. PARAMETER ESTIMATION USING GA AND SOURCE LOCALIZATION

The focalization procedure proposed by Collins *et al.*[7] was adopted in this study to account for time-varying environments. This was accomplished using the following two-step algorithm:

(1) GA estimation of geometric and environmental param-

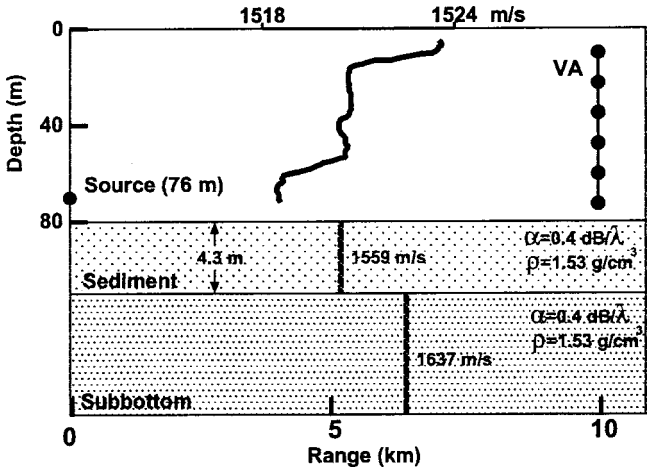


FIG. 2. Baseline model for the ADVENT'99 experiment. All parameters are range independent. The model assumes the same density and attenuation for sediment and sub-bottom.

eters: in order to determine (focus) the array position and environment to be used in the source localization—step (2).

(2) Exhaustive range-depth source localization with parameters obtained in step (1).

The matched-field localization problem usually does not include step (1) since the geometric (e.g., array position) environment (e.g., sound speed profile) and bathymetry are assumed known. There is, however, always some uncertainty in these parameters and step (1) is included here to accurately determine the optimal array position, effective bathymetry, and sound speed before application of the step-(2) localization.

One of the difficulties associated with any MFP study is the choice of the environmental model used to represent the real environment where the acoustic signal propagates. Here, we define a *baseline model* that contains a mathematical description of the real environment and constrains the attainable set of solutions.

A. The baseline model

The baseline model consists of an ocean water column overlying a sediment layer and a bottom half-space, assumed to be range independent, as shown in Fig. 2. For step (1), the forward model parameters were divided into two subsets: geometric parameters and water column parameters. The geometric parameters include source range, source depth, receiver depth, array tilt, and water depth. The parametrization of the water column will be explained later. The baseline sediment and bottom properties used for the experimental site were those estimated by Siderius *et al.*[11] using the low frequency data set. Figure 2 shows an example of a sound speed profile measured close to the vertical array at 06:38 on May 2. This profile is typical of those measured during the experiment showing a double thermocline at 10 and 55 m depth with isovelocity layers in between.

B. The objective function

The focalization of the environment and geometry [step (1)] was posed as an optimization problem, that is, to find a vector of parameters $\underline{\theta}$ that maximizes an objective function. The objective function used here is the conventional frequency incoherent broadband processor, also called the Bartlett processor, and is defined as

$$P(\underline{\theta}) = \frac{1}{N}\sum_{n=1}^{N} \underline{p}^{H}(\underline{\theta},\omega_n)\hat{\mathbf{C}}_{XX}(\omega_n)\underline{p}(\underline{\theta},\omega_n). \quad (1)$$

The factor $\hat{\mathbf{C}}_{XX}(\omega_n)$ is the sample cross-spectral matrix obtained from the observed acoustic field at frequency ω_n, N is the number of frequency bins, and $\underline{p}$ are the replica vectors to be matched with the data. All factors in (1) have norm equal 1, hence the maximum attainable value of $P(\underline{\theta})$ is 1.

The cross-spectral matrices were computed from the time series received on the 31 hydrophones using the lower frequency tones (200 to 700 Hz with 100 Hz spacing) and higher frequency tones (800, 900, 1000, 1200, 1400, 1500 Hz). Each ping of 10 s was divided into 0.5 s nonoverlapping segments, where the first and last segments were discarded, giving a total of 18 segments. Then, the 18 data segments were Fourier transformed, the bins corresponding to the multitone frequencies extracted, and the sample cross-spectral matrices computed. This procedure was repeated every 28 minutes for the 2 km track and at every 32 minutes for the 5 and 10 km tracks, for a total of 12 estimates for each track. Note that, for convenience and perceptibility, the ambiguity surfaces shown throughout this paper are only the odd numbered surfaces out of the 12 estimates.

C. Model parameter estimation and source localization

To cope with the time variability of the acoustic field the source-receiver geometry and the environmental parameters were optimized using genetic algorithm (GA) search. The GA settings were adjusted as follows: the number of iterations was set to 40 with three independent populations of 100 individuals; crossover and mutation probabilities were set to 0.9 and 0.011, respectively. These GA settings were slightly re-adjusted as additional environmental parameters were included to the search space throughout this study.

The GA optimization was carried out using a varying number of environmental parameters depending on the track. Throughout this paper it will be explained which environmental and geometric parameters were included in each inversion, but only source localization results are reported. There is no intention of validating the environmental parameters obtained throughout the various propagation tracks. The focalization step can produce a set of optimized parameters that are not necessarily correct even with a match of the replica field with the array received field. In this case a so-called equivalent model is obtained.

Table I shows the geometric parameters of the forward model and their respective search bounds for each range track. Note that source and receiving array depth are coupled to water depth, and therefore they are referenced to the depth of the bottom and referred to as source and array height.

For each source range, GA optimization was first carried out for the lower frequency MT, where the sensitivity to model mismatch is expected to be lower. Then source local-

TABLE I. Search bounds for the GA optimization of the geometric parameters; source and height of the deepest receiver on the array are measured from the water-sediment interface and are therefore coupled with the water depth search parameter. Water depth was only searched for the 10 km track and held fixed to 80 m for the 2 and 5 km tracks.

	2 km		5 km		10 km	
Parameter	min	max	min	max	min	max
Source range (km)	1.8	2.6	4.7	5.8	10.0	11.0
Source height (m)	1	10	1	10	1	10
Array height (m)	1	15	1	15	1	15
Array tilt (rad)	−0.025	0.025	−0.025	0.025	−0.025	0.025
Water depth (m)	80	80	80	80	78	82

ization ambiguity surfaces were generated for both frequency bands using replicas computed using the parameters estimated in the GA optimization. All replicas were computed using the SACLANTCEN normal mode propagation code C-SNAP.[15]

1. The 2 km track

Since model mismatch is less problematic at shorter ranges, the optimization [step (1)] for the 2 km track data set was performed directly on the higher frequency MT. The step-(1) parameter bounds for optimizing the array position are shown in Table I. The water column sound speed profile was linearly extrapolated down to the bottom using the two deepest sound speed values. No optimization was performed for the water depth since no significant water depth changes were expected in the relatively short range of 2 km.

The step-(2) range-depth ambiguity surfaces were computed for source ranges varying from 1 to 3 km, and source depth between 10 and 80 m using the baseline environmental parameters and the previously GA estimated geometry. Figure 3 shows only the odd numbered ambiguity surfaces, and Table II summarizes the results obtained through time in terms of the mean and standard deviation. The time elapsed between surfaces is 28 minutes except between surfaces 2 and 3 (where transmissions were interrupted). All surfaces show a relatively stable ambiguity pattern with a clear peak standing out from the background at the correct source position. The standard deviation is low for the source parameters and the Bartlett power has mean of 0.57 (Table II). This relatively low Bartlett power can be explained by the large frequency band being used: in order to get the maximum average match, the model has to find a parameter set that yields the best overall matched-field response, but not the best single-frequency matched-field response. If the optimization was carried out separately for each frequency, the optimum parameter set would be different and a higher matched-field response would therefore be attained for each frequency. In Fig. 4 a comparison is shown between the replicas and the received field for the tones in the best and worst match case. The best match is generally obtained at 900 Hz (0.77) and the worst match is obtained for 1500 Hz (0.56).

A full optimization including all seafloor parameters has shown little dependence of the acoustic field in this frequency band. Moreover, fluctuations in the propagation

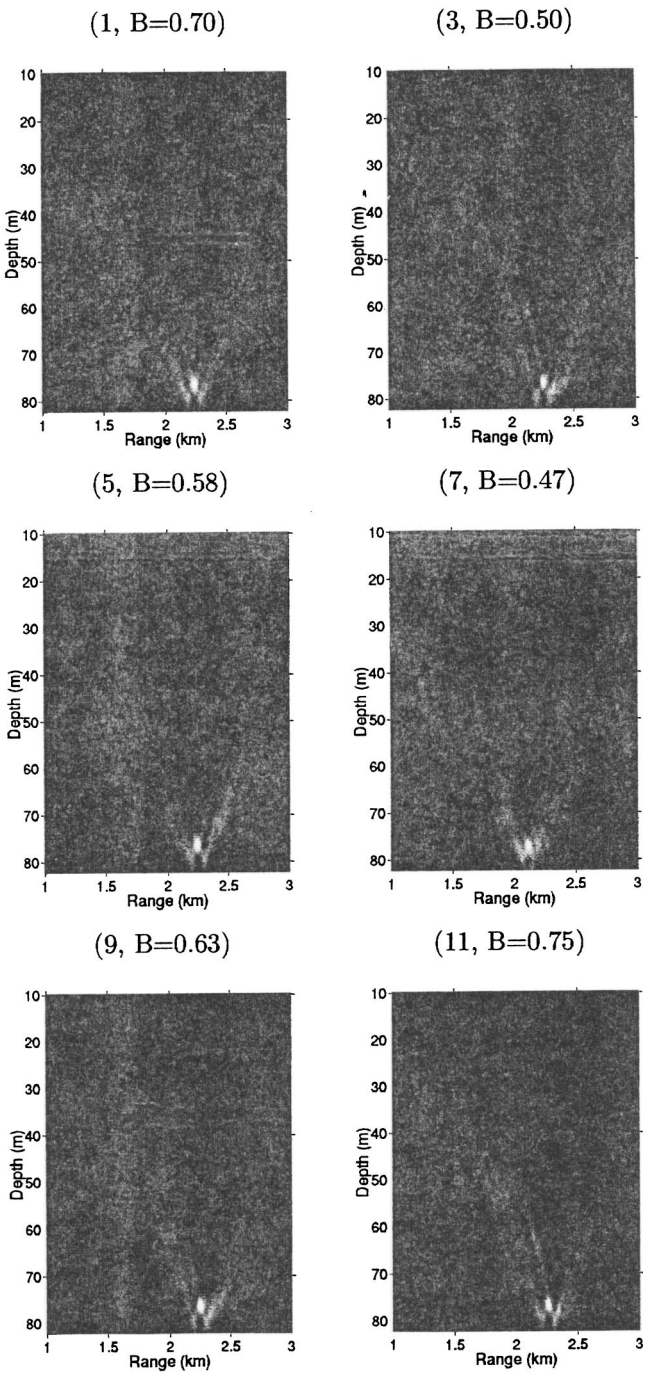


FIG. 3. Incoherent Bartlett ambiguity surfaces obtained for the 2 km track using six multitone frequency bins in the band 800 to 1500 Hz. *B* is the maximum Bartlett power obtained in each ambiguity surface.

channel cannot be taken into account by inverting the seafloor. Therefore, for the remainder of this study seafloor properties are not included in the search space.

2. The 5 km track

Having obtained stable localization results for the 2 km data set, the next step is to study the effect of a larger source-receiver range using the same procedure. First, the low frequency multitone data was processed in step (1) for the array position. Next, ambiguity surfaces were computed using the

TABLE II. Summary of the source localization results for the 2, 5, and 10 km tracks in terms of the mean and standard deviation. The data used are the higher frequency MT (800–1500 Hz) transmissions.

	2 km		5 km		10 km	
Parameter	mean	std	mean	std	mean	std
Source range (km)	2.23	0.041	5.44	0.048	10.6	0.24
Source depth (m)	76.5	0.22	75.6	0.37	74.1	1.6
Bartlett power	0.57	0.072	0.60	0.047	0.37	0.057

optimized parameters and these are shown in Fig. 5. The figure shows a well resolved peak that correctly localizes the source in both range and depth.

Ambiguity surfaces were then computed for the higher frequency MT data set [after the step-(1) otimization of the geometry] and these are shown in Fig. 6. In this case the algorithm completely failed to localize the source. Close observation of the surfaces reveals that in the first half of the run there is an ambiguity structure (peaks and sidelobes of maximum correlation) close to the expected structure from the low frequency results (see Fig. 5), and also that surfaces (4) and (5) have the peak at the correct location. The structure that appears in surface (1) to (6) disappears in surfaces (7) to (12) and this causes the true source location to be completely missed. This result suggests mismatch in environmental model used in the second half of the run.

To help mitigate the environmental mismatch problem, the water column sound speed profile was also parametrized.

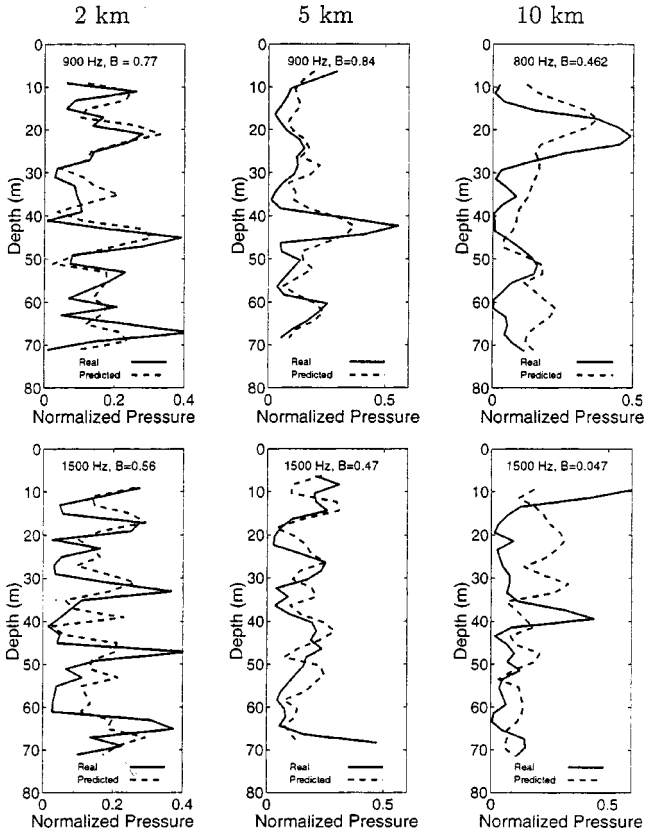


FIG. 4. Comparison of the predicted pressure with the measured pressure at frequencies with best and worst match for each range. The single frequency Bartlett power is marked on each plot.

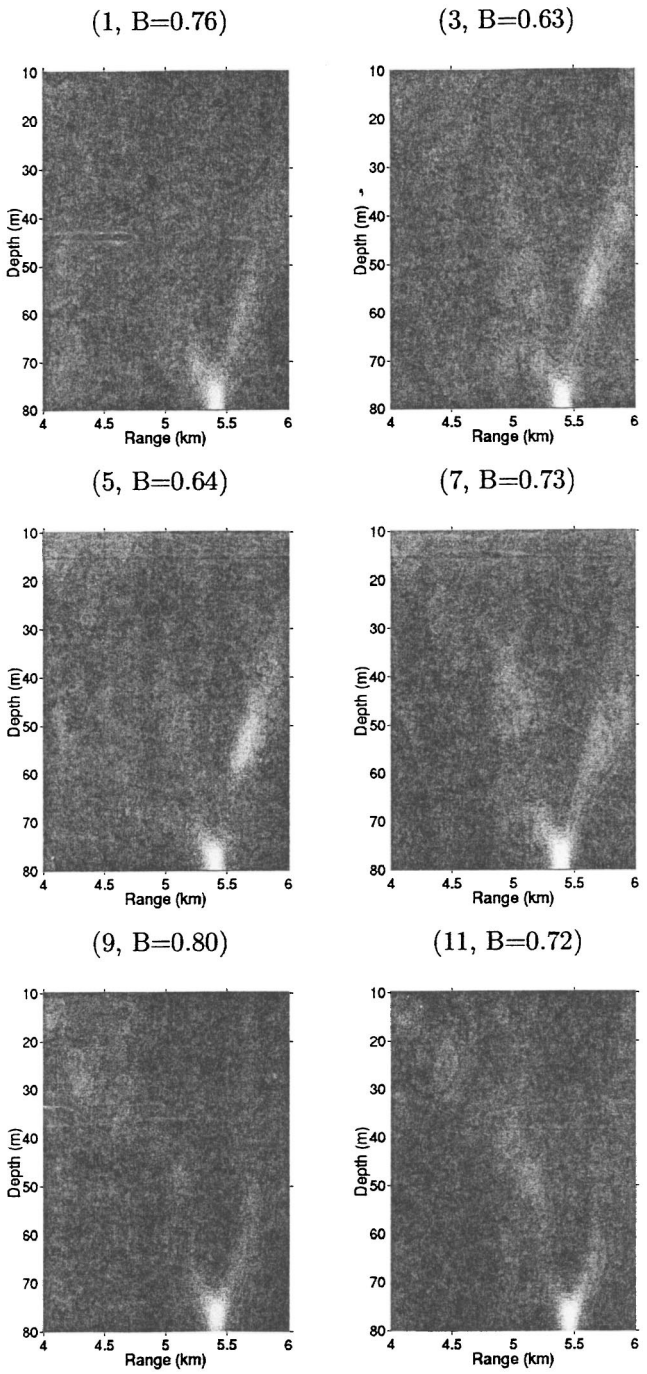


FIG. 5. Incoherent Bartlett ambiguity surfaces obtained for the 5 km track using six multitone frequency bins in the band 200 to 700 Hz. *B* is the maximum Bartlett power obtained on each ambiguity surface.

This increases the degrees of freedom of the environmental model and allows the step-(1) search to find an effective sound speed and adjust to the time varying acoustic receptions. The sound speed profile can be parametrized in several ways and combined with site measurements. To correct for possible depth errors (and short-term fluctuations) in the towed CTD chain, a search over 4 m was included in the depth of the sound speed profile measurements used in the model. Another parameter included in the search space was the gradient of the sound speed profile in the portion between 67 m and the bottom, which is a region of the water column not covered by actual measurements. This is possibly very

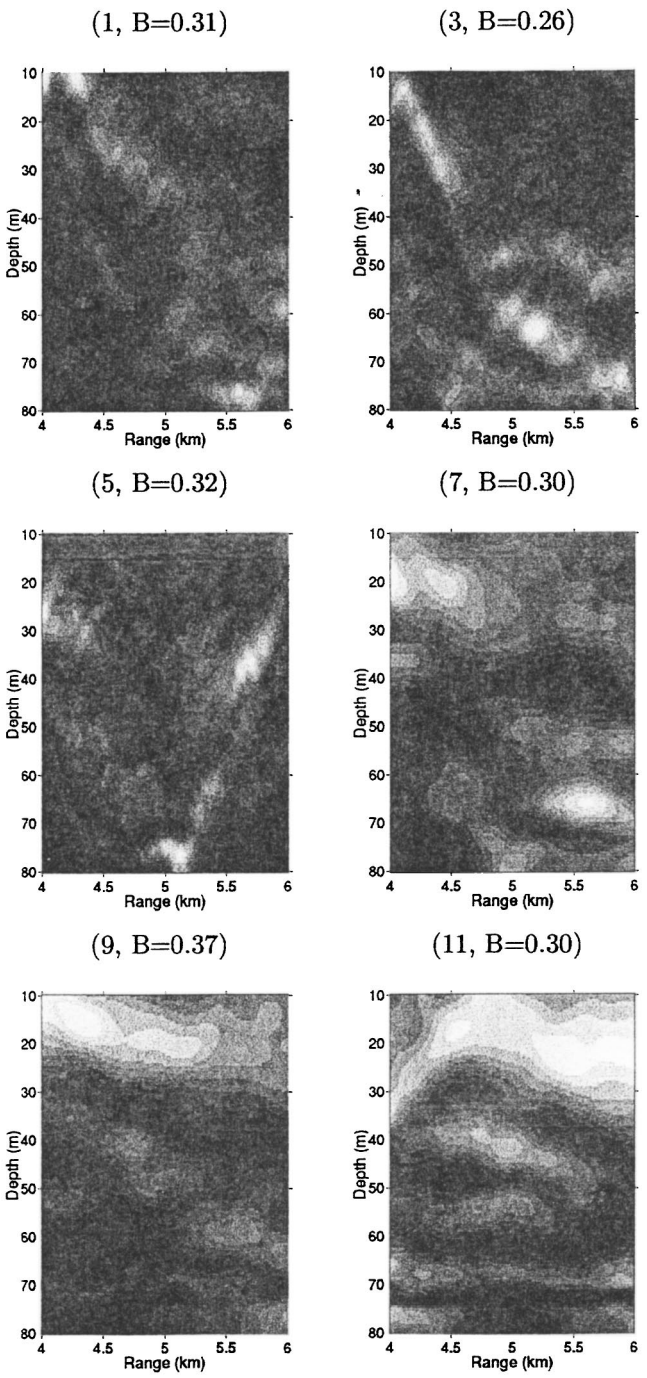


FIG. 6. Incoherent Bartlett ambiguity surfaces obtained for the 5 km track using six multitone frequency bins in the band 800 to 1500 Hz. *B* is the maximum Bartlett power obtained on each ambiguity surface.

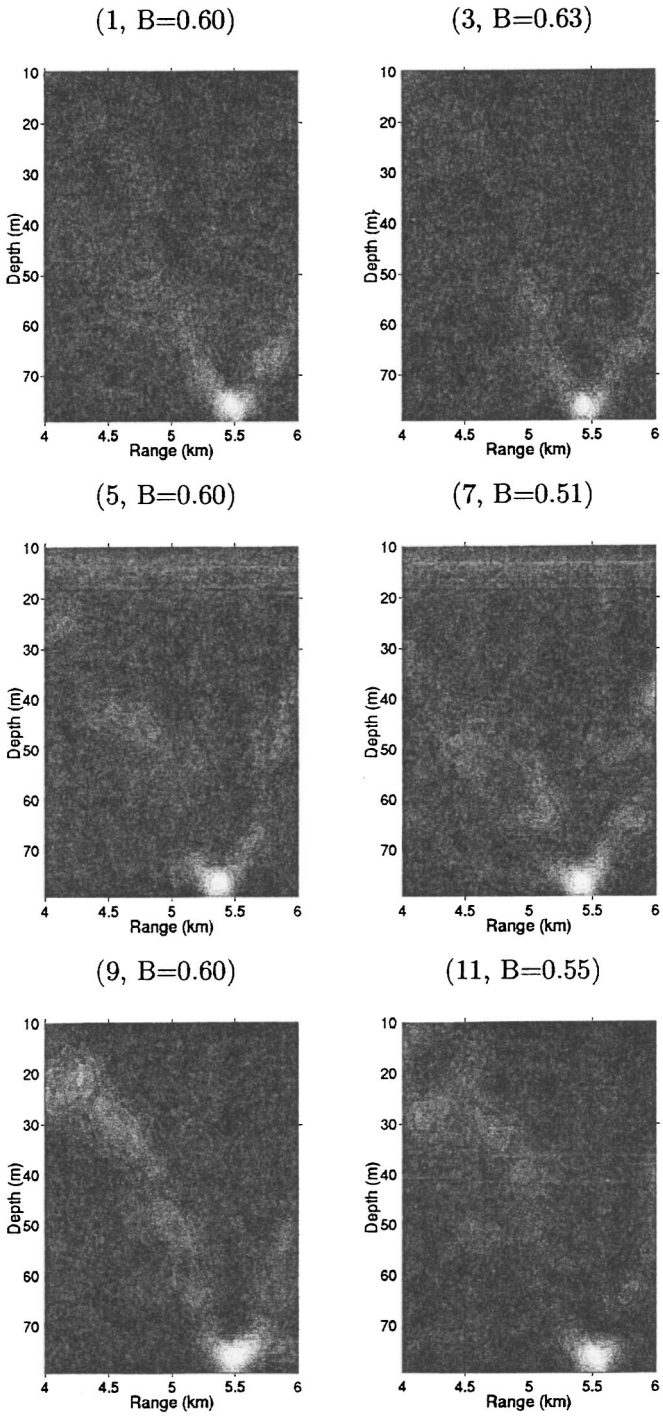


FIG. 7. Incoherent Bartlett ambiguity surfaces obtained for the 5 km track using six multitone frequency bins in the band 800 to 1500 Hz. The focalization step included estimation of the sound speed profiles via EOFs. *B* is the maximum Bartlett power obtained on each ambiguity surface.

important for predicting the acoustic field since the sound source was located near the bottom. For the 2 km track, the sound speed profile was completed by extrapolating the two deepest sound speed values down to the bottom. Including the gradient as a search parameter implicitly assumes that the sound speed behaves linearly as the depth increases but should provide a better description. Even if this is a simple and crude assumption, this parametrization was judged sufficient to model the last values of the sound speed close to the bottom. The third and last attempt to track the time variability of the acoustic data in the 5 km high frequency data set was to include a full parametrization of the sound speed profile evolution through time using a set of data based empirical orthogonal functions (EOFs). EOFs are basis functions[9] that can be obtained from a database and are very efficient to reduce the number of data points to be estimated. If historical data is available, an efficient parametrization in terms of EOFs should lead to a faster convergence and better resolution in the optimal solution since a large amount of information is already available and the search is started

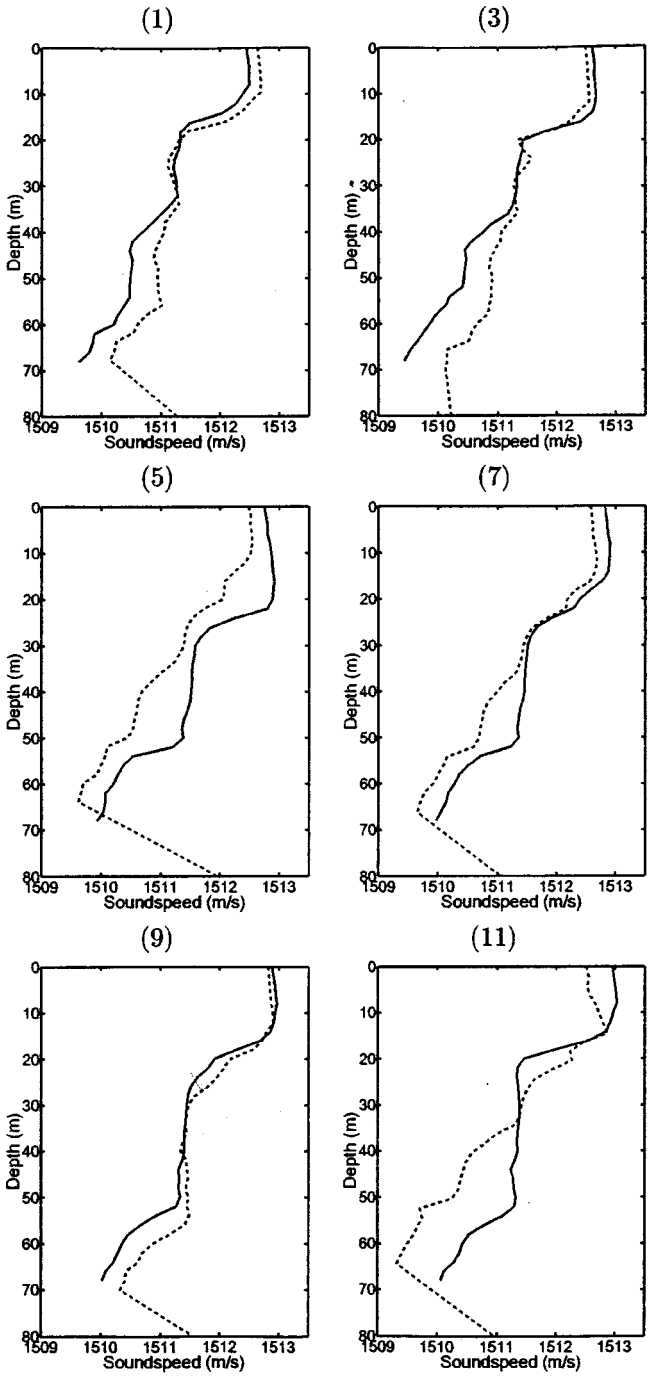


FIG. 8. Comparing sound speeds obtained by interpolation of real sound speeds with sound speeds estimated through EOFs using the higher frequency MT (800 to 1500 Hz) at the 5 km track.

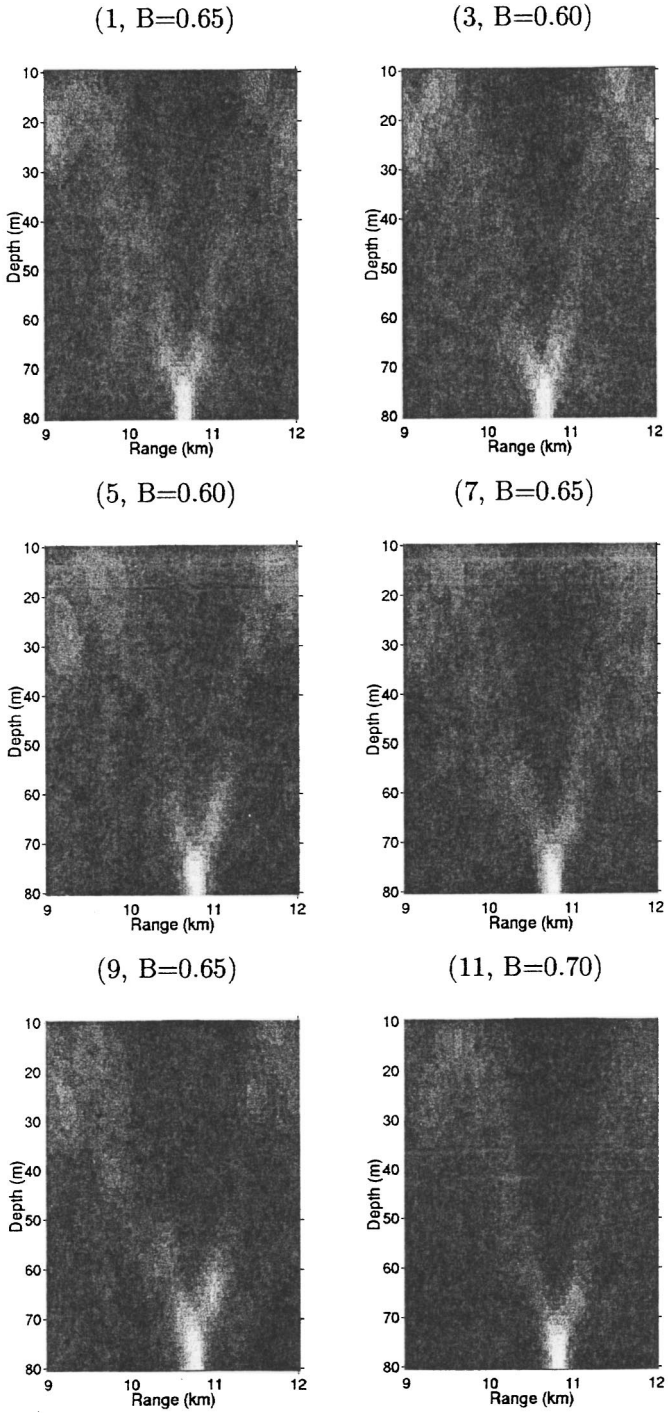


FIG. 9. Incoherent Bartlett ambiguity surfaces obtained for the 10 km track using six multitone frequency bins in the band 200 to 700 Hz. The focalization step included estimation of the sound speed profiles via EOFs. *B* is the maximum Bartlett power obtained on each ambiguity surface.

closer to the solution. The EOFs are constructed from representative data by sampling the depth dependence of the ocean sound speed. To account for the variability of the sound speed, measurements taken over time and at different locations are considered. The EOFs can be obtained from a singular value decomposition of the sample covariance matrix **C** as

$$\mathbf{C}=\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathbf{H}}, \tag{2}$$

where $\mathbf{\Lambda}=\mathrm{diag}[\lambda_1,...,\lambda_N]$ is a diagonal matrix with the eigenvalues λ_n, $\mathbf{U}=[\underline{U}_1,\ldots,\underline{U}_N]$ is a matrix with orthogonal columns, which are used as the EOFs, **H** denotes conjugate transpose, and *N* is the number of data points in the water column. Using the sound speed measurements taken close to the vertical array, it was found that three EOF's accounted for more than 80% of the water column energy. The modeled sound speed is written as

$$\underline{C}_{\mathrm{EOF}}=\overline{\underline{c}}+\sum_{n=1}^{3}\lambda_n\underline{U}_n, \tag{3}$$

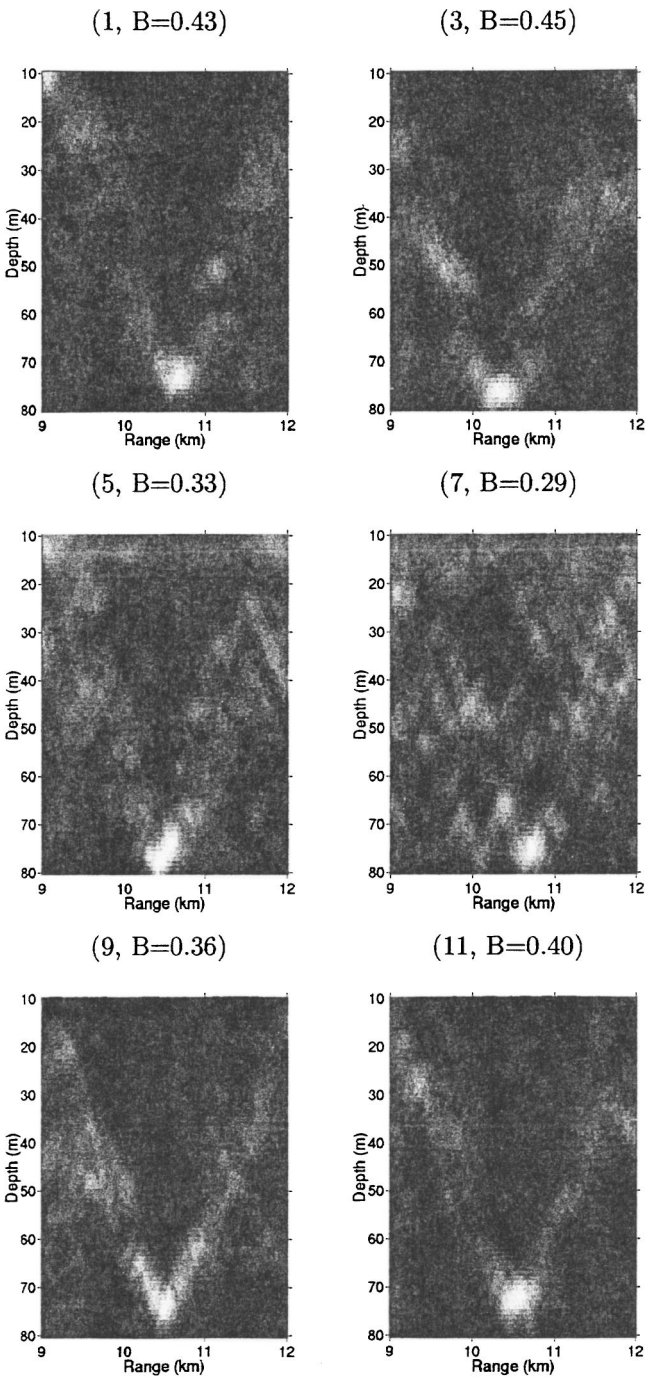


FIG. 10. Incoherent Bartlett ambiguity surfaces obtained for the 10 km track using MT frequency bins in the band 800 to 1500 Hz. The focalization step included estimation of the sound speed profiles via EOFs. *B* is the maximum Bartlett power obtained on each ambiguity surface.

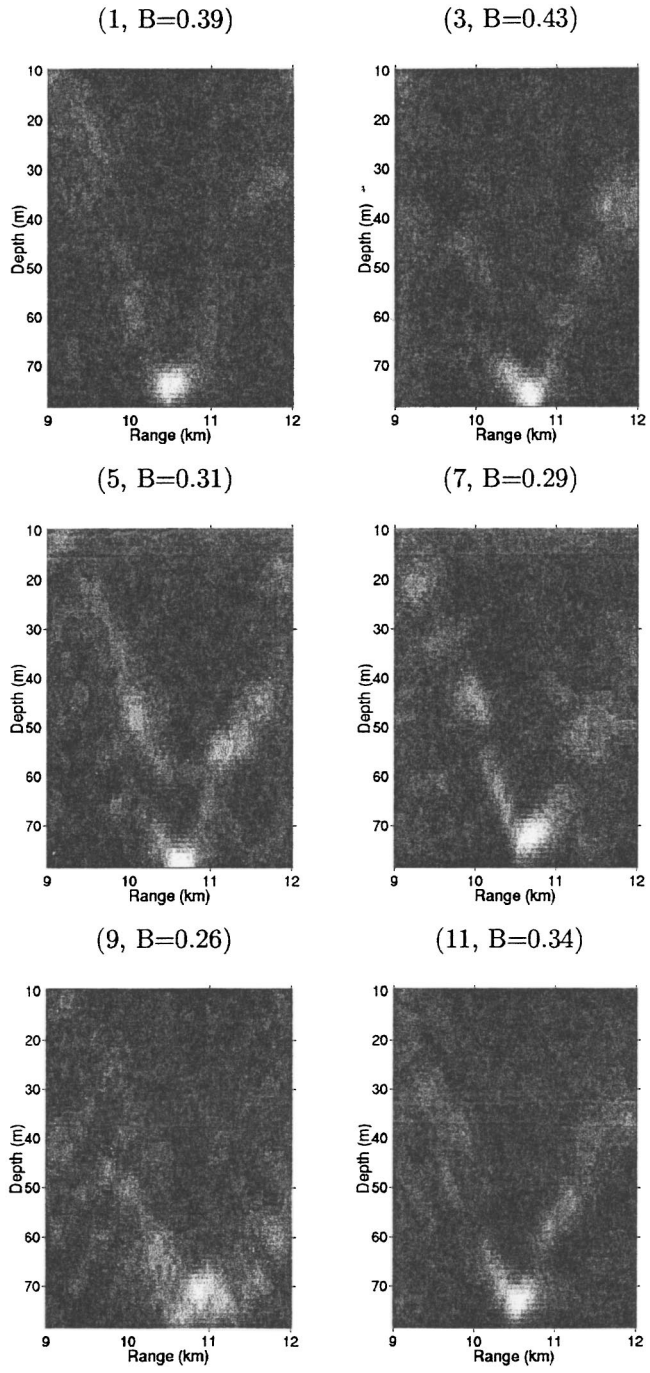


FIG. 11. Ambiguity surfaces obtained with the broadband conventional incoherent processor for the 10 km track using the higher frequency LFM signals (820 to 1500 Hz), spaced by 23.4 Hz. *B* is the maximum Bartlett power obtained on each ambiguity surface.

where $\overline{c}$ is the average sound speed profile. By trial and error, the search interval for the coefficients λ_n combining the (previously normalized) EOFs was chosen between -5 and 5.

The total number of parameters included in the focalization step is now nine, four concerning the geometry, and five concerning the sound speed in the water column, corresponding to a search space with size 2×10^{15}. Regarding the GA optimization and in order to cope with this larger search space, the number of iterations was set to 40, the number of individuals was set to 140, and the number of independent populations was three which corresponds to about 1.5×10^4 forward models to be computed.

Observing the ambiguity surfaces obtained after accomplishing the focalization step (Fig. 7), it can be seen that the main peak is always nearly at the correct location, within negligible variations, and the ambiguities are largely suppressed. In Table II it can be seen that the variability of the source parameters is of the same order of greatness as those obtained for the 2 km track and that the mean Bartlett power is even higher than that obtained for the 2 km track. Figure 8 shows plots comparing the sound speed profiles measured

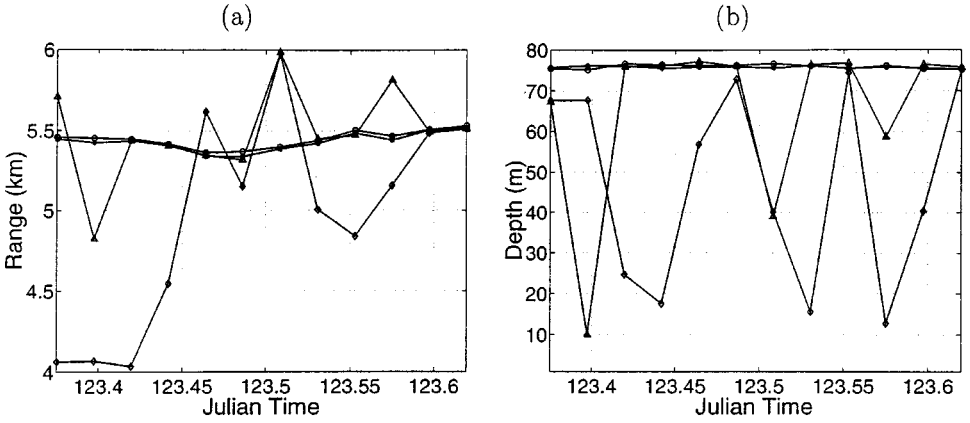


FIG. 12. Source localization results over time for the 5 km track for different array apertures using the higher frequency MT (800–1500 Hz): (a) Range; (b) depth. Each color stands for a different number of sensors: 31 (○); 16 (∗); 8 (△); 4 (◇).

close to the array and those obtained using the EOF expansion. In general it can be seen that there is a significant difference between the sound speed profiles being compared. There are at least two reasons to explain this difference: one is that the estimated profiles result from an integral of the sound speed profiles along the acoustic propagation path, and therefore include water column range dependency, if any, the other is that the estimated profiles result from a multiparameter optimization procedure and therefore may accommodate possible deviations of other environmental and/or geometric parameters, yielding the best matched-field response. An example is the sound speed profile gradient below 67 m (Fig. 8), that has often a nonphysical steep increase towards the bottom. In this paper there is no intention in validating the optimized sound speed profiles.

Finally, Fig. 4 shows that for this track the replica field at 900 Hz is well matched with the received acoustic field, and that the match at 1500 Hz is reasonable. The parametrization chosen for the sound speed allowed therefore for modeling the environment such that localization results of high quality could be obtained for the high frequency data set.

3. The 10 km track

It was shown that the modeling scheme applied at 2 km failed to accurately localize the source at 5 km for the higher frequency data set. The strategy for environmental focalization followed for the 5 km track has shown to be effective for that track, i.e., the replica fields were successfully matched with the field received by the array. Now the question is whether the focusing method applied at 5 km works also at 10 km.

The optimization strategy was therefore repeated for the 10 km track data where the only difference is the temperature data set from which the EOFs are computed. Previously, since the sound speed profiles used were only those measured close to the array, only time variability was taken into account. For this track to obtain the EOFs, a database of sound speed profiles was constructed with profiles updated every 5 minutes and comprising the whole 10 km track acquisition period. Note that, as the CTD chain was towed continuously between source and vertical array this database includes profiles that are both range and time dependent. As previously, it was found that the first three EOFs were sufficient to represent the temperature field.

The focalization step included the geometric parameters of Table I and the sound speed profile parameters as explained in Sec. II. The results obtained on the lower frequency data set are shown in Fig. 9 and indicate that the source location is well resolved over the whole time series, and the main peak is well above the background ambiguities. The source range was estimated with a mean value of 10.6 km with a small variation of approximately ± 50 m, and a source depth estimate with a variation of about 3 m around the true value of 76 m. The Bartlett power over time is larger than 0.6, which means that the agreement between the model and the measured data is relatively good.

Then, the same two step procedure was applied to the higher frequency band data set. In order to reduce the mismatch possibly due to the higher frequency and the longer range, the water depth was also included in the focalization step in agreement with the search interval shown in Table I. The ambiguity surfaces obtained for the MT analyzed are shown in Fig. 10. It can be noticed that the main peak is not completely stable in its correct position, but it always appears close to the correct location. The standard deviation is about 240 m in range and 1.6 m in depth (Table II). This corresponds to an increase relative to the 5 km track, as expected, since over a higher distance the environmental variabilities are more significant. As it can be seen in Fig. 4, there are significant difficulties, in particular at 1500 Hz, where there is no match at all (Barlett power is 0.047).

The source localization experimental results obtained for the MT are satisfactory with, however, a considerable increase of range and depth uncertainty with increasing range and with the high frequency data set. The issue is to know whether the modeling has reached its limitations, i.e., one can question whether the degree of sophistication of a range-independent environmental model is high enough to model this waveguide. For example, the hypothesis of a partially range-dependent model could be considered. Another hypothetical issue is whether it is possible to increase the amount of information inserted in the localization process by using more frequencies. Experimental results have shown that increasing the number of frequencies results in more consistent source localization estimates.[9,10,16,17] In order to test that possibility, the LFM sweeps emitted in the same time slots as the

MT were processed. A number of 30 frequencies ranging from 820 to 1500 Hz with a spacing of 23.4 Hz were selected. It can be seen that while for the tones a sustained variability in range is present, with the LFM signals it is possible to achieve periods of constancy (Fig. 11).

4. The effect of array aperture

After working with an array spanning around three quarters of the water column, and having obtained stable localization results at three different ranges, the goal now is to study the dependency of the MFP procedure on the array aperture. It is well known that, as the aperture is decreased less spatial discrimination is obtained, hence resulting in higher sidelobes relative to the main peak. In theory, even if the main peak-to-sidelobe ratio decreases with the array aperture, the main peak will always be at the correct position. However, in practice, with real data, geometric and environmental mismatch is always present and the acoustic field is observed during a limited time which creates a situation where a decrease of array aperture may indeed lead to a failure of source localization.

In the ADVENT'99 data set, reduced array aperture MFP was first applied to the higher frequency MT of the 5 km track, since for this range stable results with the full array were achieved, and could be used as reference. The array configuration used up to now had 31 elements spaced by 2 m. The procedure adopted to decrease the array aperture is to leave out sensors from the top and the bottom in such a way that the array is always centered in the water column, keeping the same spacing between sensors. The environment and geometry used were those obtained in the last focalization step for this range, hence exhaustive search is being carried out only for range and depth.

Figure 12 shows the results obtained by reducing the array aperture by successive factors of 2, i.e., from the full 31 sensors to 16, then 8 and finally to 4 sensors. The curve with circles shows the result obtained previously with the full aperture. Comparing the curves for range and depth obtained with half aperture (asterisks) to those obtained with full aperture, only negligible difference is noticed. For eight sensors (14 m of array aperture) it is still possible to obtain reasonable source localization results (curve with triangles), with, however, a larger spread in the estimates. The four sensors case (6 m aperture) represents a sampling of about 8% of the water column (curve with diamonds) where a significant degradation can be noticed.

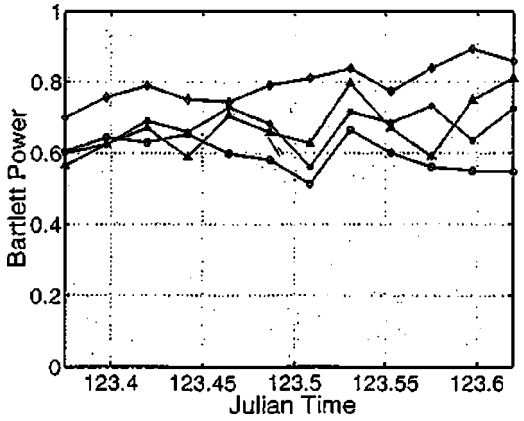


FIG. 13. Bartlett power over time for the 5 km track for different array apertures using the higher frequency MT (800–1500 Hz). Each color stands for a different number of sensors: 31 (○); 16 (*); 8 (△); 4 (◇).

Notice that the Bartlett power is highest for the four sensor configuration over the whole time even if the MFP procedure failed (Fig. 13). The increment of the Bartlett power with the reduction of the array aperture is due to the reduced complexity of the acoustic field *as seen* by the array. The matched-field response is a measure of the similarity of the real acoustic field with the replicas produced by the numerical propagation model fed with different source locations. If the aperture is small, then the uniqueness of the acoustic field along the portion of water column spanned by the field is lower, hence the match is higher.

Again, there is the possibility of compensating the lack of aperture by the addition of a large number of frequencies. Taking the LFM signals, a continuous set of frequencies is available, therefore 30 equally spaced frequencies between 820 and 1500 Hz were extracted and processed in the worst case situation of only four sensors. Figure 14 shows the improvement obtained. There are still three outliers and some variability, but now it is possible to guess that 5.4 or 5.5 km might be the correct value for range and that the source depth might be between 70 and 80 m. This case shows that it was possible to partially compensate the lack of information obtained by reduced spatial sampling by increasing the number of frequencies and obtain a reasonable source localization at

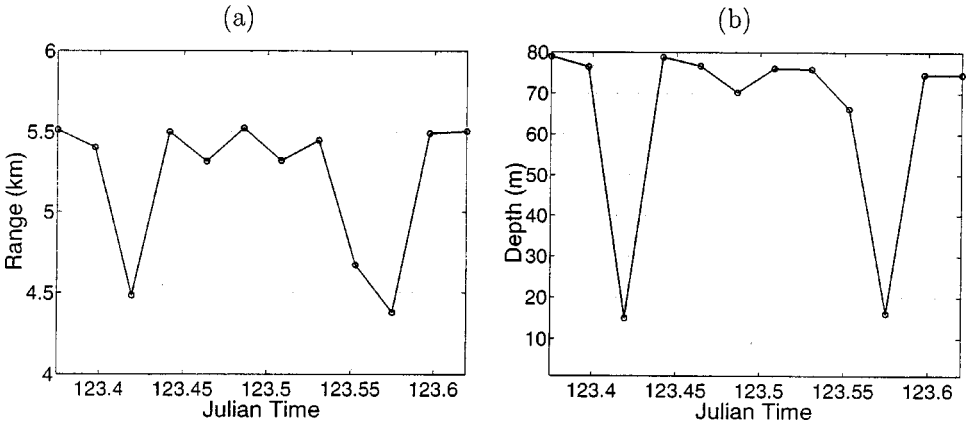


FIG. 14. Source localization results over time for an array aperture of 6 m using the higher frequency LFM signals (820–1500 Hz, spaced every 23.4 Hz): (a) range; (b) depth.

5 km from a 6 m aperture vertical array with only four sensors (8% of the water column). Using the environmental parameter set obtained during the focalization step with 31 hydrophones strongly contributed for this result.

IV. CONCLUSION

Acoustic data were collected in a 80 m depth mildly range-dependent shallow water area of the Strait of Sicily, during the ADVENT'99 sea trial in May 1999. A vertical line array was deployed at three different ranges of 2, 5, and 10 km from an acoustic source. A series of multitone and linear frequency modulated sweeps were transmitted in two frequency bands of 200 to 700 Hz and 800 to 1500 Hz.

A two stage matched-field processing (MFP) algorithm was applied throughout: (1) Focalization using genetic algorithms to search for the array position and environmental parameters giving the best fit between measured data and modeled replicas. (2) Compute the MFP ambiguity surfaces in range and depth for source localization using previously determined array position and environmental characterization.

Concerning the analysis of the ADVENT'99 data, one of the conclusions in Ref. 11 is that the environmental variability at longer ranges can destroy coherent processing and propagation prediction of acoustic data. This paper completes Ref. 11 with the following conclusions: (1) As the source range and frequency are increased, the water column variability becomes more important and this needs to be accounted for in the modeling to obtain a good match between data and replica. (2) The water column variability can be modeled using an EOF expansion and estimated with the focalization process. (3) With a full sound speed focalization precise MFP localizations could be obtained at all ranges and in both frequency bands (frequencies up to 1.5 kHz). (4) The bottom parameters can be estimated from the short range 2 km track (where variability and range dependence is less problematic) and held constant for all the other source-receiver ranges up to 10 km.

The attempt to decrease the array aperture showed that at 5 km source range, it was possible to achieve nearly correct localizations with an array 4 times smaller than the initial full aperture array, sampling only $\frac{1}{6}$ of the water column. Using an increased number of frequencies allowed reducing the array aperture even further using only four sensors ($\frac{1}{12}$ of the water column).

The results reported in this paper indicate that under controlled conditions and in a shallow water environment of low range dependence, it is possible to accurately model the acoustic field for ranges up to 10 km and frequencies up to 1.5 kHz to correctly localize an acoustic source over time if the optimized modeling process has taken into account possible time variabilities in the water column. Moreover, these results indicate that the MFP technique can be made robust to model mismatch caused by the water column time dependence, if an adequate focalization procedure is incorporated, even at high frequencies and ranges up to 10 km. This indicates that it may be feasible to carry out source localization at frequencies higher than typically considered for MFP even in the presence of strong environmental fluctuations such as internal tides, eddies, and fronts, if sufficient *a priori* information on the water column is available.

ACKNOWLEDGMENTS

The authors would like to thank SACLANTCEN for the support given during this study, in particular to Jürgen Sellschopp (the chief scientist of the ADVENT'99 cruise), Peter Nielsen and the NRV Alliance crew. This work was partially supported by PRAXIS XXI, FCT, Portugal, under the ATOMS project, Contract No. PDCTM/P/MAR/15296/1999.

[1] H. P. Bucker, "Use of calculated sound fields and matched-field detection to locate sound source in shallow water," J. Acoust. Soc. Am. **59**, 368–373 (1976).
[2] M. Hinich, "Maximum likelihood signal processing for a vertical array," J. Acoust. Soc. Am. **54**, 499–503 (1973).
[3] 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).
[4] R. M. Hamson and R. M. Heitmeyer, "An analytical study of the effects of environmental and system parameters on source localization in shallow water by matched-field processing of a vertical array," J. Acoust. Soc. Am. **86**, 1950–1959 (1989).
[5] S. Jesus, "Normal-mode matching localization in shallow water: environmental and system effects," J. Acoust. Soc. Am. **90**, 2034–2041 (1991).
[6] A. M. Richardson and L. W. Nolte, "*A posteriori* probability source localization in an uncertain sound speed deep ocean environment," J. Acoust. Soc. Am. **89**, 2280–2284 (1991).
[7] M. D. Collins and W. A. Kuperman, "Focalization: Environmental focusing and source localization," J. Acoust. Soc. Am. **90**, 1410–1422 (1991).
[8] D. F. Gingras and P. Gerstoft, "Inversion for geometric parameters in shallow water: Experimental results," J. Acoust. Soc. Am. **97**, 3589–3598 (1995).
[9] P. Gerstoft and D. F. Gingras, "Parameter estimation using multi-frequency range-dependent acoustic data in shallow water," J. Acoust. Soc. Am. **99**, 2839–2850 (1996).
[10] C. Soares, A. Waldhorst, and S. M. Jesus, "Matched-field processing: environmental focusing and source tracking with application to the North Elba data set," Oceans'99 MTS/IEEE Proceedings, Seattle, 1999, Vol. 3, pp. 1598–1602.
[11] M. Siderius, P. Nielsen, J. Sellschopp, M. Snellen, and D. Simons, "Experimental study of geo-acoustic inversion uncertainty due to ocean sound-speed fluctuations," J. Acoust. Soc. Am. **110**, 769–781 (2001).
[12] T. Fassbender, "Erweiterte Genetische Algorithmen zur globalen Optimierung multi-modaler Funktionen," Ruhr-Universität, Master thesis, Bochum, Germany, 1995.
[13] J. Sellschopp, M. Siderius, and P. Nielsen, "ADVENT'99, pre-processed acoustic and environmental cruise data," SACLANTCEN Undersea Research Center, Internal Report CD-35, La Spezia, Italy, 2000.
[14] M. Snellen, D. G. Simons, M. Siderius, J. Sellschopp, and P. L. Nielsen, "An evaluation of the accuracy of shallow water matched field inversion results," J. Acoust. Soc. Am. **109**, 514–527 (2001).
[15] C. M. Ferla, M. B. Porter, and F. B. Jensen, "C-SNAP: Coupled SACLANTCEN normal mode propagation loss model," SACLANTCEN Undersea Research Center, Memorandum SM-274, La Spezia, Italy, 1993.
[16] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," J. Acoust. Soc. Am. **83**, 571–587 (1988).
[17] S. Jesus, "Broadband matched-field processing of transient signals in shallow water," J. Acoust. Soc. Am. **93**, 1841–1850 (1993).

Necessary conditions for a maximum likelihood estimate to become asymptotically unbiased and attain the Cramer–Rao lower bound. II. Range and depth localization of a sound source in an ocean waveguide

Aaron Thode, Michele Zanolin, Eran Naftali, Ian Ingram, Purnima Ratilal, and Nicholas C. Makris
Ocean Engineering Department, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 17 October 2000; revised 24 September 2001; accepted 29 May 2002)

Analytic expressions for the first order bias and second order covariance of a maximum-likelihood estimate (MLE) are applied to the problem of localizing an acoustic source in range and depth in a shallow water waveguide with a vertical hydrophone array. These expressions are then used to determine *necessary* conditions on sample size, or equivalently signal-to-noise ratio (SNR), for the localization MLE to become asymptotically unbiased and attain minimum variance as expressed by the Cramer–Rao lower bound (CRLB). These analytic expressions can be applied in a similar fashion to any ocean-acoustic inverse problem involving random data. Both deterministic and completely randomized signals embedded in independent and additive waveguide noise are investigated. As the energy ratio of received signal to additive noise (SANR) descends to the lower operational range of a typical passive localization system, source range and depth estimates exhibit significant biases and have variances that can exceed the CRLB by orders of magnitude. The spatial structure of the bias suggests that acoustic range and depth estimates tend to converge around particular range and depth cells for moderate SANR values. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1496765]

PACS numbers: 43.30.Wi, 43.60.Rw, 43.60.Gk [DLB]

I. INTRODUCTION

In recent years, many acoustic techniques have been developed to probe the marine environment. These techniques typically require the nonlinear inversion of acoustic field data measured by a hydrophone array.[1–3] The data, however, are often randomized by the addition of natural ambient noise, or by fluctuations in the acoustic source, waveguide refractive index, and waveguide boundaries. Since the nonlinear inversion of random data often yields estimates with biases and mean-square errors that are difficult to quantify, it has become popular to simply neglect these potential biases and to compute limiting bounds on the mean-square error, since the bounds are usually much easier to obtain than the actual mean-square errors. The most widely used limiting bound is the Cramer–Rao lower bound (CRLB),[4] which describes the minimum possible variance of any unbiased estimator, and has been introduced in the ocean acoustic source localization literature via Refs. 5 and 6, for example. Other bounds, however, also exist in the literature,[7–11] that are not directly relevant to the present work.

The purpose of the present paper is not to apply a new general bound, but to demonstrate how the asymptotic properties of the maximum likelihood estimate (MLE) described in a companion paper[12] can be used to better understand the statistical errors and biases that occur in a typical ocean acoustic inverse problem. The MLE has a straightforward implementation. It is obtained by maximizing the likelihood function with respect to the parameter vector to be estimated, where the likelihood function is the conditional probability density of the data, given the unknown parameter vector, evaluated at the measured data values. The MLE is widely used in statistics, because if an estimator becomes asymptotically unbiased and attains the CRLB for large sample sizes or high signal-to-noise ratio (SNR), it is guaranteed to be the MLE.[4] It follows that an analysis of the conditions necessary for a MLE to become unbiased and attain minimum-variance, and thus optimal performance, will also reveal the conditions necessary for *any* nonlinear estimate to asymptotically achieve optimal performance.

In ocean-acoustic inverse problems, the likelihood function can be maximized by an exhaustive or directed search via forward modeling with numerical wave propagation or scattering algorithms. While the linear least squares estimator is also a widely used inversion scheme in acoustics and geophysics,[13] it is only identical to the MLE when the data and parameter vectors are linearly related, and when the data are uncorrelated, follow a multivariate Gaussian distribution, and share the same variance.[14] Since these conditions are often not satisfied in practice, the linear least squares estimator is often suboptimal even in the asymptotic regime of high SNR, making the MLE a preferable choice.

The present application concerns the classic ocean-acoustic inverse problem of localizing a source in range and depth in a shallow water waveguide, using data received on a vertical hydrophone array,[3] also known as the ''matched-field

processing'' (MFP) source localization problem.[2,16] [There are differences between some standard matched field processors and the MLE that have sometimes gone unnoticed in the MFP literature. For example, the ''minimum-variance distortionless response (MVDR) processor,'' which is presented as a MLE in Refs. 3 and 5, is neither a minimum-variance estimator nor the MLE for range and depth localization in a waveguide, as discussed by Sullivan and Middleton.[15] The MVDR processor is a MLE, however, for the particular problem of estimating the complex amplitude of a plane wave arriving from a *known* direction, in independent additive Gaussian noise.] Here the theory presented in the companion paper[12] is used to set conditions on the sample size and SNR *necessary* for the MLE to become asymptotically unbiased and attain the CRLB in MFP source localization. These conditions can also be used in experimental design and analyses to ensure that statistical biases and errors are maintained within tolerable limits set by the given scientific or engineering objective. The approach follows that given in Ref. 12 and is based on the fact that the MLE can be expanded as an asymptotic series in inverse orders of sample size[17] or equivalently an appropriately defined SNR. From this series, analytic expressions for the first order bias and second order error correlation of a general MLE can be found in terms of the joint moments of parameter derivatives of the log-likelihood function.[12] Since the first order error correlation is the CRLB, which is only valid for unbiased estimates, the second order error correlation can provide a better estimate of the MLE mean-square error that is applicable at relatively low SNR, even when the MLE is biased to first order. Necessary conditions for asymptotic optimality of the MLE are then obtained by demonstrating when the first order bias becomes negligible compared to the true value of the parameter, and when the second order error correlation becomes negligible compared to the CRLB.[12]

In this paper the localization performance of the MLE for both deterministic and randomized monopole signals embedded in independent, additive waveguide noise is analyzed. As the energy ratio of received signal to additive noise (SANR) descends to the lower operational range of a typical passive localization system, the range and depth MLE demonstrates significant bias and has a mean-square error that exceeds the CRLB by orders of magnitude.

II. ASYMPTOTIC STATISTICS

A. Preliminary definitions

Following the theory and notation adopted in paper I,[12] let the random data vector $\mathbf{X}$, given m-dimensional parameter vector $\boldsymbol{\theta}$, obey the conditional probability density function (PDF) $p(\mathbf{X};\boldsymbol{\theta})$. The log-likelihood function $l(\boldsymbol{\theta})$ is then defined as $l(\boldsymbol{\theta})=\ln(p(\mathbf{X};\boldsymbol{\theta}))$, when evaluated at the measured values of $\mathbf{X}$. The first-order parameter derivative of the log-likelihood function is then defined as $l_r=\partial l(\boldsymbol{\theta})/\partial\theta^r$, where θ^r is the rth component of $\boldsymbol{\theta}$. Moments of the log-likelihood derivatives are defined by $v_R\equiv E[l_R]$, where R is an arbitrary set of indices. If $R_1=r_{11}\ldots r_{1n_1},\ldots,\ R_m=r_{m1}\ldots r_{mn_m}$ are sets of coordinate indices in the parameter space, joint moments of the log-likelihood derivatives can be defined by $v_{R_1,R_2,\ldots,R_M}\equiv E[l_{R_1}l_{R_2}\cdots l_{R_M}]$, where, for example, $v_{s,tu}=E[l_s l_{tu}]$ and $v_{a,b,c,de}=E[l_a l_b l_c l_{de}]$.

The expected information, known as the Fisher information, is defined as $i_{rs}=E[l_r l_s]$, for arbitrary indices r,s.[4] Lifting the indices produces quantities that are defined as[17]

$$v^{R_1,R_2,\ldots,R_m}=i^{r_{11}s_{11}}i^{r_{12}s_{12}}\cdots i^{r_{1n_1}s_{1n_1}}i^{r_{21}s_{21}}i^{r_{22}s_{22}}\cdots i^{r_{mn_m}s_{mn_m}}\times v_{s_{11}s_{12}\cdots s_{1n_1},s_{21}s_{22}\cdots s_{mn_m}},\tag{1}$$

where $i^{rs}=[\mathbf{i}^{-1}]_{rs}$. Here, as elsewhere, the Einstein summation convention is used, so that whenever an index appears as both a superscript and subscript in a term, summation over that index is implied. The Fisher information matrix, $\mathbf{i}$, has an inverse $\mathbf{i}^{-1}$, known as the Cramer–Rao lower bound (CRLB),[4,14,18] which is a lower bound on the minimum variance an unbiased estimator can attain.

B. General asymptotic expansions for the bias and covariance of the MLE

With the notation presented in Sec. II A, the first-order bias of the MLE can be written as[12,17]

$$b(\hat{\theta}^r)=E[(\hat{\theta}-\theta)^r]=\underbrace{\tfrac{1}{2}i^{ra}i^{bc}(v_{abc}+2v_{a,bc})}_{O_p(n^{-1})}+O_p(n^{-2})+\cdots,\tag{2}$$

where the symbol $O_p(n^{-m})$ denotes a polynomial of exactly order n^{-m}, where n is the sample size. It is noteworthy that third derivatives of the log-likelihood function may be necessary to compute the first-order bias.

A *necessary* condition for the MLE to become asymptotically unbiased is for the first-order terms in Eq. (2) to become much smaller than the true value of the parameter θ^r. Equation (2) may then be used to determine the minimum sample size n necessary for the MLE to become effectively unbiased.

An expression for the asymptotic covariance of the MLE has been derived by Naftali and Makris in paper I,[12] who obtained the first two asymptotic orders of the MLE covariance as

$$
\begin{aligned}
\operatorname{cov}(\hat{\theta}^r,\hat{\theta}^a) &= E[(\hat{\theta}^r - E[\theta^r])(\hat{\theta}^a - E[\theta^a])] \\
&= \underbrace{[i^{ra}]}_{O_p(n^{-1})} + \underbrace{[2i^{mb}i^{nc}v_{lmn}(r^{rs}i^{la}+i^{as}i^{lr})v_{s,b,c}(n^1)+\tfrac{1}{2}i^{cd}i^{ef}(i^{rs}i^{ab}+i^{as}i^{rb})v_{bce,d,f,s}(n^2)}_{O_p(n^{-2})} \\
&+ \underbrace{i^{tu}(i^{rs}i^{ab}i^{cd}+i^{rd}i^{ab}i^{cs}+i^{ad}i^{rb}i^{cs})v_{st,u,bc,d}(n^2)}_{O_p(n^{-2})} \\
&+ \underbrace{i^{bm}i^{cq}i^{tp}v_{lmn}v_{opq}(\tfrac{1}{4}r^{rl}i^{ao}i^{sn}+\tfrac{1}{2}i^{rs}r^{al}i^{on}+\tfrac{1}{2}i^{as}i^{rl}i^{on})v_{s,t,b,c}(n^2)}_{O_p(n^{-2})} \\
&+ \underbrace{\tfrac{1}{2}i^{sm}v_{lmn}(i^{tn}i^{cd}(i^{rl}i^{ab}+i^{al}i^{rb})+2i^{bn}i^{cd}(i^{rl}i^{at}+i^{al}i^{rt})+i^{cl}i^{tn}(i^{rd}i^{ab}+i^{ad}i^{rb}))v_{s,t,bc,d}(n^2)}_{O_p(n^{-2})} \\
&+ \underbrace{\tfrac{1}{6}i^{mb}i^{nc}i^{od}v_{lmno}(i^{rs}i^{la}+i^{as}i^{rl})v_{s,b,c,d}(n^2)+4i^{bm}(i^{rs}i^{al}+i^{as}i^{rl})v_{s,m,lb}(n^1)}_{O_p(n^{-2})} \\
&- \underbrace{\tfrac{1}{4}i^{rs}i^{tu}i^{aw}i^{yz}v_{stu}v_{wyz}-i^{rs}i^{tu}i^{aw}i^{yz}v_{stu}v_{wy,z}(n^1)-i^{rs}i^{tu}i^{aw}i^{yz}v_{st,u}(n^1)v_{wy,z}(n^1)]}_{O_p(n^{-2})}.
\end{aligned}
\tag{3}
$$

The notation $v_{bce,d,f,s}(n^2)$ indicates that in the joint moment $v_{bce,d,f,s}$ only polynomial terms of order n^2 are retained. The first-order covariance term of this expansion is simply the CRLB, and the sum of the other 10 terms provides the second-order variance. The ratio of this sum to the first-order variance (CRLB) is hereafter defined as the ''second-to-first order variance ratio'' (SFOVR), which is inversely proportional to sample size *n*. A necessary criterion for the MLE to attain minimum variance is for this ratio to become negligible. Note that this is only a *necessary* condition to attain minimum variance, and not a *sufficient* condition, because there is no guarantee that higher-order variance terms neglected in Eq. (3) will not exceed the second-order variance for sufficiently small *n*. For sufficiently large *n*, the series is guaranteed to converge if a minimum variance and unbiased estimate exists. In many practical scenarios the necessary conditions specified here are also sufficient for establishing optimality.

C. Gaussian data, deterministic, and random signals

The general bias and variance expressions of Eqs. (2) and (3) are now applied to the specific case of data that obey the conditional Gaussian probability density[14]

$$
\begin{aligned}
p(\mathbf{X};\boldsymbol{\theta}) &= \frac{1}{(2\pi)^{nN/2}|\mathbf{C}(\boldsymbol{\theta})|^{n/2}} \\
&\times \exp\left\{-\frac{1}{2}\sum_{j=1}^{n}(\mathbf{X}_j-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\mathbf{C}(\boldsymbol{\theta})^{-1}(\mathbf{X}_j \right. \\
&\left. -\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}.
\end{aligned}
\tag{4}
$$

Here $\mathbf{X}_i$ is one of *n* independent and identically distributed *N*-dimensional real-valued data vectors, and $\mathbf{X}=[\mathbf{X}_1^T,\mathbf{X}_2^T,\ldots,\mathbf{X}_n^T]$, $\mathbf{C}$ is the real-valued covariance matrix, and $\boldsymbol{\mu}$ is the real-valued mean of the real random data. For the present study of underwater localization, $\mathbf{X}_i$ represents the real and imaginary parts of the narrow-band acoustic data collected across an array of $N/2$ sensors around the given harmonic-source frequency, and the parameter set $\boldsymbol{\theta}$ represents the range and depth of the acoustic source. The fact that the number of hydrophones is half the length of the data vector $\mathbf{X}_i$ follows from the use of this real vector to describe complex data as will be discussed in the next section.

In general both the data mean and covariance in Eq. (4) are functions of the desired parameter set $\boldsymbol{\theta}$, a situation that makes evaluation of the joint moments in Eqs. (2) and (3) difficult. Two limiting cases, however, are of great practical interest, since they describe a deterministic signal in additive noise and a completely randomized signal in noise, respectively. In the deterministic scenario the covariance matrix $\mathbf{C}$ is independent of the parameter vector $\boldsymbol{\theta}$, while the mean $\boldsymbol{\mu}$ depends on $\boldsymbol{\theta}$. In the randomized scenario the sample covariance $\mathbf{C}$ depends on $\boldsymbol{\theta}$ and the data mean $\boldsymbol{\mu}$ is zero. For this latter case the sample covariance of the data is thus a sufficient statistic that contains all information about the estimation parameters contained in the measured data.[4,19]

Deterministic signals are typically measured in the presence of independent, additive noise in ocean acoustics. These deterministic signals may arise from narrow-band sources on seagoing vessels, or from deployed tomographic sources. A narrow-band time series from such a combination of signal and noise would have the form $z(t)\exp(i2\pi f_c t)+\xi(t)$, when measured at a single omni-directional receiver, where $\xi(t)$ is the independent, additive noise, $z(t)$ is a deterministic envelope, and f_c is the carrier frequency. The additive noise $\xi(t)$ typically arises from a large number of independent sources distributed over the sea surface.[20] These noise sources may be generated by wind, wave interactions, or ocean-going ves-

sels. Since the total noise field is the sum of large numbers of statistically independent contributions, it follows a Gaussian probability density, by virtue of the central limit theorem. The Gaussian probability density of Eq. (4) is then a valid representation of a deterministic signal embedded in additive noise.

The second limiting case explored here is for a fully randomized signal. A particular fully randomized Gaussian signal model that is very widely used and enjoys a long history in acoustics, optics, and radar is the circular complex Gaussian random (CCGR) model.[21] A scalar signal centered around a carrier frequency f_c that obeys CCGR statistics would have the form $z(t)\exp(i2\pi f_c t)+\xi(t)$, where $\xi(t)$ is the independent, additive noise, and where $z(t)$ is an envelope whose real and imaginary parts are independent Gaussian random variables with zero mean, and equal variance. The instantaneous intensity of this signal thus obeys an exponential probability density function(PDF).[22,23] In the radar literature the Swerling II model for radar returns from a fluctuating target is equivalent to the CCGR model, since the instantaneous intensity received in that case also follows an exponential PDF.[23,24] There are many physical mechanisms for generating ocean acoustic signals with CCGR properties. Various types of mechanical and propeller noise generated by the complex source distribution of a ship or submarine generate incoherent source fluctuations that can be represented as a CCGR process in time. Even when the initial source signal is deterministic, natural disturbances in the waveguide, such as underwater turbulence or passing internal or gravity waves, lead to such randomness in the medium that the waveguide modes at the receiver can be treated as statistically independent entities. The total received field, which is the weighted sum of these modes, can then be modeled as a CCGR process in time. Randomized ocean acoustic signals have been modeled with CCGR statistics since World War II,[22,25] and as a consequence the CCGR model has become a standard assumption for analyses of MFP performance.[5,15,26,27]

Once $\mathbf{C}$ and $\boldsymbol{\mu}$ have been obtained from either signal model, the joint moments are evaluated and inserted into Eqs. (2) and (3), to compute the asymptotic bias and covariance. The appropriate formulas for the joint moments are provided in Appendixes B and C in the companion paper,[12] for both deterministic and fully randomized signals, respectively.

III. WAVEGUIDE, SIGNAL AND NOISE MODELS

Four simple shallow-water ocean waveguide environments have been modeled to examine the effects of varying bottom composition and sound-speed profile on the bias and variance of a MLE for the location of an acoustic source.

Figure 1 displays the selected array geometry, sound speed profiles, and bottom composition for each environment. All simulations employ a 100 Hz monopole source at 50 m depth, and a 10-element vertical array with 7.5 m spacing, in a 100 m deep range-independent waveguide. The shallowest element lies at 16.25 m depth, so that the array is

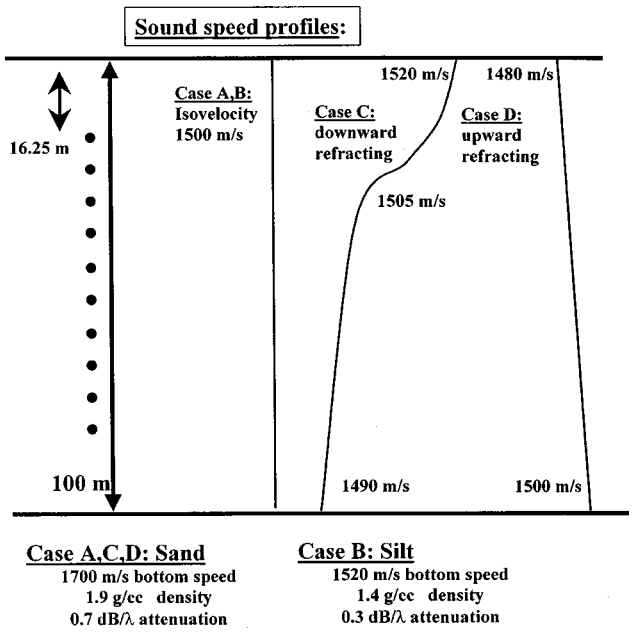


FIG. 1. Illustration of the ocean waveguide environments used in the paper. All cases employ a 100 m deep water column overlying a semi-infinite fluid half space, and a 10-element vertical array with 7.5 m spacing, with the first element positioned at 16.25 m depth. Cases A and B employ an isovelocity water profile of 1500 m/s (Pekeris profile). Case A uses geoacoustic parameters representative of a sand bottom, while case B uses parameters representative of a silt bottom. Case C uses the sand bottom, and a downward-refracting water sound–speed profile to simulate summer conditions in temperate latitudes. Case D is similar, except a linear upward-refracting profile is used to represent an arctic scenario.

centered in the waveguide. The ocean bottom is a fluid half space.

The first environment, case A, is a Pekeris waveguide with a bottom sound speed of 1700 m/s, a density of 1.9 g/cc, and an attenuation of 0.8 dB/wavelength, which are representative values for sandy environments.[3] In case B, a silt bottom is simulated, using a bottom speed of 1520 m/s, density of 1.4 g/cc, and an attenuation of 0.3 dB/wavelength. Case C retains the sand bottom parameters, but uses a downward-refracting water sound-speed profile, measured during the Swellex-93 experiments[28,29] conducted off the San Diego coast in 1993, under typical oceanic summer conditions in temperate latitudes.[29] Finally, case D illustrates the effects of propagation through an upward-refracting sound speed profile. The profile linearly decreases from 1500 m/s at the ocean bottom to 1480 m/s at the surface.

For ranges greater than a few ocean depths, the Green function for the acoustic field received by the mth hydrophone array element from a monopole source at horizontal range r and depth z at angular frequency ω can be expressed as a sum of normal modes[3]

$$\tilde{g}(z_m,z,r,\omega)=\frac{ie^{-i\pi/4}}{\rho(z)\sqrt{8\pi r}}\sum_l \Psi_l(z)\Psi_l(z_m)\frac{e^{ik_l r}}{\sqrt{k_l}}, \quad (5a)$$

where k_l is the horizontal wave number of mode l with modal amplitude $\Psi_l(z)$. Equation (5a) defines the mth element of the spatial vector $\tilde{\mathbf{g}}$ for $m=1,2,3,...,N/2$, where $N/2$

is the number of hydrophones in the receiving array. For a source with amplitude $A_j(\omega)$, the jth sample or snapshot of the received field measured across the entire hydrophone array is comprised of the components of both a complex signal vector $A_j(\omega)\tilde{\mathbf{g}}$ and an additive CCGR noise vector $\tilde{\boldsymbol{\eta}}_j$ such that

$$\mathbf{X}_j = \begin{bmatrix} \mathrm{Re}(A_j(\omega)\tilde{\mathbf{g}}) \\ \mathrm{Im}(A_j(\omega)\tilde{\mathbf{g}}) \end{bmatrix} + \boldsymbol{\eta}_j \tag{5b}$$

for $j=1,2,3,...,n$ where

$$\mathbf{g} = \begin{bmatrix} \mathrm{Re}(\tilde{\mathbf{g}}) \\ \mathrm{Im}(\tilde{\mathbf{g}}) \end{bmatrix}, \quad \eta_j = \begin{bmatrix} \mathrm{Re}(\tilde{\boldsymbol{\eta}}_j) \\ \mathrm{Im}(\tilde{\boldsymbol{\eta}}_j) \end{bmatrix}, \tag{5c}$$

and $\mathbf{X}_j$ are real vectors of length N. In the deterministic signal model $A_j(\omega)$ is a constant $A(\omega)$, for all j. In the random signal model the $A_j(\omega)$ are independent and identically distributed CCGR variables that describe a stationary random process where $\langle A_j(\omega)\rangle = 0$ and $\langle |A_j(\omega)|^2\rangle = \langle |A(\omega)|^2\rangle$ for $j=1,2,3,...,n$.

The portion of the covariance matrix due to additive background noise for a single sample $n=1$, or a single snapshot, of data across the vertical array is assumed to be spatially white, since this is typically what is measured in continental shelf environments,

$$\tilde{\mathbf{C}}^{\text{additive}} = \langle \tilde{\boldsymbol{\eta}}_j \tilde{\boldsymbol{\eta}}_j^{+} \rangle - \langle \tilde{\boldsymbol{\eta}}_j \rangle \langle \tilde{\boldsymbol{\eta}}_j^{+} \rangle = \sigma^2 \mathbf{I}, \tag{6}$$

where $\mathbf{I}$ is the identity matrix, σ^2 is the instantaneous variance of the additive noise on each hydrophone and the $+$ superscript represents a Hermitian transpose. Computations using the spatially correlated Kuperman and Ingenito waveguide noise model[30] for the given environments are not shown here, but produce results similar to those derived from Eq. (6) since the theory predicts weak spatial correlation.

In general, for a properly defined SNR, terms in Eqs. (2) and (3) that are of order n^{-m} must also be of order SNR^{-m}. In the matched field processing literature, the signal to additive noise ratio (SANR) is typically used, which is not necessarily the SNR nor is it necessarily proportional to the SNR. For a single sample, $n=1$, we define the SANR as

$$\mathrm{SANR}[1]_{z_s,r} = \frac{\sum_{u=1}^{N/2} \langle |A(\omega)|^2 \rangle |\tilde{g}(z_u, z_s, r, \omega)|^2}{\mathrm{trace}\ (\tilde{\mathbf{C}}^{\text{additive}})} \tag{7}$$

which is a function of source range r and depth z. For multiple independent and identically distributed samples in the deterministic signal model, $\mathrm{SANR} = n\mathrm{SANR}[1]$, for a given source range and depth. Moreover, terms in Eqs. (2) and (3) that are of order n^{-m} are also of order SANR^{-m} when the additive noise is zero-mean Gaussian and spatially *uncorrelated*, in the deterministic signal model, as shown in appendix B. This is not the case when the noise is spatially *correlated*. In the correlated case, terms in Eqs. (2) and (3) must be expanded in powers of a more generalized quantity that cannot simply be factored into the ratio of a signal term and a noise term. The tenant that SANR and SNR should conform to such a factorable ratio must be abandoned in this more generalized framework.

For a completely randomized signal in additive noise, the SANR and SNR are not equal, however, for consistency we still define $\mathrm{SANR} = n\mathrm{SANR}[1]$ in this case. Consider, for example, the case of a scalar measurement of a randomized signal with no additive noise. Here the SNR is independent of signal intensity because the variance of an instantaneous intensity measurement equals the square of the expected intensity. The SNR, which is the ratio of these two quantities, is unity for an instantaneous intensity measurement.[22] The SANR of this same signal, however, is infinite, because no additive noise is present. More generally, for finite time measurements of intensity, the SNR for a scalar measurement of a completely randomized signal is defined as the ratio of the square of the expected intensity to the variance of this intensity, as has been described in detail in Refs. 22 and 19. Under such circumstances the SNR is then approximately equal to the number of independent samples n in a measurement time T where asymptotic convergence to n occurs for $n \gg 1$. For measurement times much greater than the coherence time τ_c for fluctuations in the received field, n is well approximated by T/τ_c.[19,21,22] The number of fluctuations n is also equivalent to the time-bandwidth product of the received signal, where the bandwidth of the fluctuating field is $1/\tau_c$ in the limit as $T \gg \tau_c$.

For example, if a fast-Fourier transform (FFT) is applied to a data segment from a completely randomized signal of duration T, the effective number of independent samples n is T times the signal bandwidth, which is another way of stating that the signal is expected to fluctuate n times during the measurement. If T is much shorter than the coherence time τ_c of the random signal, then this measurement represents a single statistical sample where $n=1$.

While the SANR has traditionally been the quantity of practical interest in MFP, the difference between SANR and SNR is important because a signal with high SANR might still have low SNR, due to signal-dependent fluctuations during measurements with low time-bandwidth products. All the randomized signal examples in this paper are computed for an instantaneous measurement, where $n=1$ and the SNR is 1. Measuring the data over longer periods leads to a linear increase in the sample size n if the signal is deterministic or a nearly linear increase in n which is asymptotically linear for $n \gg 1$ if the signal can be described as a CCGR process.

Since SANR is a function of both source range and depth, for consistency, we adopt the convention of setting the SANR[1] of the field across the array to unity *for a source located at r=1 km range and any depth z* for all simulations presented in this paper. This implies that the SANR is made constant over source depth for any fixed range separation between the source and receiver array by appropriately varying the source amplitude with source depth.

We define the complex mean $\tilde{\boldsymbol{\mu}}$ and covariance $\tilde{\mathbf{C}}$ that are related to the real mean $\boldsymbol{\mu}$ and covariance $\mathbf{C}$ of Eq. (4) by the following expressions:[14]

$$\boldsymbol{\mu} = \begin{bmatrix} \mathrm{Re}(\tilde{\boldsymbol{\mu}}) \\ \mathrm{Im}(\tilde{\boldsymbol{\mu}}) \end{bmatrix}, \quad \mathbf{C} = \frac{1}{2} \begin{bmatrix} \mathrm{Re}(\tilde{\mathbf{C}}) & -\mathrm{Im}(\tilde{\mathbf{C}}) \\ \mathrm{Im}(\tilde{\mathbf{C}}) & \mathrm{Re}(\tilde{\mathbf{C}}) \end{bmatrix} \tag{8}$$

which are valid under the assumption that the complex data measured at each hydrophone follow a circularly complex Gaussian random density[21] when the mean is subtracted.

When modeling deterministic signals, the complex mean vector becomes $\tilde{\boldsymbol{\mu}}=A(\omega)\tilde{\mathbf{g}}$, and the complex covariance matrix $\tilde{\mathbf{C}}$ is taken to be $\tilde{\mathbf{C}}^{\text{additive}}$, for a single sample $n=1$.

When modeling randomized signals due to random source amplitude $A_j(\omega)$, $\tilde{\boldsymbol{\mu}}$ is set to zero, and $\tilde{\mathbf{C}}$ is defined for a single sample $n=1$ as

$$\tilde{\mathbf{C}}=\langle|A(\omega)|^2\rangle\tilde{\mathbf{g}}\tilde{\mathbf{g}}^{+}+\tilde{\mathbf{C}}^{\text{additive}}. \tag{9}$$

It is noteworthy that $\tilde{\mathbf{C}}^{-1}$ becomes singular when the additive noise vanishes in the random signal model given by Eq. (9) when $\tilde{\mathbf{g}}$ is *not* a scalar. For sufficiently high SANR, the error in estimating range and depth from *nonscalar* measurements of a randomized signal then approaches zero as the additive noise approaches zero, as is shown in Appendix B. If the signal randomness is due to fluctuations in the waveguide rather than at the source, Eq. (9) is not an appropriate model. A more appropriate model would replace $\tilde{\mathbf{g}}\tilde{\mathbf{g}}^{+}$ in Eq. (9) with a diagonal matrix that has the same diagonal elements as $\tilde{\mathbf{g}}\tilde{\mathbf{g}}^{+}$, as might arise from equipartion of modes in a fully saturated waveguide.[19,22] This random signal model due to waveguide fluctuations would be well defined even in the absence of additive noise.

Note that the definition presented in Eq. (7) does not account for potential improvements in the SANR from array gain, which is $10\log(N/2)$ for the ideal case of a plane wave signal embedded in spatially uncorrelated white noise. The 10-element array modeled here could then have an ''array-gain-augmented SANR'' that is up to 10 dB greater than that indicated by Eq. (7) for the given array. For this reason, it will sometimes be necessary to distinguish the SANR of Eq. (7) as ''input SANR'' as opposed to ''array-gain-augmented SANR.''

In the given signal models, evaluation of Eqs. (2) and (3) requires knowledge of the higher-order derivatives of $\mathbf{g}$ with respect to parameters r and z. The normal-mode depth derivatives must be computed to obtain the Green function depth derivatives. Since numerical differentiation of the modes can lead to instabilities, the modes at a given source depth are decomposed instead into upward and downward-propagating plane waves, so that depth derivatives can be derived by analytic differentiation. This procedure is discussed in detail in Appendix A.

If source amplitude is sought as well as source position, a three-dimensional parameter estimation problem must be solved. However, as shown in Appendix B, the additional uncertainties introduced by estimating the source amplitude do not noticeably affect the localization performance of either deterministic or randomized signals. One reason behind this is that the source amplitude parameter $A(\omega)$ for the standard monopole source assumed in MFP[3,5,28] is linearly related to the measured data, which in this case is the complex pressure field across the array. Because the second-order derivative of the measurement with respect to amplitude is zero, many of the higher-order joint moments that appear in Eq. (2) and Eq. (3) are zero, and those that remain do not noticeably alter the localization estimates, as illustrated in the Appendix B figures.

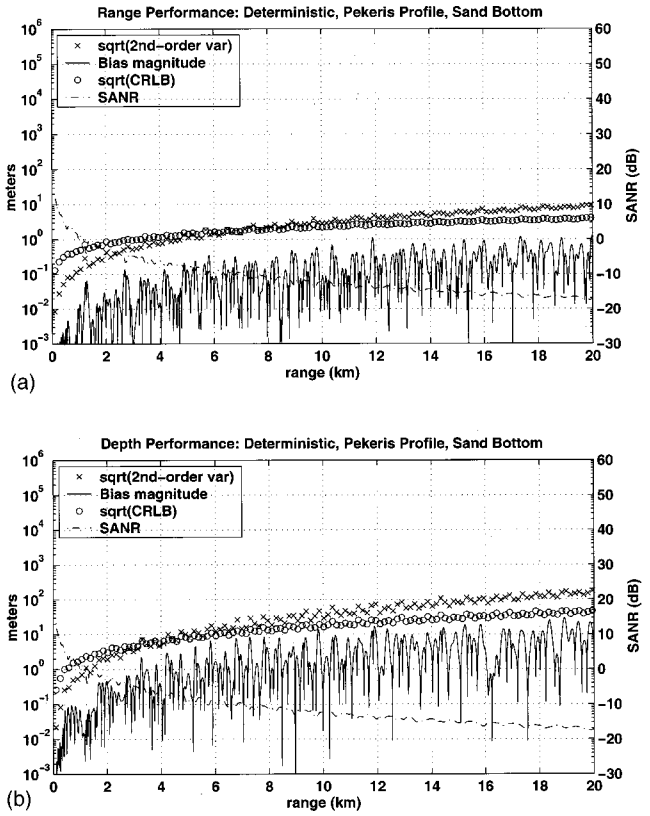
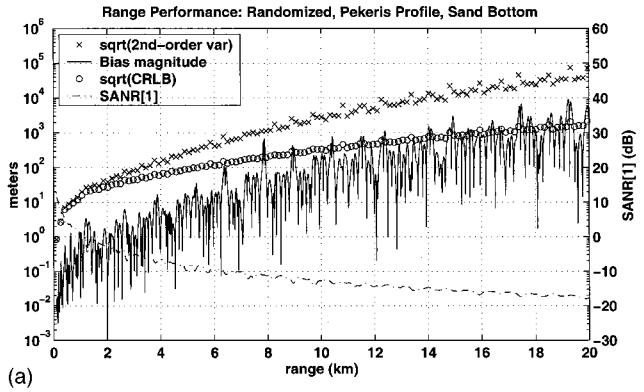


FIG. 2. Deterministic ocean acoustic localization MLE performance for (a) range estimation and (b) depth estimation versus source range, for a 100 Hz source placed at 50 m depth in the case A environment, a Pekeris waveguide with a sand bottom. The first-order bias magnitude (solid line), the square root of the Cramer–Rao lower bound (circles), the square root of the second-order variance (crosses), and the average input signal-to-additive noise ratio (SANR, dashed–dotted line) into the array are shown. Note that the input SANR plotted here does not incorporate array gain effects. All quantities are expressed in units of meters, except for the SANR, which is plotted in dB units. The background noise level has been scaled so that the input SANR is 0 dB at 1 km source range. Whenever the second-order variance attains roughly 10% of the CRLB, the total variance of the estimate will not attain the CRLB.

IV. ILLUSTRATIVE EXAMPLES

The asymptotic biases and variances depend on the source signal characteristics, measurement geometry, SANR and SNR, and surrounding propagation environment. To isolate and illustrate these contributions, a number of simulations are performed. First, the source level and source depth are held fixed, and the first-order bias and second-order variance are computed as a function of source-receiver range for various waveguide environments, using both deterministic and randomized signals. Localization estimates are shown to degrade rapidly as (10 log of) the array-gain augmented SANR descends below 0 dB. (When SANR or SNR are discussed in decibels, 10 log of SANR is assumed.) Next, the minimum sample size necessary for the MLE to attain the CRLB is computed. A fundamental difference between the sample size requirements for randomized and deterministic



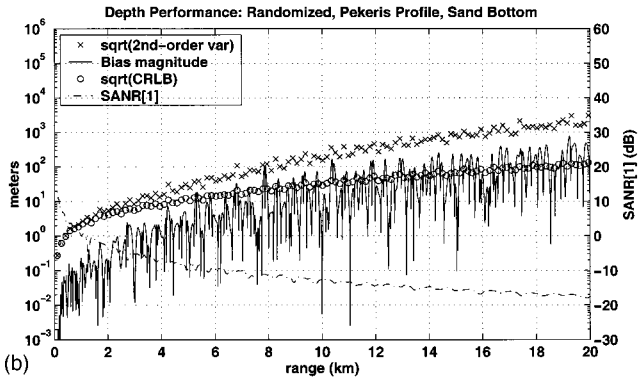
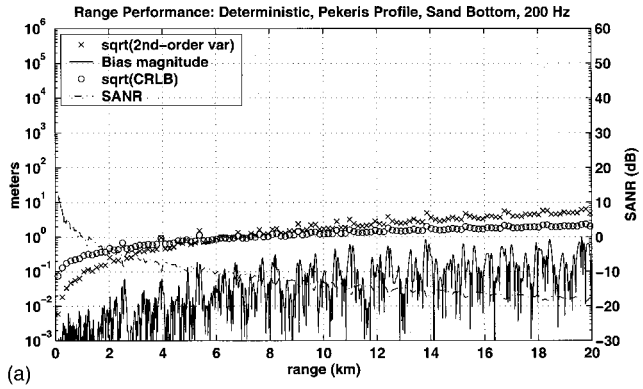


FIG. 3. Same as Fig. 2, recomputed for a single sample ($n=1$) of a completely randomized signal.



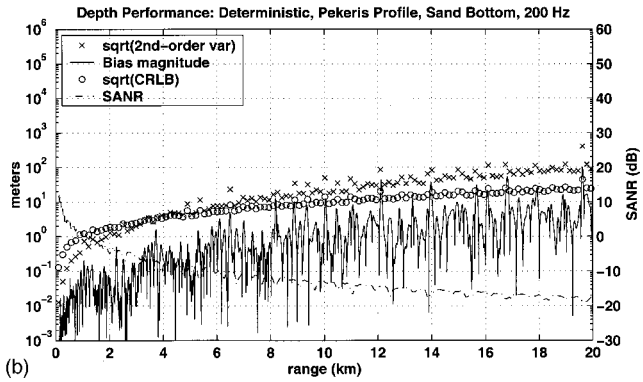


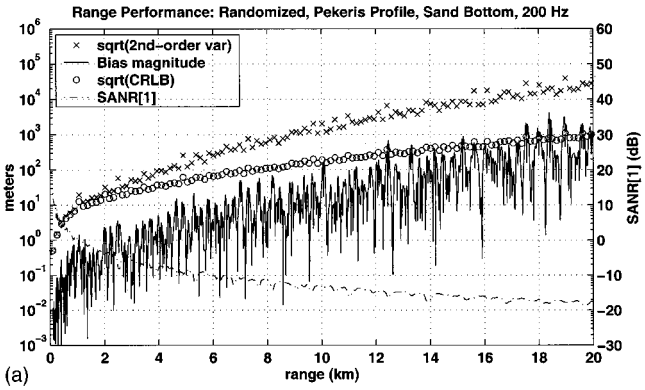
FIG. 4. Same as Fig. 2, except the deterministic source signal now radiates at 200 Hz.

signals is demonstrated. Finally, the spatial structures of the asymptotic bias and variance are investigated. It is shown that range and depth estimates tend to be biased toward certain waveguide locations as the array-gain augmented SANR falls below 0 dB.

In all cases where the deterministic signal model is used, the bias and variance terms are plotted as a function of $\mathrm{SANR}=n\mathrm{SANR}[1]$, since these terms are all proportional to SANR^{-m}, as shown in Appendix B, where m is the integer order of the term, i.e., m is either 1 or 2 in the illustrative examples. In all cases where the random signal model is used, the bias and variance terms are plotted as a function of SANR[1], since these terms depend on n and SANR[1] in different ways, as discussed in Appendix B. Since in both the deterministic and random signal models, the bias and variance terms always depend on n^{-m}, where m is the integer order of the term, all one must do to obtain a result for arbitrary n in any of these figures is to shift the curve according to the power m of the term involved and the value of n desired for a given SANR[1].

A. Localization performance versus range and waveguide environment

The magnitude of the first-order bias, the CRLB, and second-order variance typically follow an increasing trend as a function of source range in Figs. 2–11, where the source depth is fixed at 50 m, the frequency is 100 Hz, and the sample size is unity, $n=1$. The input SANR for the deterministic signal examples and SANR[1] for the random signal examples, computed from Eq. (7), are shown as a



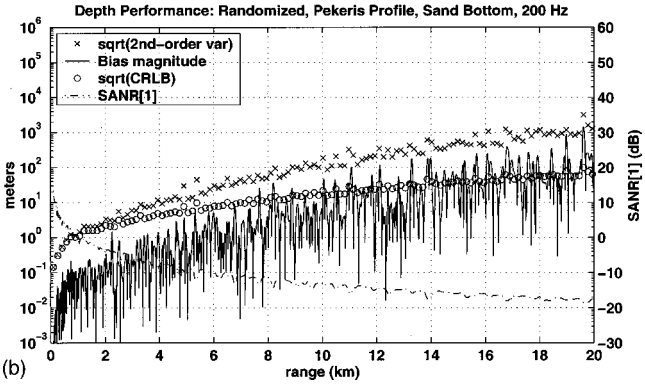
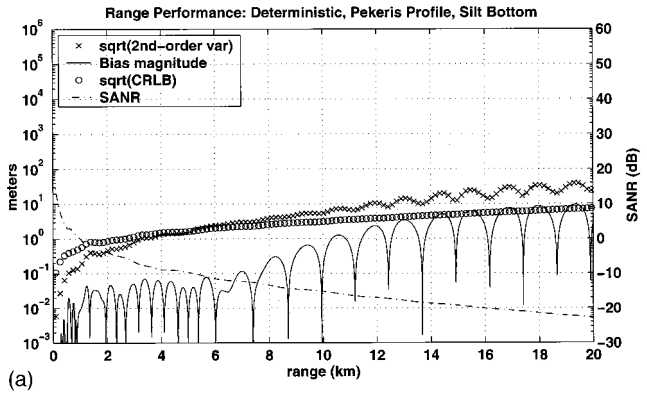


FIG. 5. Same as Fig. 3, except the randomized source signal now radiates at 200 Hz center frequency.



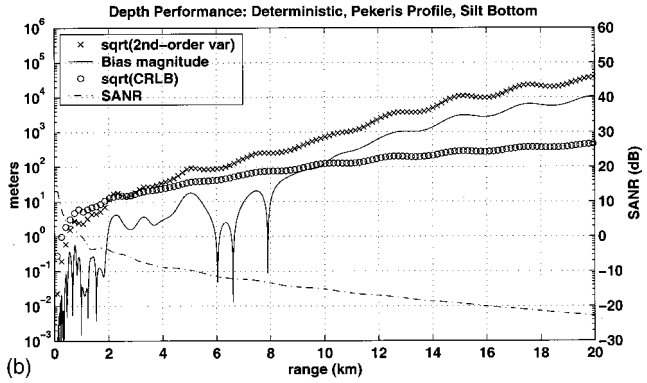
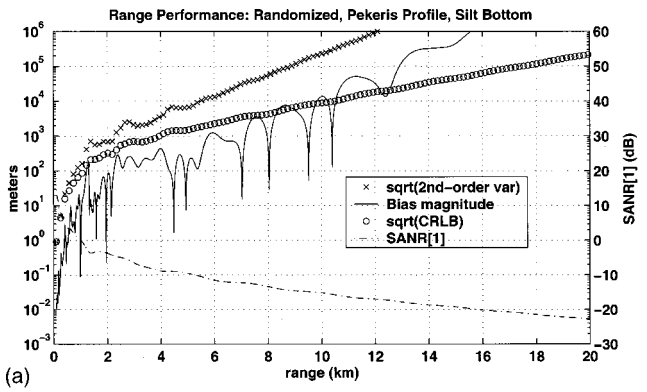


FIG. 6. Deterministic ocean acoustic localization MLE performance for (a) range estimation and (b) depth estimation versus source range, for a 100 Hz signal placed at 50 m depth in the case B environment, a Pekeris waveguide with a silt bottom. See Fig. 2 caption for plot descriptions.



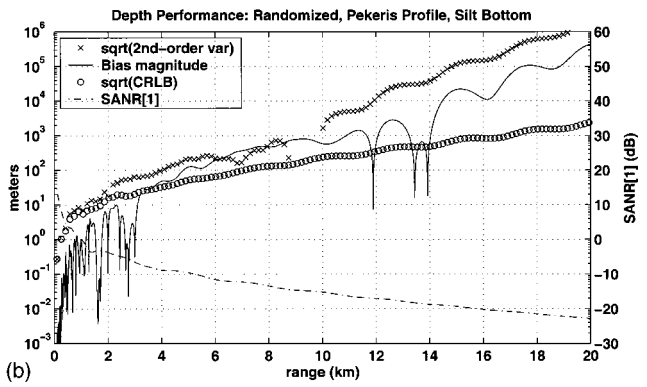
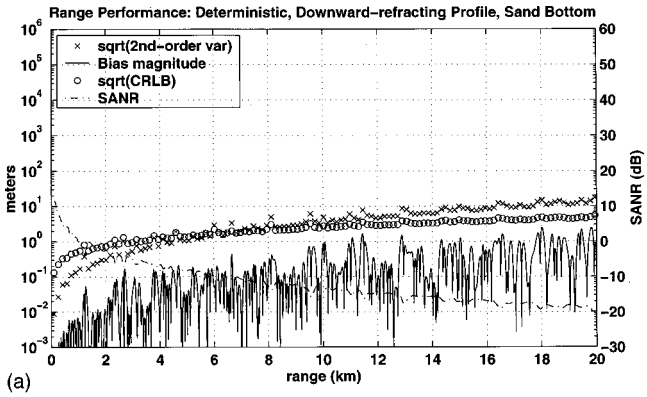


FIG. 7. Same as Fig. 6, recomputed for a single sample of a completely randomized signal.



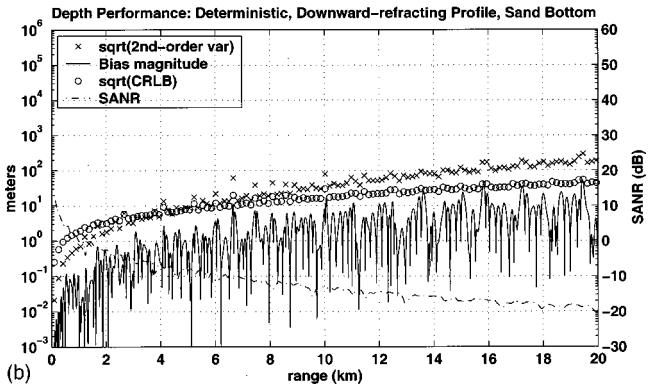
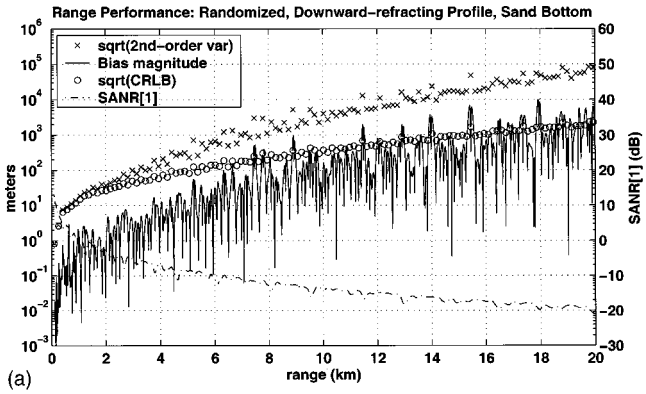


FIG. 8. Deterministic ocean acoustic localization MLE performance for (a) range estimation and (b) depth estimation versus source range, for a 100 Hz signal placed at 50 m depth in the case C environment, which consists of a downward-refracting profile over a sand bottom. See Fig. 2 caption for plot descriptions.



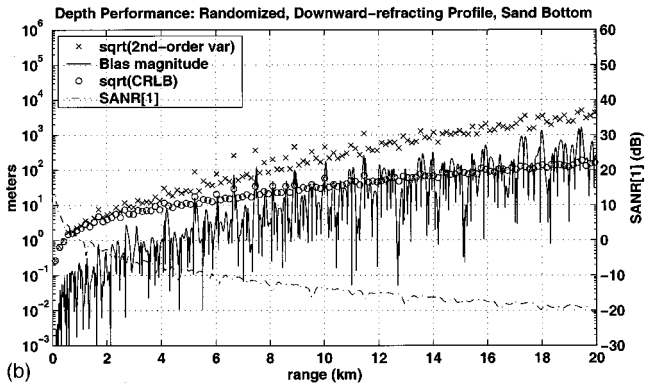


FIG. 9. Same as Fig. 8, recomputed for a single sample of a completely randomized signal.

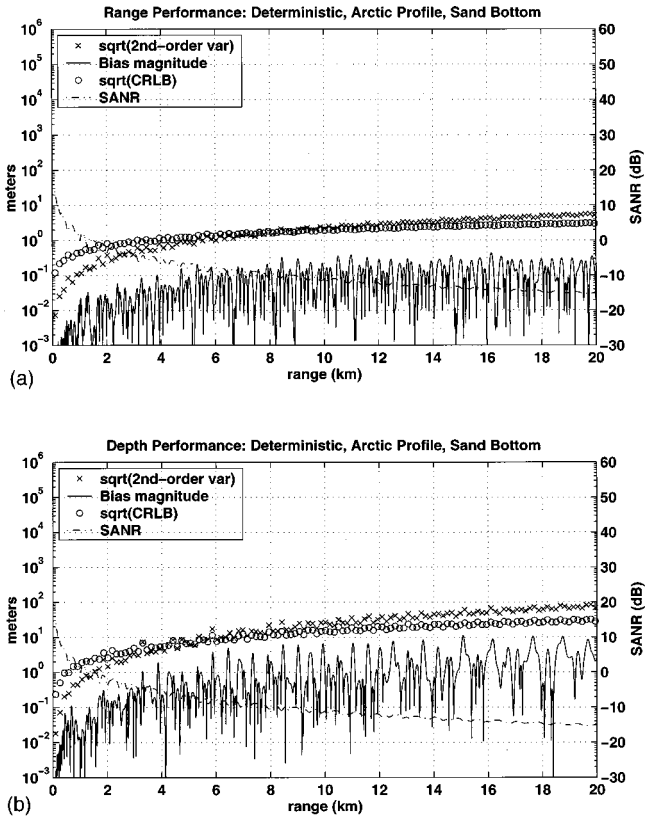


FIG. 10. Deterministic single-sample ocean acoustic localization MLE performance for (a) range estimation and (b) depth estimation versus source range, for a 100 Hz signal placed at 50 m depth in the case D environment, which consists of an upward-refracting profile over a sand bottom. See Fig. 2 caption for plot descriptions.

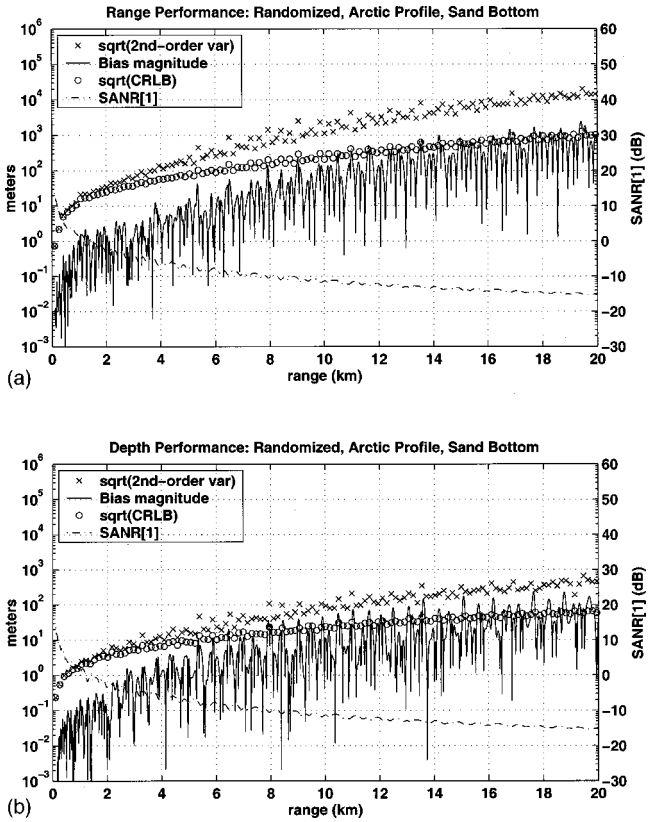


FIG. 11. Same as Fig. 10, recomputed for a single sample of a completely randomized signal.

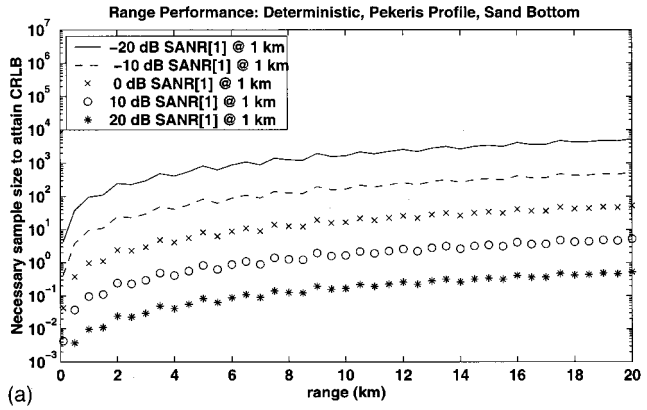
dashed–dotted line plotted relative to the right vertical axis. In all figures the source level has been adjusted so that SANR or SANR[1] at 1 km is 0 dB. Note that it is the square root of the CRLB and the second-order variance that has been plotted in the figures.

Figures 2 and 3 show the results of propagation through the case A environment, which consists of an isovelocity sound speed profile over a sandy bottom. The randomized signal MLE biases are much larger than those from an equivalent deterministic signal and the degradation in the range estimation performance is especially notable. Hundreds to thousands of data samples are required to reduce the randomized bias to less than 10 m at 20 km range. At ranges less than 6 km, when the input SANR is greater than -10 dB, the range biases are negligible for deterministic signals, less than 1 m, but roughly 10 times more significant, 10 m, for randomized signals with SANR[1] greater than -10 dB at ranges of 6 km or less. For deterministic signals, as the input SANR descends below -10 dB, the bias magnitude increases by an order of magnitude, so that at 20 km range, where the input SANR drops to -19 dB, the asymptotic range and depth biases reach maxima of 1 and 30 m, respectively. The corresponding maxima for the randomized signal biases are 8000 m and 800 m, for a 20 km range source. Even at a typical operational range of 6 km, where the array-gain augmented SANR[1] is roughly 0 dB, the randomized signal localization biases are greater than 10 m.

A similar pattern is evident for the second-order variance. As the input SANR descends below -10 dB, the second-order variance magnitude grows much more rapidly than the CRLB magnitude, for both signal types. Beyond 6 km range the second-order localization variances of the deterministic signal equal or exceed those of the CRLB. For randomized signals, the second-order localization variances are even greater, exceeding the CRLB by nearly two orders of magnitude at 6 km range. In both cases the CRLB underestimates the true parameter variance, and tens to thousands of data samples are required to make the second to first-order variance ratio (SFOVR) negligible, and so have the MLE asymptotically attain minimum variance.

Figures 4 and 5 show the results of using a 200 Hz signal to estimate source position in the case A environment. One might expect the greater number of available propagating modes to improve the MLE localization performance. Indeed, the localization bias magnitude for both signal types decreases slightly. Doubling the frequency also reduces the range and depth second-order variance terms by factors of 4, for both signal types. However, the deterministic CRLB is also reduced by a similar factor. Therefore, while the localization variance decreases with increasing frequency, the SFOVR remains unaffected, so there is no reduction in the number of deterministic data samples required to attain the CRLB. In contrast, the randomized signal SFOVR does decrease with increasing frequency.

Figures 6 and 7 show the effects of a different bottom composition, in this case, silt, on the localization performance, using the original 100 Hz source. The localization performance has worsened noticeably relative to that of the sand bottom, due to the absence of higher-order modes. For



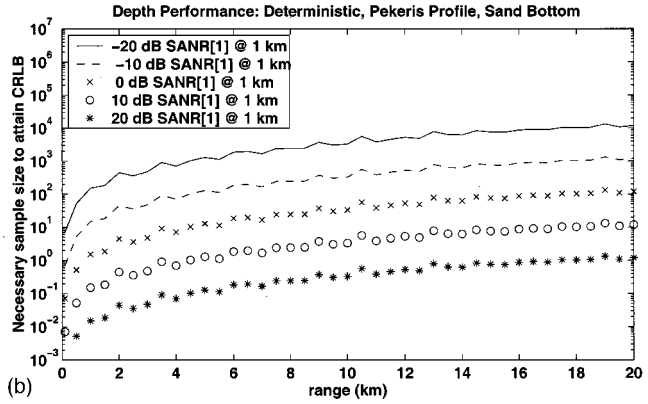
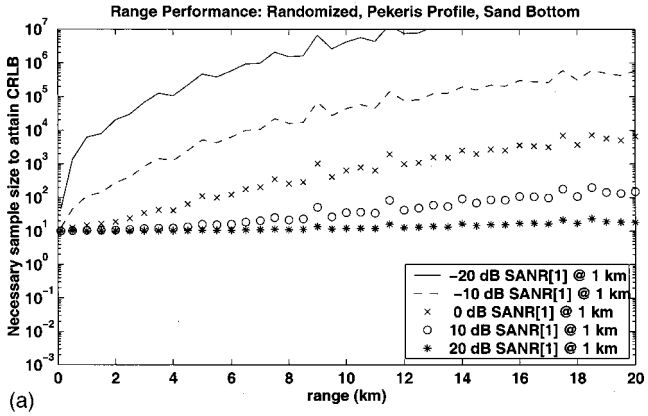


FIG. 12. Minimum independent data sample size (MSS) necessary for a deterministic ocean acoustic MLE to effectively attain the CRLB. The MSS is defined in Eq. (13) as the minimum sample size necessary for the second-order variance to be reduced to 10% of the CRLB. The MSS is plotted for (a) range estimation and (b) depth estimation versus source range and source level, for the case A environment. The 100 Hz source is located at 50 m depth. Values less than 1 indicate that a single data sample is sufficient to attain the CRLB.



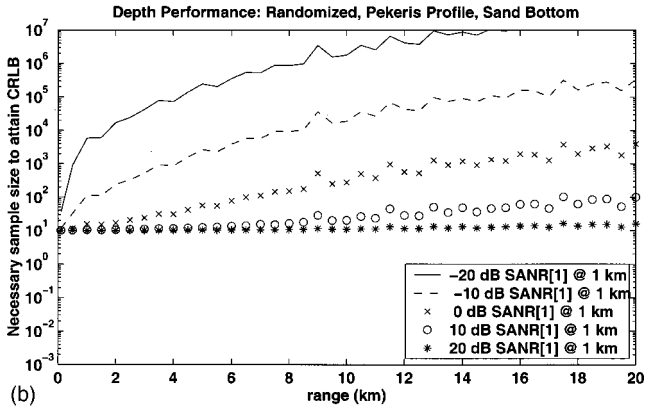


FIG. 13. Same as Fig. 12, recomputed for a completely randomized signal.

example, the deterministic depth bias exceeds the 100 m waveguide depth beyond 9 km range, whereas for a signal propagating over a sand bottom, the bias is less than 100 m out to 20 km range. The randomized depth bias exceeds the waveguide depth beyond 5 km range in the silt environment, compared with 12–14 km for the sand environment. Even at typical operational ranges of a few kilometers, randomized signals received with an array-gain augmented SANR[1] of 0 dB will have range and depth biases on the order of tens of meters.

Figures 8–11 illustrate the effects of including the more complex sound speed profiles of cases C and D. A comparison between these figures and Figs. 2 and 3 suggests that the variations in the sound speed profile illustrated in Fig. 1 have relatively minor effects on localization performance, as compared with changes in bottom composition.

B. Minimum sample size necessary to attain CRLB

Figures 12–15 show the minimum sample sizes necessary for the second-to-first-order variance ratio (SFOVR) to be less than 0.1. We take this as a necessary condition for an MLE estimate to approximately attain the CRLB in the asymptotic regime. For convenience a necessary minimum sample size (MSS) is defined here as 10 times the ratio of the second-to-first order variance for a single sample. In other words, if $\mathrm{var}_1(n)$ represents the first-order variance term in Eq. (3), derived from n samples, and $\mathrm{var}_2(n)$ represents the sum of all the second-order terms, also derived from n samples, then our necessary criterion becomes

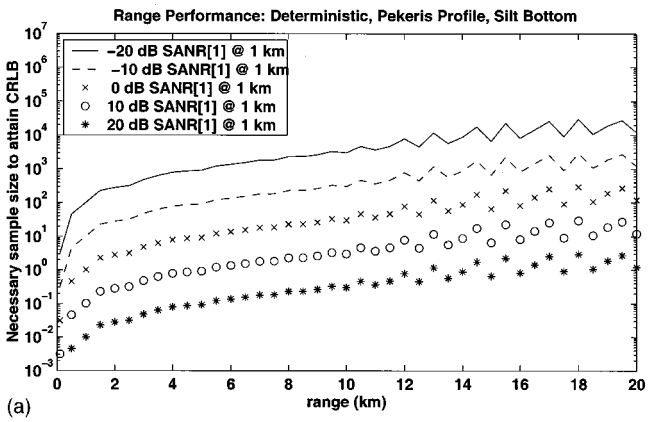
$$\frac{\mathrm{var}_2(n)}{\mathrm{var}_1(n)} = \frac{\mathrm{var}_2(1)}{n\,\mathrm{var}_1(1)} \leqslant 0.1, \tag{10a}$$

where

$$\mathrm{MSS} = 10\frac{\mathrm{var}_2(1)}{\mathrm{var}_1(1)}. \tag{10b}$$

A single sample is sufficient to achieve the CRLB, for MSS values less than or equal to unity. The required MSS has been plotted as a function of range for the case A environment in Figs. 12 and 13, for both deterministic and randomized signals. For a deterministic signal with an input SANR of 0 dB at 1 km, as modeled in Fig. 2, an MSS of 20 is necessary to attain the CRLB at a range of 10 km and a depth of 50 m. If the input SANR increases by a factor of 10 (10 dB), only two samples would be required to attain the deterministic CRLB, since MSS is inversely related to SNR[1] and SANR[1] for deterministic signals, as discussed in Appendix B.

The randomized signal results in Fig. 13 display some fundamental differences from their deterministic counter-



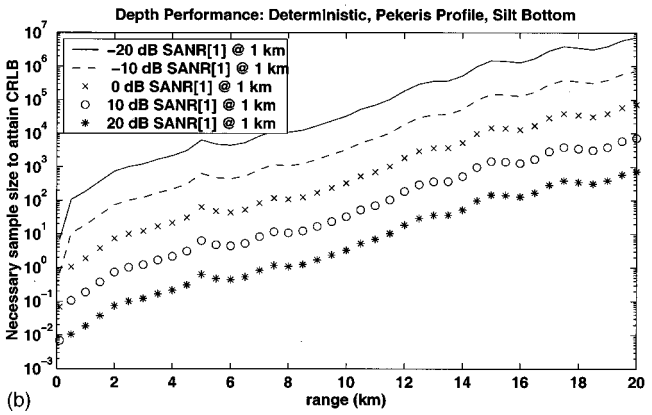
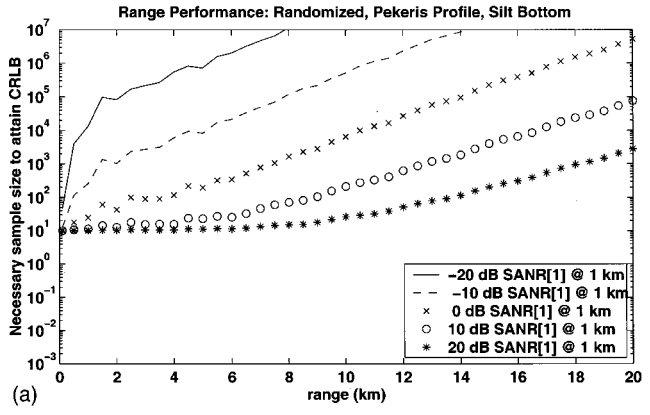


FIG. 14. Minimum independent data sample size necessary for a deterministic ocean acoustic MLE to effectively attain the CRLB for the case B environment. All other parameters remain unchanged from Fig. 12.



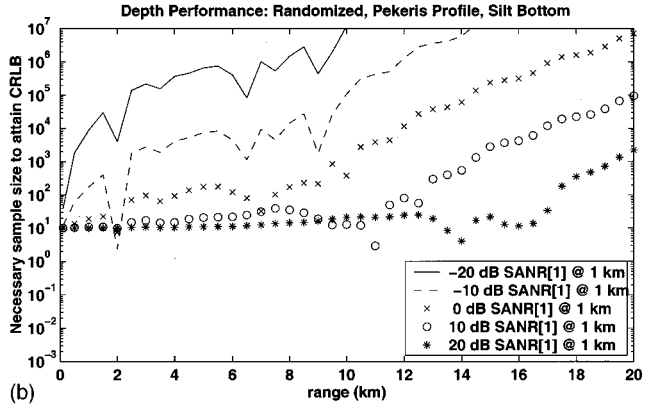


FIG. 15. Same as Fig. 14, recomputed for a completely randomized signal.

parts. First, the randomized MSS is always much larger than those required for deterministic signals. For example, roughly 1000 data samples are required to attain the CRLB at 10 km range and 50 m depth, for a signal with an input SANR[1] of 0 dB at 1 km range. The randomized MSS eventually plateaus to approximately 10 as the SANR[1] increases. Indeed, the MSS uniformly converges to a value of 10 at all ranges, as the SANR[1] increases past 20 dB at 1 km range. In other words, the second-order variance converges to the same magnitude as the first-order variance at large SANR[1] values. When the additive noise term is zero and $N/2>1$, however, the random signal model is not well defined as is discussed in Appendix B since the covariance given in Eq. (9) is no longer invertible.

Figures 14 and 15 show the MSS for deterministic and randomized signals, respectively, in the case B environment, a Pekeris waveguide with a silt bottom. While the MSS values are generally much greater than the values obtained for a signal propagating over a sand bottom, the overall trends visible in the MSS are the same as observed for the sand bottom. The deterministic signal MSS follows the expected inverse relationship with SNR[1], and the randomized signal MSS asymptotically approaches 10 as the input SANR[1] becomes large, converging SNR[1] to 0 dB.

C. Effect of source depth on localization performance

Another factor that influences localization performance is the source depth. Figures 16 and 17 illustrate the CRLB, bias magnitude, bias sign, and the MSS for a deterministic signal in the case A environment. As discussed in Sec. III, the noise matrix has been rescaled so that SANR from a 1 km range source will be 0 dB for every source depth. Figures 18 and 19 show the corresponding results for randomized signals in terms of SANR[1].

The general features of the contour plots are similar for both signal types. As SANR follows a decreasing trend with range, both the CRLB and MSS increase together in a highly correlated fashion. Locations with a large CRLB also tend to require a large MSS to attain the bound. For a given source range in this environment, the waveguide center and boundaries tend to produce the lowest CRLB and MSS. In Fig. 17 only the depth bias displays any significant depth dependence, where it increases sharply near the waveguide boundaries. An examination of the depth bias sign of both signal types reveals that the large biases at the boundaries tend to shift the estimates toward the waveguide center, so that localizations near the surface will be positively biased, and localizations near the ocean bottom will be negatively biased.

The depth bias shows other sign reversals with depth, for both deterministic and randomized signals. The 100 Hz signal in the case A environment generates 11 distinct sign reversals over the waveguide depth, out to ranges of 5 km. Beyond this range the bias sign structure dissipates as the input SANR falls below −10 dB. The consequence of these

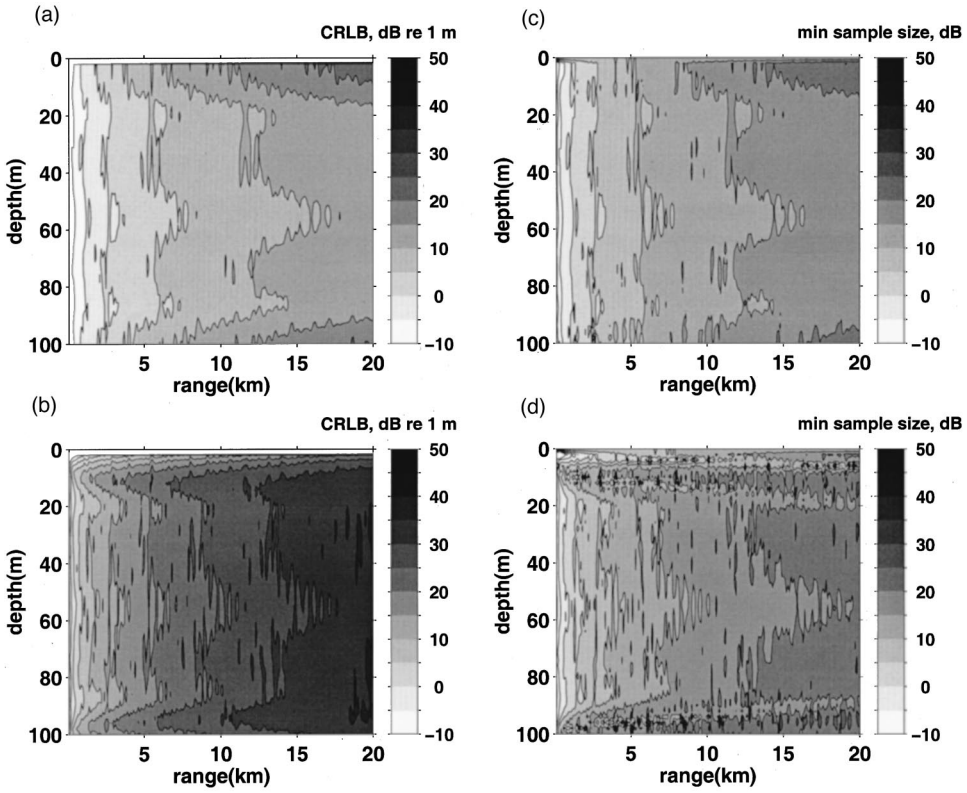


FIG. 16. Contour plots of (a) deterministic range CRLB, (b) deterministic depth CRLB, (c) minimum sample size required for a deterministic range estimate to attain the CRLB, and (d) minimum sample size required for a deterministic depth estimate to attain the CRLB. The 100 Hz acoustic source has been placed in the case A environment, and the received SANR at 1 km source range is 0 dB at all source depths, as described in Sec. III. All quantities are expressed in terms of dB units, and the contour interval is 5 dB *re* 1 m.

depth patterns is that MLE depth estimates will have a tendency to converge toward depths where the bias sign switches from positive to negative, with increasing depth. In terms of the contour plot, these regions lie wherever a black (positive bias) layer overlies a white (negative bias) layer. The range bias sign also shows alternating patterns indicative of range convergence, particularly at the waveguide middepth.

These spatial effects are more dramatic when an environment with fewer propagating modes is analyzed. Figures 20–23 illustrate the results of applying the same asymptotic bias and variance computations to the case B environment,

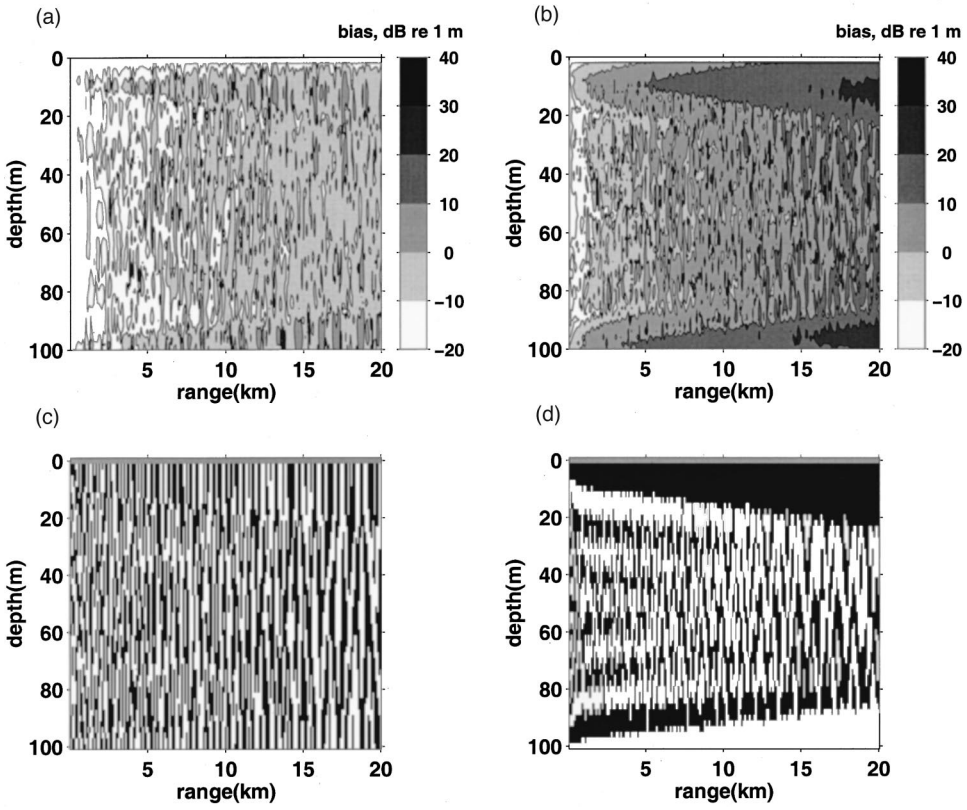


FIG. 17. Contour plots of (a) deterministic range bias magnitude, (b) deterministic depth bias magnitude, (c) deterministic range bias sign, and (d) deterministic depth bias sign, for the same scenario described in Fig. 16. The top row is in dB units, with a contour interval of 10 dB *re* 1 m. The bottom row displays positive values as black, negative values as white. Note the horizontal layers of alternating sign in the depth bias sign plot.

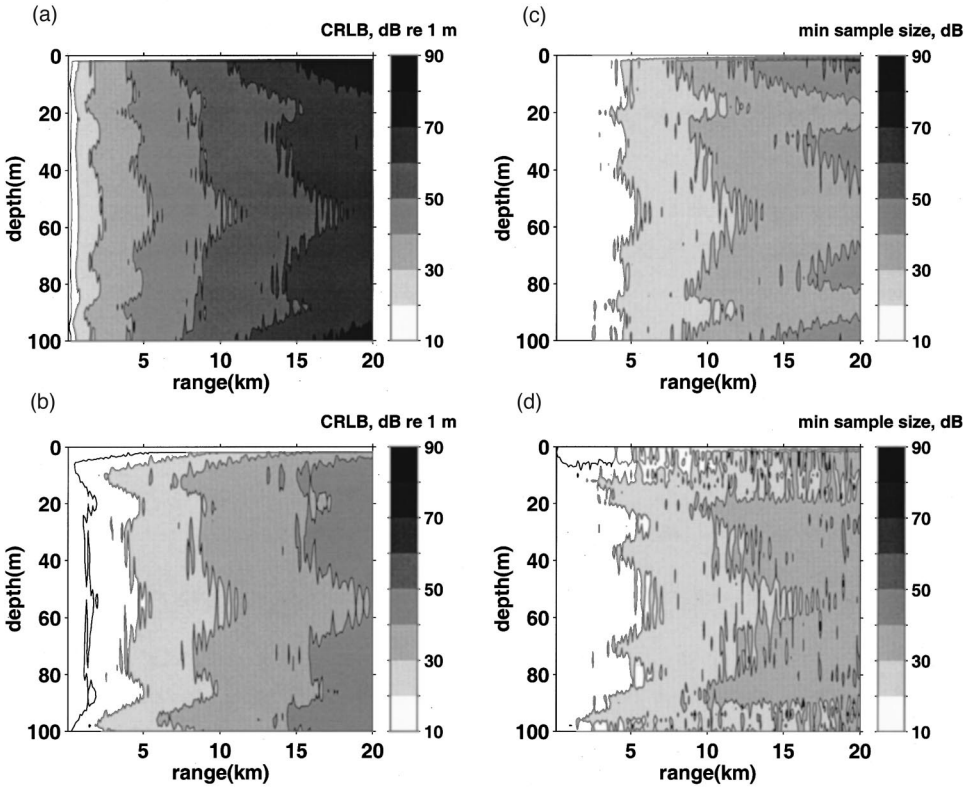


FIG. 18. Same as Fig. 16, recomputed for a completely randomized source signal. The contour interval is 10 dB.

characterized by a silt bottom. In both propagation environments depth localizations near the waveguide boundaries lead to the best performance, in terms of both the CRLB and MSS. For the silt case, the CRLB and MSS reach their maxima at the waveguide midpoint. The localization biases clearly display strong tendencies to converge at certain ranges and depths. For example, deterministic signal localizations will tend to converge toward the waveguide middepth and toward range cells evenly separated 2.5 km apart, which is the modal interference length between the two propagating modes in the system. The randomized signal localizations display similar features.

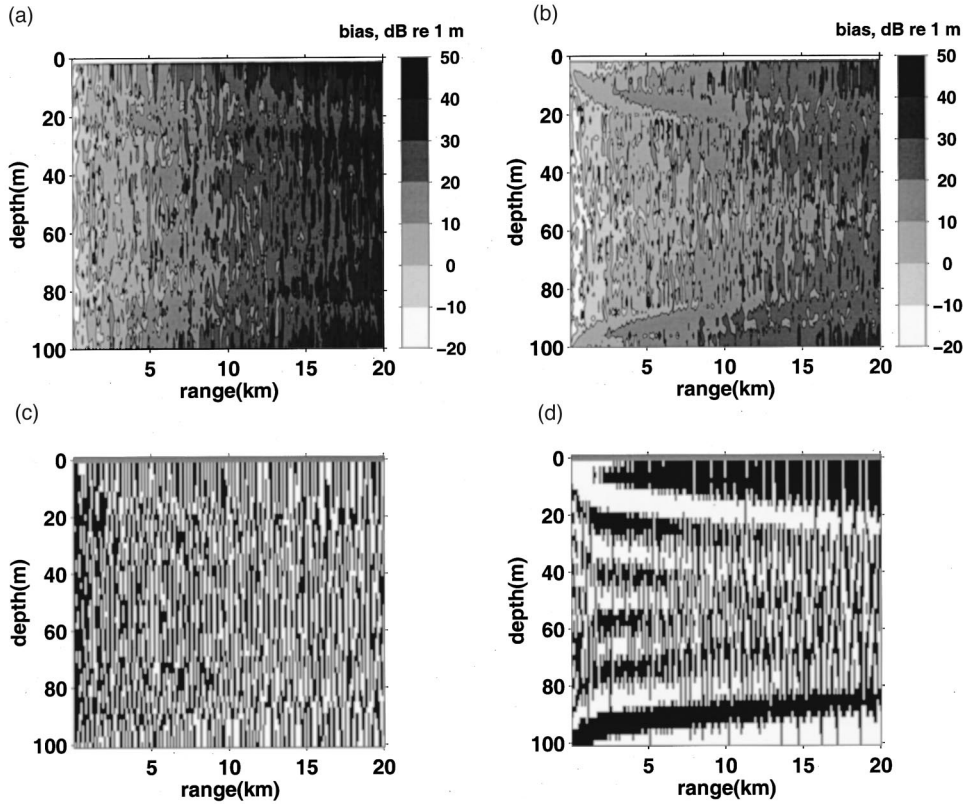


FIG. 19. Same as Fig. 17, recomputed for a completely randomized source signal. The contour interval for the top row is 10 dB *re* 1 m, and the bottom row displays positive values as black, negative values as white.

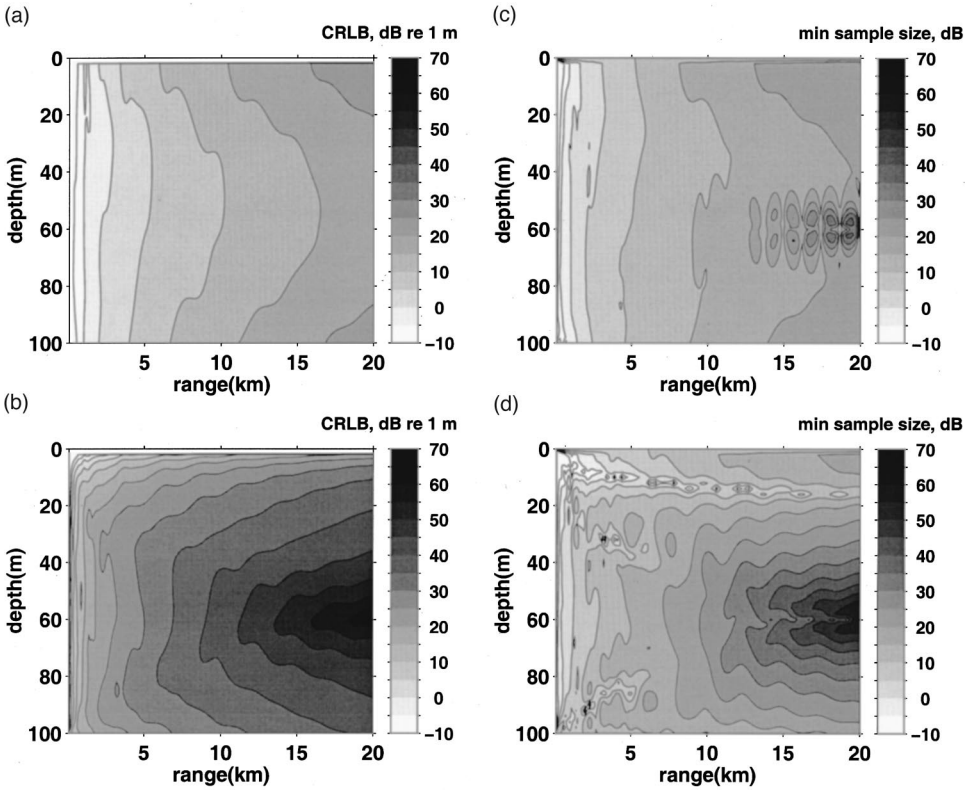


FIG. 20. Same as Fig. 16, but recomputed for a deterministic signal propagating through the case B environment. The contour interval is 5 dB.

V. DISCUSSION

It is convenient to frame the discussion in terms of SANR, starting in the high SANR regime. The results presented in Figs. 2–23 have been computed using a 10 element array. As indicated above, an increase in the number of array elements used to perform the inversion is expected to lead to a reduction in the biases and variances presented here, due to the effects of array gain. In the presence of spatially uncorrelated white noise, an $N/2$-element array increases the array-gain augmented SANR by $10\log(N/2)$ over the input SANR defined in Eq. (7). Since the bias and variance have been shown to be primarily functions of the signal SNR, these

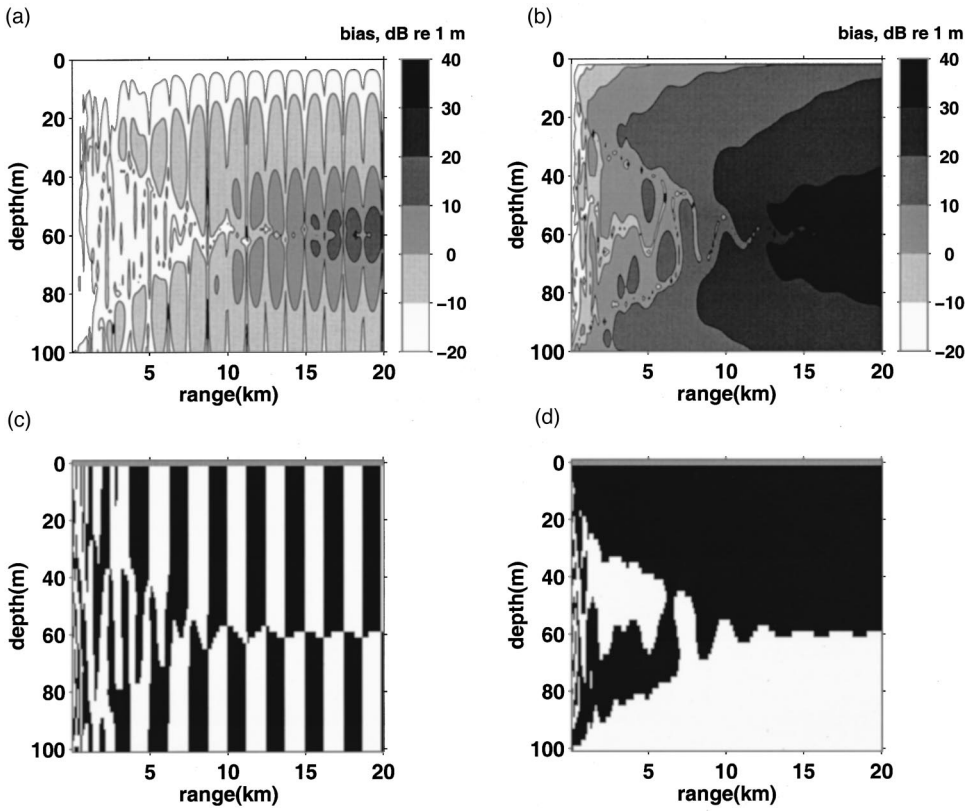


FIG. 21. Same as Fig. 17, but recomputed for a deterministic signal propagating through the case B environment. The contour interval for the top row is 10 dB *re* 1 m, and the bottom row displays positive values as black, negative values as white.

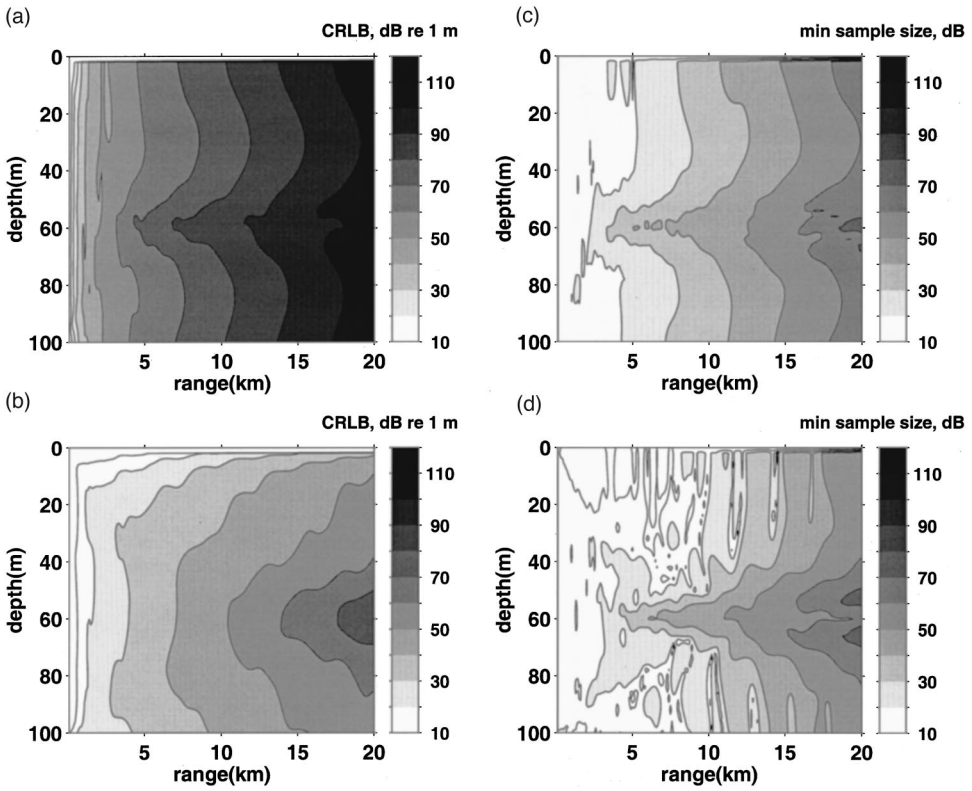


FIG. 22. Same as Fig. 16, but recomputed for a completely randomized signal propagating through the case B environment. The contour interval is 10 dB.

quantities decrease with increasing array gain. For example, a 10-fold increase in the number of hydrophone elements is expected to generate a 10-fold decrease in the deterministic bias and MSS, provided that the background noise covariance has similar characteristics to uncorrelated white noise.

Estimates extracted from deterministic signals can attain the CRLB with a single sample or snapshot of the field across the array, for sufficiently high SANR values, as Figs. 12 and 14 attest. In other words, at high SANR levels the second-order variance computed from Eq. (3) is negligible relative to the first-order variance, even when $n=1$. However, Figs. 13 and 15 illustrate how the situation for ran-

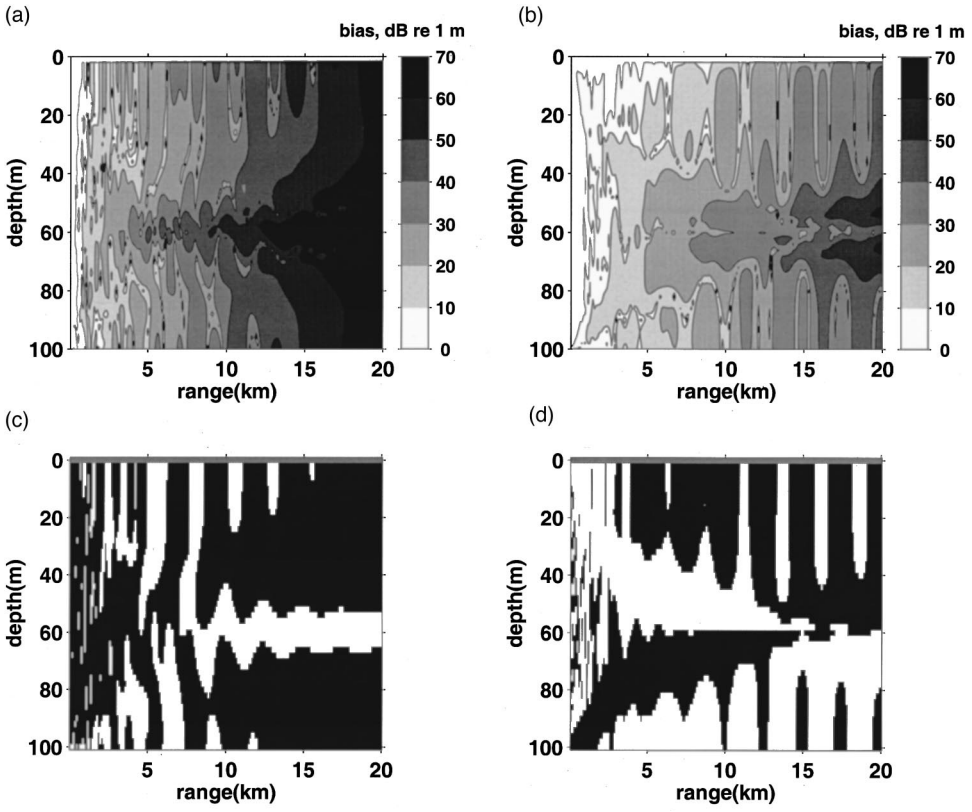


FIG. 23. Same as Fig. 17, but recomputed for a completely randomized signal propagating through the case B environment. The contour interval for the top row is 10 dB *re* 1 m, and the bottom row displays positive values as black, negative values as white.

domly fluctuating signals is more complex, in that estimates derived from these signals do not attain minimum variance, unless at least 10 data samples are used to construct the estimate. The reason behind this limitation is outlined in Sec. III, where the difference between the SANR, SANR[1] and the SNR of a measurement is discussed.

For example, Figs. 12 and 14 show that as the input SANR of a deterministic signal descends below -10 dB, the magnitude of the second-order variance begins to exceed the CRLB, so a deterministic estimate derived from a single data sample fails to attain the CRLB. As stated earlier, a -10 dB input SANR across a 10-element array yields an approximate array-gain augmented SANR of 0 dB.

For every 10 dB decrease in SANR, the deterministic SFOVR increases by an order of magnitude, resulting in a 10-fold increase in the MSS. This inverse relationship between the MSS and SANR[1] is a consequence of the fact that given spatially white noise in the deterministic signal model $\mathrm{SNR}=\mathrm{SANR}=n\mathrm{SANR}[1]$, as discussed in Sec. III. The MSS itself is relatively insensitive to changes in source frequency, sound–speed profile, and even bottom composition, for the cases investigated. For example, a 6 km range source in a sandy environment yields the same input SANR (-10 dB) as a 4 km range source in a silt environment. Examination of Figs. 12 and 14 show that the MSS is the same for both situations, suggesting that the deterministic signal results presented here can be used to guide analysis in other propagation environments and array geometries, if the transmission loss curves are known.

No simple relationship between SANR[1] and MSS exists for randomized signals, because the SANR is not proportional to SNR in this case. As Figs. 13 and 15 demonstrate, the SFOVR, and thus the MSS, are nonlinearly related to SANR[1] and are very sensitive to propagation effects. Environments dominated by only a few propagating modes, in particular, seem to create situations where the SFOVR and MSS can change by an order of magnitude with only a small change in source range. However, at high SANR[1] levels the MSS asymptotically approaches a value of 10 at all ranges and for all environments investigated, which seems to imply that for $n=1$ the CRLB cannot be attained for large SANR[1]. Our analysis indicates that this is only the case when $\frac{N}{2}=1$ because for more than one receiver the error tends to zero for large SANR[1] as is discussed in Appendix B. This asymptote is due to the fact that with our definition of SANR[1], in the random signal model terms in Eqs. (2) and (3) that are of order n^{-m} are not necessarily also of order $\mathrm{SANR}[1]^{-m}$. Both 1st and 2nd order terms may then appraoch zero with the same power law in SANR[1].

As the array-gain augmented SANR descends below 0 dB the localization biases are no longer negligible for either signal type. The spatial distribution of the bias sign reveals that the maximum-likelihood localization estimates tend to converge toward particular ranges and depths at low SANR. The exact convergence locations depend on the propagation environment; however, the localizations are generally biased away from the waveguide boundaries.

In some sense, the results provided here are still optimistic, as the ocean environment has been assumed perfectly characterized. In most practical situations, the waveguide parameters are insufficiently known, and this environmental uncertainty will further degrade the localization performance. The effects of this environmental uncertainty can be incorporated into Eqs. (2) and (3) by adding geoacoustic parameters to the parameter vector $\boldsymbol{\theta}$, and then computing derivatives of Eq. (5) with respect to these parameters. This differentiation may be accomplished either via numerical methods or by perturbation theory.[31]

VI. CONCLUSION

Asymptotic expressions for the first-order bias and second-order variance of a MLE have been applied to the problem of localizing an acoustic source in an ocean waveguide, for the cases of deterministic and randomized signals received with independent and additive background noise. The results suggest that as the array-gain augmented signal-to-additive noise ratio (SANR) at the array output descends below 0 dB, the MLE exhibits significant biases and variances that can exceed the CRLB by orders of magnitude. The localization biases tend to concentrate the estimates around particular source ranges and depths for moderate SANR values.

In principle, if enough data samples are available, unbiased estimates can be derived from low SANR signals. However, if the acoustic source is changing position with time, as is usually the case, the number of independent data samples available to construct a localization estimate is limited, because the estimation parameters themselves are changing with time. Therefore, under many practical operational scenarios, localization estimates are expected to be significantly biased, and the CRLB will underestimate the true variance by orders of magnitude.

ACKNOWLEDGMENT

This research was sponsored by ONR.

APPENDIX A: ANALYTICAL DERIVATIVES OF MODES USING PLANE-WAVE DECOMPOSITION

The moments presented in Appendixes B and C in paper I[12] require expressions for multiple-order depth derivatives of the waveguide normal modes. These are achieved by decomposing an individual mode into an upward and downward propagating plane wave at the desired source depth. Suppose that the values of mode Ψ_m at depths z and $z+H$ are known, where H is a small depth increment. Assuming that the sound–speed is constant between the two depths, the coefficients of the upgoing and downgoing plane waves connecting the two points are obtained by a matrix equation:

$$\begin{bmatrix} \Psi_m(z) \\ \Psi_m(z+H) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ e^{ik_{m,z}H} & e^{-ik_{m,z}H} \end{bmatrix} \begin{bmatrix} A^+ \\ A^- \end{bmatrix}, \tag{A1}$$

where $k_{m,z}=\sqrt{k^2-k_{r,m}^2}$ is the vertical modal wave number. The above matrix is easily inverted to solve for the coeffi-

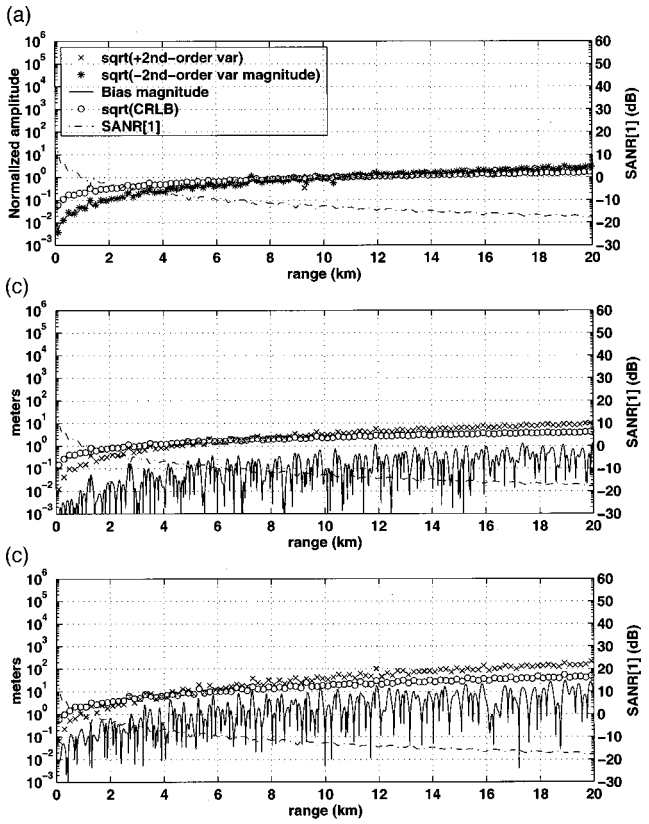


FIG. 24. Deterministic signal single-sample ($n=1$) ocean acoustic MLE performance for (a) amplitude A, (b) range, and (c) depth estimation versus source range, for a 100 Hz signal placed at 50 m depth in the case A environment, a Pekeris waveguide with a sand bottom. The true source amplitude A is 1. A comparison of these results with Fig. 2 shows that the addition of amplitude as an estimation parameter has negligible effect on the MLE localization performance.

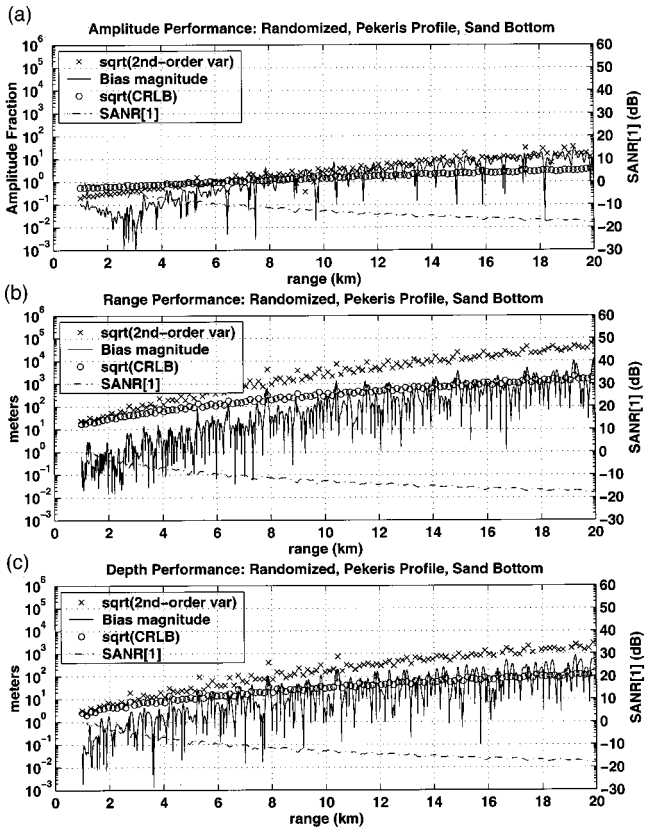


FIG. 25. Random signal single-sample ($n=1$) ocean acoustic MLE performance for (a) amplitude A, (b) range, and (c) depth estimation versus source range, for a 100 Hz signal placed at 50 m depth in the case A environment, a Pekeris waveguide with a sand bottom. The true source amplitude A is 1. A comparison of these results with Fig. 3 shows that the addition of amplitude as an estimation parameter has negligible effect on the MLE localization performance.

cient vector **A**. The modal derivatives with respect to depth can then be written in terms of **A**:

$$\mathbf{U}_z=\begin{bmatrix}\Psi'_m(z)\\ \Psi'_m(z+H)\end{bmatrix}$$
$$=\begin{bmatrix}ik_{m,z} & -ik_{m,z}\\ ik_{m,z}e^{ik_{m,z}H} & -ik_{m,z}e^{-ik_{m,z}H}\end{bmatrix}\begin{bmatrix}A^+\\ A^-\end{bmatrix}$$
$$\equiv\mathbf{D}_1\mathbf{A}. \tag{A2}$$

Second and third-order derivatives can be computed by defining matrices $\mathbf{D}_2$ and $\mathbf{D}_3$:

$$\mathbf{U_{zz}}=\mathbf{D}_2\mathbf{A},\quad D_{ij,2}=[D_{ij,1}]^2,$$
$$\mathbf{U_{zzz}}=\mathbf{D}_3\mathbf{A},\quad D_{ij,3}=[D_{ij,1}]^3. \tag{A3}$$

Use of Eqs. (A1)–(A3) allows computation of the modal derivatives using only the values of the modes at fixed points, without having to recourse to numerical differentiation.

APPENDIX B: THE EFFECT OF THE SANR AND SOURCE AMPLITUDE ESTIMATION ON LOCALIZATION PERFORMANCE

It is demonstrated here that the addition of the modulus of source amplitude $\alpha=|A(\omega)|$ for the deterministic signal model and source power $I_\alpha=\langle|A(\omega)|^2\rangle$ for the randomized signal model as a third estimation parameter in addition to source range and depth, has only a minor effect on localization performance. The bias and the variance terms are independent of α for the deterministic case and I_α for the randomized case, provided that the SANR[1] remains fixed. Variations between two-parameter and three-parameter-case estimates of range and depth are negligible in our localization scenarios, as we show in the simulations presented in the Figs. 24 and 25. No new derivatives are required to apply Eqs. (2) and (3) to the three-parameter estimation problem, since the source amplitude is linearly related to the data.

We first present analysis for the deterministic scenario and then we discuss the completely randomized case.

1. Deterministic signal model

In the deterministic signal model, according to Eqs. (5a), (5b), and (8), the expected complex vector field can be expressed as a real vector

$$\boldsymbol{\mu}=\alpha\begin{bmatrix}\cos(\phi)\mathbf{I} & -\sin(\phi)\mathbf{I}\\ \sin(\phi)\mathbf{I} & \cos(\phi)\mathbf{I}\end{bmatrix}\begin{bmatrix}\mathrm{Re}(\tilde{\mathbf{g}}(r,z))\\ \mathrm{Im}(\tilde{\mathbf{g}}(r,z))\end{bmatrix}$$
$$=\alpha\boldsymbol{\Omega}\mathbf{g}(r,z), \tag{B1}$$

where ϕ is the phase of $A(\omega)$, $\mathbf{\Omega}$ is a rotation matrix of dimension N, $\mathbf{I}$ is the identity matrix of dimension $N/2$ and $\mathbf{g}(r,z)$ is formed with the real part and the imaginary part of $\tilde{\mathbf{g}}(r,z)$. In our complex representation the source phase conveniently scales out of the problem since the matrix $\mathbf{\Omega}$ is orthogonal ($\mathbf{\Omega}^T\mathbf{\Omega}=1$). It follows that $\boldsymbol{\mu}_r=\alpha\mathbf{\Omega}\mathbf{g}_r(r,z)$, $\boldsymbol{\mu}_z=\alpha\mathbf{\Omega}\mathbf{g}_z(r,z)$, $\boldsymbol{\mu}_a=\mathbf{\Omega}\mathbf{g}(r,z)$, and $\boldsymbol{\mu}_{\alpha\alpha}=0$, where in this Appendix only the subscripts r, z, and α represent derivatives with respect to range, depth, and source amplitude, respectively. The noise covariance, according to Eqs. (6) and (8) is a diagonal matrix whose trace is given by $\mathrm{tr}(\mathbf{C}^{\mathrm{additive}})=\sigma^2(N/2)$. From Eqs. (2) and (3), Eqs. (B6)–(B15), and Eqs. (B18)–(B21), we then observe that the first-order bias and the CRLB are proportional to 1/SANR, where $\mathrm{SANR}=\mathrm{SANR}[n]=n(2\mu^T\mu/N\sigma^2)=n\ \mathrm{SANR}[1]$, while the second-order variance is proportional to $1/\mathrm{SANR}^2$. This property explains the dependence of the curves in Figs. 12 and 14 on SANR.

We first consider the problem of estimating two parameters, amplitude and range, to illustrate the issues. We then extend the results to the three-parameter case of amplitude, range, and depth.

From the Appendix of Ref. 12 and the definition above for the mean $\boldsymbol{\mu}$, the Fisher information becomes

$$i=\mathrm{SANR}\frac{N}{2}\begin{bmatrix}\dfrac{1}{\alpha^2} & \dfrac{\mathbf{g}^T\mathbf{g}_r}{\alpha\mathbf{g}^T\mathbf{g}}\\ \dfrac{\mathbf{g}^T\mathbf{g}_r}{\alpha\mathbf{g}^T\mathbf{g}} & \dfrac{\mathbf{g}_r^T\mathbf{g}_r}{\mathbf{g}^T\mathbf{g}}\end{bmatrix} \tag{B2}$$

from which the Cramer–Rao bound becomes

$$\mathrm{CRLB}=i^{-1}=\frac{1}{\mathrm{SANR}}\frac{1}{\left(\mathbf{g}_r^T\mathbf{g}_r-\dfrac{(\mathbf{g}^T\mathbf{g}_r)^2}{\mathbf{g}^T\mathbf{g}}\right)\dfrac{N}{2}}\times\begin{bmatrix}\overbrace{\alpha^2\mathbf{g}_r^T\mathbf{g}_r}^{\text{amplitude}} & -\alpha\mathbf{g}^T\mathbf{g}_r\\ -\alpha\mathbf{g}^T\mathbf{g}_r & \underbrace{\mathbf{g}^T\mathbf{g}}_{\text{range}}\end{bmatrix}. \tag{B3}$$

While the CRLB for the source amplitude scales with α^2, the CRLB for range is independent of α, for fixed SANR. For range estimation, the square root of the CRLB, which represents a first approximation to the error, becomes

$$\sqrt{\mathrm{CRLB}(r)}=\frac{1}{\sqrt{\dfrac{N}{2}\mathrm{SANR}\left|\dfrac{\mathbf{g}_{r,\perp}}{\mathbf{g}}\right|}}, \tag{B4}$$

where

$$|\mathbf{g}_{r,\perp}|^2=\mathbf{g}^T\mathbf{g}\mathbf{g}_r^T\mathbf{g}_r-(\mathbf{g}_r^T\mathbf{g})^2. \tag{B5}$$

As expected, the error decreases for increasing SANR. However, in Eq. (B4) there is also a geometrical interpretation: the error is not controlled directly by $\mathbf{g}_r^T\mathbf{g}_r$, but by the component of $\mathbf{g}_r$ orthogonal to $\mathbf{g}$.

A similar analysis is performed for the components of the bias when the tensors

$$v_{\alpha\alpha\alpha}=0, \tag{B6}$$

$$v_{\alpha\alpha r}=-2n\frac{\mathbf{g}^T\mathbf{g}_r}{\sigma^2}, \tag{B7}$$

$$v_{\alpha rr}=-n\alpha\left(\frac{\mathbf{g}_{rr}^T\mathbf{g}}{\sigma^2}+2\frac{\mathbf{g}_r^T\mathbf{g}_r}{\sigma^2}\right), \tag{B8}$$

$$v_{rrr}=-3n\alpha^2\frac{\mathbf{g}_{rr}^T\mathbf{g}_r}{\sigma^2}, \tag{B9}$$

$$v_{\alpha,rr}=n\alpha\frac{\mathbf{g}_{rr}^T\mathbf{g}}{\sigma^2}, \tag{B10}$$

$$v_{\alpha,r\alpha}=n\frac{\mathbf{g}^T\mathbf{g}_r}{\sigma^2}, \tag{B11}$$

$$v_{\alpha,\alpha\alpha}=0, \tag{B12}$$

$$v_{r,\alpha\alpha}=0, \tag{B13}$$

$$v_{r,r\alpha}=n\alpha\frac{\mathbf{g}_r^T\mathbf{g}_r}{\sigma^2}, \tag{B14}$$

$$v_{r,rr}=n\frac{\mathbf{g}_{rr}^T\mathbf{g}_r}{\sigma^2}, \tag{B15}$$

are substituted in (B2). The first-order bias then becomes

$$\begin{bmatrix}b^\alpha(1)\\ b^r(1)\end{bmatrix}=\frac{1}{\mathrm{SANR}\xi\dfrac{N}{2}}\left[\begin{bmatrix}\alpha\mathbf{g}_r^T\mathbf{g}_r\\ \mathbf{g}^T\mathbf{g}_r\end{bmatrix}\left(\mathbf{g}_{rr}^T\mathbf{g}-2\frac{(\mathbf{g}_r^T\mathbf{g})^2}{\mathbf{g}^T\mathbf{g}}\right)+\begin{bmatrix}\alpha\mathbf{g}^T\mathbf{g}_r\\ \mathbf{g}^T\mathbf{g}\end{bmatrix}\left(\mathbf{g}_{rr}^T\mathbf{g}_r-2\frac{\mathbf{g}_r^T\mathbf{g}\mathbf{g}_r^T\mathbf{g}_r}{\mathbf{g}^T\mathbf{g}}\right)\right], \tag{B16}$$

where $\xi=\mathbf{g}_r^T\mathbf{g}_r-[(\mathbf{g}^T\mathbf{g}_r)^2/\mathbf{g}^T\mathbf{g}]$. For fixed SANR, the range bias $b^r(1)$ is independent of α while the amplitude bias is proportional to α.

The analysis of the second-order covariance is simpler if we express the terms of order $O_p(n^{-2})$ in Eq. (3) as

$$\begin{aligned}\mathrm{cov}_2(i,j) = &-i^{ij} - i^{im}i^{jn}i^{pq}(v_{nq,m,p} + v_{mq,n,p} + v_{nmpq} + 3v_{nq,pm} + 2v_{nmp,q} + \tfrac{1}{2}v_{mpq,n} + \tfrac{1}{2}v_{npq,m}) \\ &+ i^{im}i^{jn}i^{pz}i^{qt}(\tfrac{1}{2}v_{npt}v_{m,q,z} + \tfrac{1}{2}v_{mpt}v_{n,q,z} + v_{npm}v_{qzt} + \tfrac{5}{2}v_{npq}v_{mzt} \\ &+ v_{n,qz}v_{mtp} + v_{m,qz}v_{ntp} + 2v_{nmt}v_{qp,z} + 3v_{nt,z}v_{mpq} + 3v_{npq}v_{mt,z} + \tfrac{1}{2}v_{n,mt}v_{pqz} \\ &+ \tfrac{1}{2}v_{nt,m}v_{pqz} + v_{nq,z}v_{pt,m} + v_{mq,z}v_{pt,n} + v_{nq,m}v_{pt,z} + v_{mq,n}v_{pt,z} + v_{nq,p}v_{mt,z}),\end{aligned} \tag{B17}$$

where $i,j = r$, z or α and we use the explicit expression of the tensors in (B3)–(B12) as well as the tensors

$$\begin{aligned}v_{mnpq} = \frac{n}{\sigma^2}(&\boldsymbol{\mu}^T_{mnp}\boldsymbol{\mu}_q + \boldsymbol{\mu}^T_{mnq}\boldsymbol{\mu}_p + \boldsymbol{\mu}^T_{npq}\boldsymbol{\mu}_m \\ &+ \boldsymbol{\mu}^T_{mpq}\boldsymbol{\mu}_n + \boldsymbol{\mu}^T_{mn}\boldsymbol{\mu}_{pq} + \boldsymbol{\mu}^T_{mp}\boldsymbol{\mu}_{nq} + \boldsymbol{\mu}^T_{mq}\boldsymbol{\mu}_{np}),\end{aligned} \tag{B18}$$

$$v_{mnp,q} = \frac{n}{\sigma^2}\boldsymbol{\mu}^T_{mnp}\boldsymbol{\mu}_q, \tag{B19}$$

$$v_{mn,pq} = i_{mn}i_{pq} - \frac{n}{\sigma^2}\boldsymbol{\mu}^T_{mn}\boldsymbol{\mu}_{pq}, \tag{B20}$$

$$v_{mn,p,q} = -i_{mn}i_{pq}, \tag{B21}$$

where m, n, p, q are equal to α or r. We find that $\mathrm{cov}_2(\alpha,\alpha) \propto \alpha^2/\mathrm{SANR}^2$ and $\mathrm{cov}_2(r,r) \propto 1/\mathrm{SANR}^2$, where $\propto$ means *linearly proportional to*. The second-order covariance for the range MLE is then independent of α, while the second-order amplitude variance scales with α^2.

If we include depth as a third parameter, the bias and variances of both range and depth will still only depend on SANR and not on α alone. For example, if we consider the Fisher information matrix for the three-parameter case, we have

$$i = \mathrm{SANR}\frac{N}{2}\begin{bmatrix} \frac{1}{\alpha^2} & \frac{\mathbf{g}^T\mathbf{g}_r}{\alpha\mathbf{g}^T\mathbf{g}} & \frac{\mathbf{g}^T\mathbf{g}_z}{\alpha\mathbf{g}^T\mathbf{g}} \\ \frac{\mathbf{g}^T\mathbf{g}_r}{\alpha\mathbf{g}^T\mathbf{g}} & \frac{\mathbf{g}_r^T\mathbf{g}_r}{\mathbf{g}^T\mathbf{g}} & \frac{\mathbf{g}_r^T\mathbf{g}_z}{\mathbf{g}^T\mathbf{g}} \\ \frac{\mathbf{g}^T\mathbf{g}_z}{\alpha\mathbf{g}^T\mathbf{g}} & \frac{\mathbf{g}_r^T\mathbf{g}_z}{\mathbf{g}^T\mathbf{g}} & \frac{\mathbf{g}_z^T\mathbf{g}_z}{\mathbf{g}^T\mathbf{g}} \end{bmatrix} \tag{B22}$$

from which it can be inferred that $\mathrm{CRLB}(\alpha,\alpha) \propto \alpha^2/\mathrm{SANR}$, $\mathrm{CRLB}(r,r) \propto 1/\mathrm{SANR}$, and $\mathrm{CRLB}(z,z) \propto 1/\mathrm{SANR}$.

The functional dependencies of these biases and variances on SANR, clearly differs from the two-parameter case. A comparison of Fig. 2 and Fig. 24 for deterministic signals and Fig. 3 with Fig. 25 for randomized signals, however, shows that this difference is negligible for our estimation scenarios.

2. Random signal model

For a randomized signal, the complex covariance matrix is given by

$$\tilde{\mathbf{C}} = \sigma^2\mathbf{K}, \tag{B23}$$

where

$$\mathbf{K} = \frac{I_\alpha}{\sigma^2}\tilde{\mathbf{g}}\tilde{\mathbf{g}}^+ + \mathbf{I} \tag{B24}$$

and $\mathrm{tr}(\mathbf{I}) = N/2$. We define $\mathrm{SANR}[1] = 2I_\alpha\mathbf{g}^T\mathbf{g}/N\sigma^2 = I_\alpha\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}/N\sigma^2$. Let us consider the properties of $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{C}}^{-1}$ as the additive noise vanishes. Note that $\tilde{\mathbf{g}}\tilde{\mathbf{g}}^+$ is a matrix that projects $N/2$-dimensional real vectors onto the subspace parallel to $\tilde{\mathbf{g}}$, annihilating the components in all other directions. If we introduce a unitary matrix $\mathbf{Q}$ that rotates $\tilde{\mathbf{g}}$ in the first component of an $N/2$-dimensional complex vector space, we can then write $\tilde{\mathbf{C}}$ as

$$\tilde{\mathbf{C}} = \mathbf{Q}\mathbf{D}\mathbf{Q}^+, \tag{B25}$$

where

$$\begin{aligned}\mathbf{D} &= I_\alpha|\tilde{\mathbf{g}}|^2\begin{bmatrix} 1 & & & 0 \\ & 0 & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & 0 \end{bmatrix} + \sigma^2\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & 1 \end{bmatrix} \\ &= \begin{bmatrix} I_\alpha|\tilde{\mathbf{g}}|^2 + \sigma^2 & & & 0 \\ & \sigma^2 & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \sigma^2 \end{bmatrix},\end{aligned} \tag{B26}$$

where $|\tilde{\mathbf{g}}|^2 = \tilde{\mathbf{g}}^+\tilde{\mathbf{g}}$. From this expression the inverse can be written as $\mathbf{C}^{-1} = \mathbf{Q}\mathbf{D}^{-1}\mathbf{Q}^+$, where

$$\mathbf{D}^{-1} = \begin{bmatrix} \frac{1}{I_\alpha|\mathbf{g}|^2 + \sigma^2} & & & 0 \\ & \frac{1}{\sigma^2} & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \frac{1}{\sigma^2} \end{bmatrix}. \tag{B27}$$

Since the determinant of $\tilde{\mathbf{C}}$ is equal to the determinant of $\mathbf{D}$ which tends to zero in proportion to $\sigma^{2[(N/2)-1]}$ as $\sigma \to 0$, the rank of $\tilde{\mathbf{C}}$ approaches unity as $\sigma \to 0$. From (B26) and (B27)

we observe that for $\sigma=0$, in the scalar case, where one receiver only is employed, the model remains well defined because the data vector is one dimensional, but $\tilde{\mathbf{C}}^{-1}$ does not exist for $N\geqslant 2$ because the determinant of $\tilde{\mathbf{C}}$ vanishes.

With the present definition of SANR in the random signal model, terms in Eqs. (2) and (3) of order n^{-m} are generally not of order in SANR^{-m}, in contrast to the situation found for a deterministic signal.

For randomized signals, however, the first order bias, the CRLB, and the second-order variance for range and depth still depend only on the SANR[1] and not on signal amplitude alone. This can be seen, from the definition of the Fisher information matrix for the randomized case $i_{nm}=\mathrm{tr}(\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{C}}_m\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{C}}_n)$, where $m=\alpha$, r and $n=\alpha$, r, $\tilde{\mathbf{C}}_\alpha=\partial\tilde{\mathbf{C}}/\partial I_\alpha=\tilde{\mathbf{g}}\tilde{\mathbf{g}}^+$ and $\tilde{\mathbf{C}}_r=\partial\tilde{\mathbf{C}}/\partial r=I_\alpha(\tilde{\mathbf{g}}_r\tilde{\mathbf{g}}^++\mathbf{g}\mathbf{g}_r^+)$, and

$$\mathbf{C}^{-1}=\frac{1}{\sigma^2}\left[I-\frac{\frac{I_\alpha}{\sigma^2}\tilde{\mathbf{g}}\tilde{\mathbf{g}}^+}{1+\frac{I_\alpha}{\sigma^2}|\mathbf{g}|^2}\right]. \tag{B28}$$

The Fisher information matrix becomes

$$i=n\left(\frac{\frac{N}{2}\mathrm{SANR}[1]}{1+\frac{N}{2}\mathrm{SANR}[1]}\right)^2\begin{bmatrix}\frac{1}{I_\alpha^2} & \frac{1}{I_\alpha}\frac{(\tilde{\mathbf{g}}_r^+\tilde{\mathbf{g}}+\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}_r)}{|\tilde{\mathbf{g}}|^2}\\ \frac{1}{I_\alpha}\frac{(\tilde{\mathbf{g}}_r^+\tilde{\mathbf{g}}+\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}_r)}{|\tilde{\mathbf{g}}|^2} & \frac{1}{|\tilde{\mathbf{g}}|^4}\left((\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}+\tilde{\mathbf{g}}^+\tilde{\mathbf{g}})^2+2\left(1+\frac{N}{2}\mathrm{SANR}[1]\right)(|\tilde{\mathbf{g}}_r|^2|\tilde{\mathbf{g}}|^2-\tilde{\mathbf{g}}_r^+\tilde{\mathbf{g}}\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}_r)\right)\end{bmatrix} \tag{B29}$$

with CRLB,

$$\mathrm{CRLB}=\frac{\Delta}{n\left(\frac{N}{2}\mathrm{SANR}[1]\right)^2}\begin{bmatrix}I_\alpha^2\frac{1}{|\tilde{\mathbf{g}}|^4}\left((\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}+\tilde{\mathbf{g}}^+\tilde{\mathbf{g}})^2+2\left(1+\frac{N}{2}\mathrm{SANR}[1]\right)(|\tilde{\mathbf{g}}_r|^2|\tilde{\mathbf{g}}|^2-\tilde{\mathbf{g}}_r^+\tilde{\mathbf{g}}\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}_r)\right) & -2I_\alpha\frac{\tilde{\mathbf{g}}_r^+\tilde{\mathbf{g}}+\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}_r}{|\tilde{\mathbf{g}}|^2}\\ -2I_\alpha\frac{\tilde{\mathbf{g}}_r^+\tilde{\mathbf{g}}+\tilde{\mathbf{g}}^+\tilde{\mathbf{g}}_r}{|\tilde{\mathbf{g}}|^2} & 1\end{bmatrix}, \tag{B30}$$

where

$$\Delta=\frac{1}{2}\left(1+\frac{N}{2}\mathrm{SANR}[1]\right)\left|\frac{\tilde{\mathbf{g}}_{r,\perp}}{\tilde{\mathbf{g}}}\right|^{-2}.$$

This explicitly shows that the range CRLB is independent of I_α, while the intensity CRLB scales as I_α^2 for fixed SANR[1].

As in the deterministic signal model, the moments of the asymptotic expansion can be expressed as a product of a function depending on SANR[1] and a function depending on the geometric properties of the received signal vector. For instance from Eq. (B30) we find that

$$\mathrm{CRLB}(r)=\frac{1}{2n}\frac{1+\frac{N}{2}\mathrm{SANR}[1]}{\left(\frac{N}{2}\mathrm{SANR}[1]\left|\frac{\tilde{\mathbf{g}}_{r,\perp}}{\tilde{\mathbf{g}}}\right|\right)^2}. \tag{B31}$$

Equation (B31) is only valid for $N/2>1$, and SANR[1] not equal to infinity, so that it does not apply to scalar parameter estimates from scalar data.

After similar analysis, too detailed to present for the full three-parameter randomized signal case, we find that the first order bias, the CRLB, and the second-order variance for range and depth depend only on the SANR[1] and not on I_α alone as expected.

[1] O. Diachok, A. Caiti, P. Gerstoft, and H. Schmidt, *Full Field Inversion Methods in Ocean and Seismic Acoustics* (Kluwer, Dordrecht, 1995).

[2] 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).

[3] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).

[4] C. R. Rao, *Linear Statistical Inference and its Applications* (Wiley, New York, 1973).

[5] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: source localization in correlated noise as an optimum parameter estimation problem," J. Acoust. Soc. Am. **83**, 571–587 (1988).

[6] M. J. Hinich and E. J. Sullivan, "Maximum-likelihood passive localization using mode filtering," J. Acoust. Soc. Am. **85**, 214–219 (1989).

[7] A. J. Weiss and E. Weinstein, "Fundamental limitations in passive time delay estimation," IEEE Trans. Acoust., Speech, Signal Process. **ASSP-31**, 472–485 (1983).

[8] A. J. Weiss and E. Weinstein, Lower bounds in parameter estimation-summary of results, presented at ICASSP, Tokyo, 1986.

[9] D. Chazan, M. Zakai, and J. Ziv, "Improved lower bounds on signal parameter estimation," IEEE Trans. Inf. Theory **IT-21**, 90–93 (1975).

[10] E. W. Barankin, "Locally best unbiased estimates," Ann. Math. Stat. **20**, 477–501 (1949).

[11] J. Tabrikian, "Barankin Bounds for Source Localization in an Uncertain Ocean Environment," IEEE Trans. Signal Process. **47**, 2917–2927 (1999).

[12] E. Naftali and N. C. Makris, "Necessary conditions for a maximum-likelihood estimate to become asymptotically unbiased and attain the Cramer-Rao lower bound. I. General approach with an application to time-delay and doppler shift estimation," J. Acoust. Soc. Am. **110**, 1917-1930 (2001). [Errata: in Sec. IV the order of the polynomials in all equations should have SNR replace sample size *n*, since $n=1$ in Eq. (12). Also in Eq. (23) the factor $8\omega_c^2$ should be replaced with $6\omega_c^2$ and the factor $3\omega_c^4$ should be replaced with ω_c^4.]

[13] R. L. Parker, *Geophysical Inverse Theory* (Princeton University Press, Princeton, New Jersey, 1994).

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1993), Vol. 1.

[15] E. J. Sullivan and D. Middleton, ''Estimation and detection issues in matched-field processing,'' IEEE J. Ocean. Eng. **18**, 156–167 (1993).
[16] 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).
[17] O. E. Barndorff-Nielsen and D. R. Cox, *Inference and Asymptotics* (Chapman and Hall, London, 1994).
[18] L. L. Van Trees, *Detection, Estimation and Modulation Theory* (Wiley, New York, 1970), Vol. III.
[19] N. C. Makris, ''Parameter resolution bounds that depend on sample size,'' J. Acoust. Soc. Am. **99**, 2851–2861 (1996). The misstatement in this paper that the distribution for the sample covariance of a CCGR data vector is unknown is corrected in Ref. 20 below. The distribution in question is the complex Wishart distribution.
[20] N. C. Makris, ''The statistics of ocean-acoustic ambient noise,'' in *Sea Surface Sound*, edited by T. Leighton (Kluwer Academic, Dordrecht, 1997).
[21] J. H. Goodman, *Statistical Optics* (Wiley-Interscience, New York, 1985).
[22] N. C. Makris, ''The effect of saturated transmission scintillation on ocean acoustic intensity measurements,'' J. Acoust. Soc. Am. **100**, 769–783 (1996).
[23] N. Levanon, *Radar Principles* (Wiley, New York, 1988).
[24] P. Swerling, ''Probability of detection for fluctuating targets,'' Trans. IRE Prof. Group Inf. Theory **IT-6**, 269–308 (1960).
[25] P. G. Bergmann, ''Intensity fluctuations,'' in *The Physics of Sound in the Sea, Part I: Transmission* (National Defense Research Committee, Washington, DC, 1946).
[26] J. L. Krolik, ''Matched-field minimum variance beamforming in a random ocean channel,'' J. Acoust. Soc. Am. **92**, 1408–1419 (1992).
[27] J. Krolik and G. Niezgoda, ''Wideband matched-field processing in a random ocean channel,'' presented at Acoustic Signal Processing for Ocean Exploration, 1993.
[28] N. O. Booth, P. A. Baxley, J. A. Rice, P. W. Schey, W. S. Hodgkiss, G. L. D'Spain, and J. J. Murray, ''Source localization with broad-band matched-field processing in shallow water,'' IEEE J. Ocean. Eng. **21**, 402–412 (1996).
[29] J. Rice and J. Determan, ''SWellEx-3 Data Report,'' NCCOSC RDT&E Division, Code 541, Tech Rept. 1995.
[30] W. A. Kuperman and F. Ingenito, ''Spatial correlation of surface generated noise in a stratified ocean,'' J. Acoust. Soc. Am. **67**, 1988–1996 (1980).
[31] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. II.

Sound propagation in concentrated emulsions: Comparison of coupled phase model and core-shell model

Jonathan M. Evans and Keith Attenborough[a)]
Faculty of Technology, The Open University, Milton Keynes MK7 6AA, United Kingdom

(Received 26 January 2002; revised 19 July 2002; accepted 1 August 2002)

The predictions of two models of sound propagation in concentrated emulsions are compared with experimental measurements of ultrasonic velocity and attenuation in emulsions with volume fractions up to 0.7. The core-shell model includes irreversible heat transfer, viscoinertial forces, and multiple scattering [McClements *et al.*, J. Acoust. Soc. Am. **105**, 915–918 (1999)]. This model accounts for the effect of thermal interactions between neighboring particles by introducing an effective medium, and is valid for all volume fractions. The coupled phase model includes irreversible heat transfer and viscoinertial forces, and also is valid for all volume fractions, since it is derived from volume-averaged balance equations [J. M. Evans and K. Attenborough, J. Acoust. Soc. Am. **102**, 278–282 (1997)]. This model has a significantly simpler formulation than the core-shell model and does not require the assumption of an effective medium. The coupled phase model is shown to be a good approximation to the core-shell model when the acoustic radius is small. Despite the fact that it does not include thermal interactions, the coupled phase model is shown to give at least as good agreement as the core-shell model with the experimental data, for all volume fractions, as long as the acoustic radius is less than 0.01. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1510142]

PACS numbers: 43.35.Bf [RR]

LIST OF SYMBOLS

a	radius of particles
c	speed of sound
C, C_p	specific heat at constant volume and constant pressure
f	frequency
k	complex wavenumber $k=\omega/c+i\alpha$
$\tilde{p}$	pressure $\tilde{p}=p+p'\exp[i(kz-\omega t)]$
S, S_T	frequency-dependent momentum and heat transfer factors
$\tilde{T}$	temperature $\tilde{T}=T+T'\exp[i(kz-\omega t)]$
$\tilde{u}$	velocity $\tilde{u}=u'\exp[i(kz-\omega t)]$
α	attenuation coefficient
β	coefficient of thermal (volume) expansion
γ	ratio of specific heats
ϕ_1	volume fraction of continuous phase $1-\phi_2$
δ, δ_T	viscous and thermal boundary layer thickness
$\tilde{\phi}_2$	volume fraction of particulate phase $\tilde{\phi}_2=\phi_2+\phi_2'\exp[i(kz-\omega t)]$
κ	adiabatic compressibility
λ	wavelength
μ	dynamic viscosity
$\tilde{\rho}$	density $\tilde{\rho}=\rho+\rho'\exp[i(kz-\omega t)]$
τ	thermal conductivity
ω	angular frequency

Subscripts

1	continuous phase
2	particulate phase
va	volume-averaged quantity $x_{va}=\phi_1 x_1+\phi_2 x_2$

I. INTRODUCTION

Long-wavelength sound propagation in emulsions is influenced by the density and compressibility of each phase, interphase heat transfer, viscoinertial forces, and the intrinsic absorption of the constituent liquids. Various, related, theoretical approaches have been used to model the relation between these effects and the complex wave number. The most widely successful is scattering theory,[1] which is applicable at arbitrarily high frequencies. Considerable simplifications can be made to scattering theory when the wavelength is much greater than the particle size.[2] Multiple scattering theory[3] takes into account that, when the volume fraction or the extinction cross section or the propagation distance is large, the wave incident on successive particles cannot be assumed to be the same.

Coupled phase theory[4] uses a volume-averaging procedure to account for the fact that the two phases occupy the same control volume. This gives a formulation that is ''self-consistent,'' i.e., it is valid for all values of volume fraction. Because it assumes that the particulate phase is a continuum, it is only applicable in the long-wavelength regime. Coupled phase theory can be modified to account for many different phenomenon, including particle nonsphericity,[5] chemical reactions,[6] or mass transfer[7] between the phases, flow,[8] viscoelasticity,[9] and phoresis.[8]

[a)]Author to whom correspondence should be addressed. Current address: Department of Engineering, University of Hull, Hull HU6 7RX, UK. Electronic mail: k.attenborough@hull.ac.uk

Most theoretical models have assumed that the emulsion particles are isolated and therefore the thermal wave around a particle does not interact with the thermal waves of its neighbors. This is only strictly valid for low volume fractions, or when the ratio of the thermal boundary layer thickness to the particle size is small.

Several researchers have attempted to model interactions between the thermal waves. Fukumoto and Izuyama[10] developed a rigorous model for periodic emulsions. This requires considerable computation and has limited applicability. Hemar *et al.*[11–13] used an approximate ''cell'' model, the ''core-shell'' model. The standard problem of thermal wave scattering from a spherical particle was solved with additional boundary conditions on a ''cell'' concentric with the particle. The boundary conditions represent the effect of neighboring particles. Both Fukumoto and Hemar calculated the complex wave number using a volume integral.

Hemar *et al.*[11] compared predictions of their core-shell model (CS), and predictions of the model neglecting interactions, to measurements of the ultrasonic attenuation in oil-in-water emulsions. The frequency range was 0.5–10 MHz, the particle radius range was 0.046–0.9 μm, and the volume fraction range was 0.05–0.5. The predictions of CS were in very close agreement with these experimental data for the whole range of volume fractions. The inclusion of the interaction effect improved the agreement at the lower frequencies, and only a small discrepancy appeared at the highest frequencies. McClements *et al.*[13] added multiple scattering and viscoinertial effects to the model, which corrected this high-frequency discrepancy. The resulting model will be described here subsequently as the core-shell model (CS). The velocity predictions of the model were not shown, or compared to any experimental data.

There can also be interactions between the viscous waves associated with neighboring particles; these tend to be more important for sound propagation in suspensions. Viscous wave interactions in steady-state flows have been modeled by Kuwabara[14] and Happel.[15] Strout[16] extended this work to oscillating flows to investigate the effect of viscous wave interactions on sound propagation. The wave number for sound propagation in suspensions was calculated from coupled phase theory.

In this paper, attenuation and velocity predictions of the McClements *et al.*[13] core-shell model, including multiple scattering (CS), which includes thermal interactions, are compared with predictions of the Evans and Attenborough[4] coupled phase model (CP), for three different emulsions in the long-wavelength regime, and for volume fractions up to 0.7. Both models are valid for all values of volume fraction, and so they are applicable to these high volume fractions. However, only the CS model includes the effect of thermal interactions, which are expected at higher volume fractions, and when the ratio of the thermal boundary layer thickness to the particle size is large.

II. THEORY

A. Core-shell model

To calculate the complex wave number in an emulsion, McClements *et al.*[13] used the multiple scattering model of Waterman and Truell. This was modified with the Hemar *et al.*[11] expression, to account for thermal wave interactions. The complex wave number in the emulsion k is given by

$$\left(\frac{k}{k_1}\right)^2=\left(1-\frac{3i\phi_2A_0}{(k_1a)^3}\right)\left(1-\frac{9i\phi_2A_1}{(k_1a)^3}\right), \tag{1}$$

where $k=\omega/c+i\alpha$, c is the speed in the emulsion, and α is the attenuation coefficient in the emulsion.

A_0 and A_1 are the monopole and dipole scattering coefficients of a particle, given by

$$A_0=-\frac{ik_1a}{3}\left((k_1a)^2-(k_2a)\frac{\rho_1}{\rho_2}\right)-\frac{k_1^3a(\gamma_1-1)\kappa_1}{\omega\rho_1C_{p1}\kappa_{\mathrm{va}}}\times\left(1-\frac{\beta_2\rho_1C_{p1}}{\beta_1\rho_2C_{p2}}\right)^2B, \tag{2}$$

$$A_1=\frac{i(k_1a)^3}{9}\frac{(\rho_2-\rho_1)(1+S_{\mathrm{CS}})}{\rho_2+\rho_1S_{\mathrm{CS}}}, \tag{3}$$

$$S_{\mathrm{CS}}=\frac{9i}{4}\left(\frac{\delta}{a}\right)^2+\frac{9}{4}(1+i)\frac{\delta}{a}+\frac{1}{2},\quad \delta=\sqrt{\frac{2\mu_1}{\rho_1\omega}}.$$

It should be noted that the numerical factor 9 in the denominator in the expression for A_1 is missing in Ref. 13 [their Eq. (3)]. Here the subscript 1 indicates the continuous phase and the subscript 2 indicates the particulate phase. The subscript va indicates a volume-averaged quantity $x_{\mathrm{va}}=\phi_1x_1+\phi_2x_2$. It should be noted also that the scattering coefficients given here are appropriate for fluid droplets and account explicitly for the difference in fluid densities.

The expression B represents the thermal wave, including interactions:

$$B=\frac{\kappa_1\tau_2}{\kappa_{\mathrm{va}}}\frac{n_2a-\tanh(n_2a)}{EC+FD}\left[2\tau_3\frac{g_1-g_3}{g_2-g_1}n_1b(n_3b+1)+C(1+n_1a)+D(1-n_1a)\right],$$

$$n_j=\frac{1-i}{\delta_{Tj}},\quad \delta_{Tj}=\sqrt{\frac{2\tau_j}{\omega\rho_jC_{pj}}},\quad g_j=\frac{\beta_j}{\rho_jC_{pj}},$$

$$C=\exp(n_1(b-a))[\tau_1(n_1b-1)+\tau_3(n_3b+1)],$$

$$D=\exp(-n_1(b-a))[\tau_1(n_1b+1)-\tau_3(n_3b+1)], \tag{4}$$

$$E=\tau_2n_2a+[\tau_1(n_1a+1)-\tau_2]\tanh(n_2a),$$

$$F=\tau_2n_2a-[\tau_1(n_1a-1)+\tau_2]\tanh(n_2a).$$

Here b is the radius of the ''shell'' (see Fig. 1) and is given by $b=a\phi_2^{-1/3}$. By introducing the volume fraction, through b, the Hemar model becomes valid for all values of the particulate phase volume fraction. The attenuation tends to zero as the volume fraction approaches 0 and 1. The subscript 3 indicates the ''effective medium'' (see Fig. 1) whose physical properties are

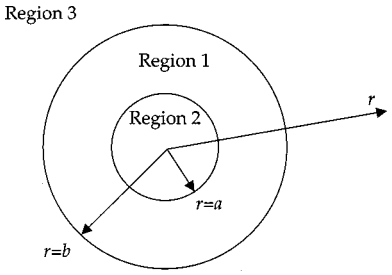


FIG. 1. Core-shell model.

$$\rho_3=\rho_{\mathrm{va}},\quad \kappa_3=\kappa_{\mathrm{va}},\quad C_{p3}=\frac{(\rho C_p)_{\mathrm{va}}}{\rho_{\mathrm{va}}},$$

$$\tau_3=\tau_1\frac{1+2\phi_2\chi-2\phi_1\zeta\chi^2}{1-\phi_2\chi-2\phi_1\zeta\chi^2},$$

$$\chi=\frac{\tau_2-\tau_1}{\tau_2+2\tau_1},\quad \zeta=0.210\,68\phi_2-0.046\,93\phi_2^2.$$

B. Coupled phase model

Reference 4 details the coupled phase model (CP) for sound propagation in emulsions. This is outlined again here but with slight modifications. The expression for the complex wave number is calculated from the linearized, volume-averaged equations for conservation of mass, momentum, and energy, and the equations of state for the continuous and particulate phases:

$$\phi_j\frac{\partial\tilde{\rho}_j}{\partial t}+i^{2j}\rho_j\frac{\partial\tilde{\phi}}{\partial t}+\phi_j\rho_j\frac{\partial\tilde{u}_j}{\partial z}=0,\quad j=1,2, \tag{5}$$

$$\phi_j\rho_j\frac{\partial\tilde{u}_j}{\partial t}=i^{2j-1}\omega\phi_2\rho_1 S(\tilde{u}_1-\tilde{u}_2)-\phi_j\frac{\partial\tilde{p}}{\partial z}, \tag{6}$$

$$\rho_j C_j\frac{\partial\tilde{T}_j}{\partial t}+\frac{\rho_j(\gamma_j-1)C_j}{\beta_j}\frac{\partial\tilde{u}_j}{\partial z}=i^{2j-1}\omega\rho_1 S_T(\tilde{T}_1-\tilde{T}_2), \tag{7}$$

$$\rho_j'+\rho_j\beta_j T_j'-\gamma_j\rho_j\kappa_j p'=0, \tag{8}$$

where S is a factor representing the momentum transfer,

$$S=\frac{9i}{4}\frac{\delta^2}{a^2}+\frac{9}{4}(1+i)\frac{\delta}{a}+\frac{1+2\phi_2}{2\alpha}. \tag{9}$$

This factor represents the effect of viscoinertial forces. It is calculated on the assumption that the particle is rigid and hence does not depend on relative fluid density. The core-shell model has a similar factor [see after Eq. (3)], which shares the first two terms, and has been denoted S_{CS}. The difference between these expressions is in the final, frequency-independent term. This represents the induced mass for a single rigid sphere. In the core-shell model, which has been derived for an isolated sphere, its value is 0.5. In the coupled phase model, an expression is used that accounts for the effect of the other particles on the induced mass of a given particle.[17] In the low volume fraction limit this approaches 0.5.

S_T represents the irreversible heat transfer,

$$S_T=\frac{-3\tau_1}{i\omega a^2\rho_1}\left[\frac{1}{1-iq_1a}-\frac{\tau_1}{\tau_2}\frac{\tan(q_2a)+(3/q_2a)-[3\tan(q_2a)/(q_2a)^2]}{\tan(q_2a)-q_2a}\right]^{-1}, \tag{10}$$

$$q_j=\frac{1+i}{\delta_{Tj}}.$$

It should be noted that the second 3 in this expression was erroneously stated as a 2 in Ref. 4. This expression for the irreversible heat transfer between the phases comes from Gumerov *et al.*[7] It is derived from the single sphere thermal wave scattering problem in exactly the same way as the Hemar *et al.* result. The expression in tan functions for the internal heat transfer (heat transfer in the particle) is of a different form because the original expression has been volume averaged over the particle. This replaces the *microtemperature* field with a single value *macrotemperature*. This is necessary because, in the coupled phase approach, the particulate phase is treated as continuum where the field variables fluctuate only on the acoustic wavelength scale.

The set of Eqs. (5)–(8) can be written as a matrix equation. Previously k was found by numerically inverting the matrix.[4] However, here the exact dispersion equation is derived from the set of equations, by eliminating the field variables,

$$\left(\frac{k^2}{\omega^2}\right)=\frac{(\gamma\kappa)_{\mathrm{va}}(\phi_1\rho_2+\rho_{\mathrm{va}}S)(\phi_1\rho_2C_1C_2+(C\rho)_{\mathrm{va}}S_T)}{(\phi_1\rho_2\rho_b+S)(C\rho)_{\mathrm{va}}S_T+\phi_1\rho_2C_1C_2(\phi_1\rho_2\rho_\infty+\gamma_{\mathrm{va}}S)+\beta_{\mathrm{va}}S_T[(\rho C(\gamma-1)/\beta)_{\mathrm{va}}S+(C(\gamma-1)/\beta)_{\mathrm{va}}\phi_1\rho_2]}. \tag{11}$$

Here $\rho_b=(\rho^{-1})_{\mathrm{va}}$ and $\rho_\infty=(\gamma/\rho)_{\mathrm{va}}$.

Equation (11) shows that the expressions for A_b and A_{eq}, introduced in Ref. 4, can be written more concisely in terms of volume-averaged physical constants:

$$A_b=\rho_b(C\rho)_{\mathrm{va}}+\beta_{\mathrm{va}}\left(\frac{C(\gamma-1)}{\beta}\right)_{\mathrm{va}},$$

$$A_{\mathrm{eq}}=(C\rho)_{\mathrm{va}}+\beta_{\mathrm{va}}\left(\frac{\rho C(\gamma-1)}{\beta}\right)_{\mathrm{va}}.$$

In the CP model, as the core-shell model, it is assumed that the particulate phase size distribution is monodisperse. All comparisons in this paper are in terms of the excess attenuation, so in this model we also neglect the intrinsic absorption of the two liquids.

The expression for the momentum transfer, S, in the coupled phase model, representing the effect of the viscoinertial force, assumes rigid (but movable) particles in a fluid. This implies a large density contrast, since S does not depend on the particle density, and a large relative velocity. However, in the emulsions considered, the density contrast between the two liquids is minor, and there is a small relative velocity. Consequently, the best approximation for the viscoinertial effect in the emulsions is obtained by assuming there is no momentum transfer between the phases, i.e., $S=0$. In the core-shell model this regime is automatically accounted for by the $\rho_2-\rho_1$ term that multiplies the viscoinertial term in Eq. (3). The coupled phase model could be improved by including the correct expression for the viscoinertial force on an oscillating liquid sphere in a liquid.

When $S=0$, the conservation of momentum equations (7) and (8) become

$$\rho_j \frac{\partial \tilde{u}_j}{\partial t} = -\frac{\partial \tilde{p}}{\partial z}, \quad j=1,2,$$

and the wave number from Eq. (11) is given by

$$\left(\frac{k^2}{\omega^2}\right) = \frac{(\gamma\kappa)_{va}(\phi_1\rho_2 C_1 C_2 + (C\rho)_{va} S_T)}{\phi_1\rho_2 C_1 C_2 \rho_\infty + [\rho_b (C\rho)_{va} + \beta_{va}(C(\gamma-1)/\beta)_{va}]S_T}. \tag{12}$$

Strout[16] modified coupled phase theory to account for viscous wave interactions by modifying the momentum transfer factor S so that the frequency-dependent terms were also dependent on the volume fraction. It has not yet been demonstrated how to include viscoinertial interactions in a core-shell model such as that of Hemar and co-workers.

The predictions of Eqs. (1) and (12) for the complex wave number in an emulsion will be compared in Sec. III. This calculation requires the input of values for frequency, volume fraction, particle radius, six physical constants for each phase, and the equilibrium temperature.

C. Comparison of core-shell model and coupled phase model

In principle, both the core-shell model and the coupled phase model are valid for all values of particulate phase volume fraction, but only the core-shell model accounts for thermal interactions. The volume fraction dependence in the coupled phase model is introduced from the beginning by volume averaging the balance equations, following the principles of two-phase continuum mechanics. In the core-shell model, the volume fraction dependence is introduced by assuming an effective medium on the boundary of the cell. CS is the long-wavelength limit of the general scattering model, which is valid for arbitrary frequencies. CP was developed from hydrodynamics and has steady-state low-frequency limits (see below).

From Ref. 13, when thermal interactions are neglected,

$$A_0 = -\frac{ik_1 a}{3}\left((k_1 a)^2 - (k_2 a)^2 \frac{\rho_1}{\rho_2}\right) - \frac{k_1^3 a(\gamma_1-1)\tau_1}{\omega\rho_1 C_{p1}} \times \left(1 - \frac{\beta_2\rho_1 C_{p1}}{\beta_1\rho_2 C_{p2}}\right)^2 H, \tag{13}$$

where

$$H = \left[\frac{1}{(1-iq_1 a)} - \frac{\tau_1}{\tau_2}\frac{\tan(q_2 a)}{\tan(q_2 a) - q_2 a}\right]^{-1} \tag{14}$$

represents the heat transfer. We see that this expression is analogous to the expression in the CP model in square brackets in Eq. (10), i.e.,

$$H \approx \frac{-i\omega a^2 \rho_1}{3\tau_1} S_T \tag{15}$$

or

$$A_0 = -\frac{ik_1 a}{3}\left((k_1 a)^2 - (k_2 a)^2 \frac{\rho_1}{\rho_2}\right) + \frac{i(k_1 a)^2(\gamma_1-1)}{3C_{p1}}\left(1 - \frac{\beta_2\rho_1 C_{p1}}{\beta_1\rho_2 C_{p2}}\right)^2 S_T. \tag{16}$$

As $\omega\to 0$, the first expression in Eq. (14), which represents the external heat transfer,

$$\frac{1}{1-iq_1 a} \to 1.$$

The second expression in (14), which represents the internal heat transfer, grows indefinitely as $\omega\to 0$, so $H\to 0$ and the heat transfer disappears.

The corresponding CP expression, Eq. (10), has the same expression for the external heat transfer, and, as $\omega\to 0$, the expression for the internal heat transfer $\to \tau_1/5\tau_2$ so

$$H_{CP} \to \left[1 + \frac{\tau_1}{5\tau_2}\right]^{-1},$$

which is the steady-state heat transfer result.[7]

Using (3) and (16) we can write Eq. (1) in terms of S_T, S_{CS}, and physical constants, for comparison with the coupled phase model [Eq. (11)]:

$$\left(\frac{k}{k_1}\right)^2 = \left[1 - \phi_2\left(1 - \frac{k_2^2\rho_1}{k_1^2\rho_2}\right) + \frac{(\gamma_1-1)\phi_2}{C_{p1}} \times \left(1 - \frac{\beta_2 C_{p1}\rho_1}{\beta_1 C_{p2}\rho_2}\right)^2 S_T\right] \times \left[1 + \frac{\phi_2(\rho_2-\rho_1)(1+S_{CS})}{\rho_2 + \rho_1 S_{CS}}\right]. \tag{17}$$

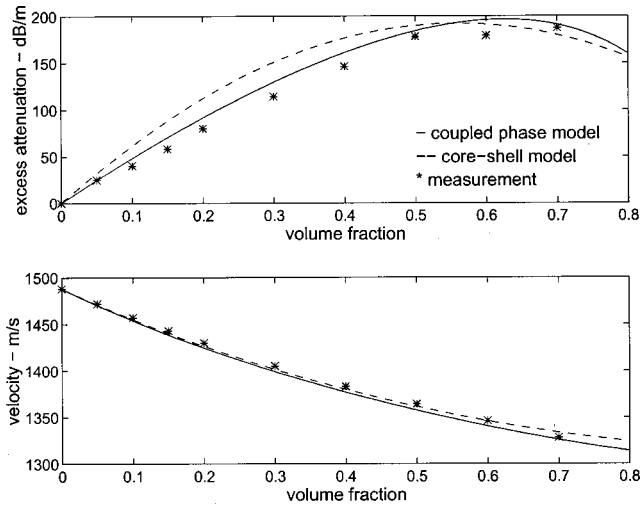


FIG. 2. Excess attenuation in dB m^{-1} and ultrasonic velocity at 500 kHz versus the oil volume fraction for an oil-in-water emulsion with mean particle radius 0.96 μm. A comparison of data of Nesse (Ref. 18) and predictions of the coupled phase model and core-shell model.

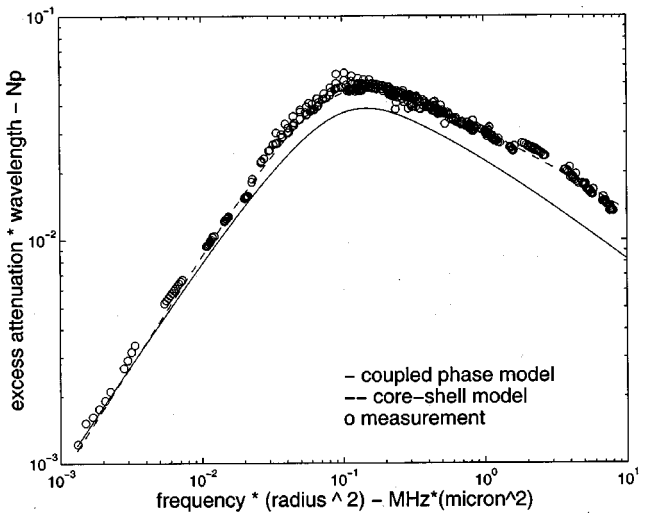


FIG. 3. Excess attentuation×λ in Np vs fa^2 for a hexadecane oil-in-water emulsion with volume fraction 0.2. A comparison of data of Hemar *et al.* (Ref. 13) and predictions of the coupled phase model and the core-shell model.

Neglecting viscous forces and heat transfer, $S_{CS}=0.5$ and $S_T=0$, and neglecting intrinsic absorption, $\alpha_1=\alpha_2=0$, gives the well-known lossless scattering result,

$$c^{-2}=c_1^{-2}\left(1+\phi_2\frac{\kappa_2-\kappa_1}{\kappa_1}\right)\left(1+3\phi_2\frac{\rho_2-\rho_1}{2\rho_2+\rho_1}\right). \qquad (18)$$

If $\rho_1\approx\rho_2$, such as in an emulsion, then (18) becomes

$$c^{-2}=\kappa_{va}\rho_{va}.$$

In the coupled phase model, if $S_T=0$, (12) gives

$$c^{-2}=\frac{(\kappa\gamma)_{va}}{\rho_\infty}.$$

III. COMPARISONS WITH EXPERIMENTAL DATA

Nesse[18] measured the ultrasonic attenuation and velocity for a variety of oil-in-water emulsions. In one experiment the attenuation and velocity in an emulsion of Exxsol D80 with a Sauter mean radius of 0.96 μm were measured in the frequency range 250–900 kHz and for volume fractions from 0.02 to 0.7. The experimental data at 500 kHz are compared to predictions of the coupled phase model (CP) and the core-shell model (CS) in Fig. 2. The predictions are calculated, assuming that the particle size distribution is monodisperse, with all the particle radii equal to the mean particle radius.

The predictions of both models agree well with the measured velocity. There is good qualitative agreement between the predictions of the models and the measured attenuation. The discrepancy in the attenuation predictions is likely to be the result of polydispersity in the experimental emulsion.

One indicator of the interaction effect is the overlap of the thermal boundary layers of neighboring particles. For the Exxsol emulsion at 500 kHz, this occurs for volume fractions above 0.3.[1,18] Reference 18 compared these experimental data with multiple scattering theory predictions (not including interactions), which were found to overpredict the attenuation significantly.

Hemar *et al.*[11–13] measured the ultrasonic attenuation for hexadecane oil-in-water emulsions. The frequency range was 0.5–10 MHz, the radii of the emulsion particles varied between 0.046 and 0.9 μm, and the volume fraction range was 0.05–0.5. The experimental data were compared with predictions of two models: one including interactions (CS) and one neglecting interactions (NI). At the lower frequencies, the predictions of CS were closer to the experimental data than the predictions of NI. At the highest frequencies, the predictions of CS and NI converged, both slightly underpredicting the attenuation. The inclusion of multiple scattering and viscoinertial effects corrected this high-frequency discrepancy.[13] The predictions of CS were in very close agreement with the experimental data for the whole range of volume fractions. This close agreement was achieved by fitting the theoretical prediction versus frequency to the experimental data at a volume fraction of 0.1, using the monodisperse particle radius, which was not measured, as the adjustable parameter. The velocity predictions of the model were not shown or compared to any experimental data.

To determine how well the models predict the frequency, particle radius, and volume fraction dependence of the attenuation, predictions of CP and CS are compared to Hemar's experimental data in Figs. 3 and 4, for volume fractions of 0.2 and 0.45, respectively. The figures plot the product of the attenuation α and the wavelength λ against fa^2. Hemar *et al.* fitted the CS predictions to the data by varying a. For both models, the complex wave number is a function of fa^2, so the predictions of CP have been calculated using the fa^2 values of the fitted CS curve.

The CS predictions agree well with the measured attenuation and velocity in Figs. 2–4. In Fig. 2, the CP model, which neglects momentum transfer between the phases and neglects thermal interactions, agrees well with the predictions of the CS model and the experimental data. Figures 3 and 4 show that predictions of the CP model agree with the measured attenuation and the CS predictions at low values of fa^2, but tend to underestimate the attenuation as the fre-

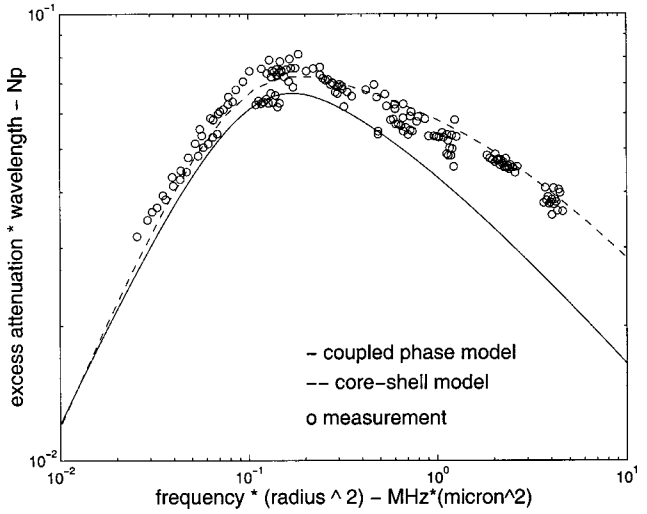


FIG. 4. Exess attenuation×λ in Np vs fa^2 for a hexadecane oil-in-water emulsion with volume fraction 0.45. A comparison of data of Hemar *et al.* (Ref. 13) and predictions of the coupled phase model and core-shell model.

quency increases. This is to be expected, as CP is a long-wavelength model. The validity of the CP model can be assessed by considering the value of the acoustic radius k_1a. For Fig. 2 $k_1a=2.0\times10^{-3}$ and, for Figs. 3 and 4, k_1a varies between 9.8×10^{-5} and 3.8×10^{-2}. From these comparisons it can be concluded that the CP prediction gives a good approximation to the CS prediction when $k_1a<0.01$.

CP agrees with CS at low values of k_1a by including volume averaging but neglecting thermal interactions. This suggests that, in this regime, the main contribution of the interactions in the core-shell model is the volume fraction dependence. For low k_1a, this dependence is reproduced by the coupled phase model, but by means of a significantly simpler formulation. The frequency-dependent contribution of the interactions is less important in this regime.

More generally, both CP and CS are only valid for low values of k_1a. CP represents the particulate phase as a continuum, so the wavelength must be greater than the particle size. CS assumes that the pressure is uniform in the region surrounding the particle, requiring a similar constraint. For this reason, the comparisons here have been restricted to data at low frequencies and particle sizes. The models will not be valid at the higher frequencies at which data are available, or for the larger particle sizes that are encountered in practical situations.

Another, more precise, condition for the validity of these models is the constraint that the acoustic depth $N\sigma L\ll1$. Here N is the number of particles per unit volume, σ is the total scattering cross section of a particle, i.e., the total energy removed from the incident wave by the particle, in this case by irreversible heat transfer and viscosity, and L is the distance the incident wave travels in the emulsion. This condition arises from the assumption that the amplitude of the pressure wave incident on all the particles is the same and equal to the incident wave amplitude.

IV. CONCLUSIONS

The core-shell model relates the complex wave number in an emulsion to the effects of irreversible heat transfer, viscoinertial forces, multiple scattering, and thermal interactions. The coupled phase model relates the complex wave number in an emulsion to the effects of irreversible heat transfer and viscoinertial forces. Both models are valid for all values of volume fraction, but only the core-shell model includes the effects of thermal interactions. Predictions of the core-shell model and the coupled phase model have been compared with measurements of ultrasonic velocity and attenuation in two different emulsions for volume fractions up to 0.7. The coupled phase model used here neglects the effect of viscoinertial forces, or momentum transfer. The reason for this is that the expression for the viscoinertial force in the coupled phase model assumes that the density contrast between the two phases is large. For emulsions the density contrast is low, so the model is more accurate if this expression is neglected. It has been shown that the coupled phase model not only provides a good approximation to the core-shell model for low values of the acoustic radius k_1a, but does so with a significantly simpler formulation.

ACKNOWLEDGMENTS

This work was supported by Grant No. GR/L52505 from the EPSRC (UK). The authors would like to thank Olga Umnova, Norbert Herrmann, and Julian McClements for their assistance and data.

[1] D. J. McClements, "Comparison of multiple scattering theories with experimental measurements in emulsions," J. Acoust. Soc. Am. **91**, 849–853 (1992).
[2] J. R. Allegra and S. A. Hawley, "Attenuation of sound in suspensions and emulsions: theory and experiments," J. Acoust. Soc. Am. **51**, 1545–1564 (1972).
[3] V. Twersky, "On scattering of waves by random distributions. I. Free space scatter formalism," J. Math. Phys. **3**, 700–715 (1962).
[4] J. M. Evans and K. Attenborough, "Coupled phase theory for sound propagation in emulsions," J. Acoust. Soc. Am. **102**, 278–282 (1997).
[5] A. S. Ahuja and W. R. Hendee, "Effects of particle shape and orientation on propagation of sound in suspensions," J. Acoust. Soc. Am. **63**, 1074–1080 (1978).
[6] T. S. Margulies and W. H. Schwartz, "Acoustic wave propagation in fluids," in *Frontiers in Fluid Mechanics*, edited by S. H. Davis and J. L. Lumley (Springer-Verlag, New York, 1985), pp. 219–280.
[7] N. A. Gumerov, A. I. Ivandaev, and R. I. Nigmatulin, "Sound waves in monodisperse gas-particle or vapor-droplet mixtures," J. Fluid Mech. **193**, 53–74 (1988).
[8] W. H. Schwartz and T. S. Margulies, "Sound wave propagation through emulsions, colloids, and suspensions using a generalized Fick's law," J. Acoust. Soc. Am. **90**, 3209–3217 (1991).
[9] T. S. Margulies and W. H. Schwartz, "A multiphase continuum theory for sound wave propagation through dilute suspensions of particles," J. Acoust. Soc. Am. **96**, 319–331 (1994).
[10] Y. Fukumoto and T. Izuyama, "Thermal attenuation and dispersion of sound in a periodic emulsion," Phys. Rev. A **46**, 4905–4921 (1992).
[11] Y. Hemar, N. Herrmann, P. Lemarechal, R. Hocquart, and F. Lequeux, "Effective medium model for ultrasonic attenuation due to the thermo-elastic effect in concentrated emulsions," J. Phys. II **7**, 637–647 (1997).
[12] N. Herrmann, P. Boltenhagen, and P. Lemarechal, "Experimental study of sound attenuation in quasi-monodisperse emulsions," J. Phys. II **6**, 1389–1403 (1996).
[13] D. J. McClements, Y. Hemar, and N. Herrmann, "Incorporation of thermal overlap effects into multiple scattering theory," J. Acoust. Soc. Am. **105**, 915–918 (1999).
[14] S. Kuwabara, "The forces experienced by randomly distributed parallel circular cylinders or spheres in a viscous flow at small Reynolds number," J. Phys. Soc. Jpn. **14**, 527–532 (1959).
[15] J. Happel and H. Brenner, *Low Reynolds Number Hydrodynamics–With*

Special Applications to Particulate Media (Prentice–Hall, Englewood Cliffs, NJ, 1965).
[16] T. A. Strout, ''Attenuation of sound in high-concentration suspensions: Development and application of an oscillatory cell model,'' Ph.D. thesis, University of Maine, 1991.
[17] A. H. Harker and J. A. G. Temple, ''Velocity and attenuation of ultrasound in suspensions of particles in fluids,'' J. Phys. D **21**, 1576–1588 (1988).
[18] O. Nesse, ''Sound propagation in emulsions,'' Ph.D. thesis, University of Bergen, 1998.

Two-frequency driven single-bubble sonoluminescence

Dagmar Krefting, Robert Mettin, and Werner Lauterborn
Drittes Physikalisches Institut, Universität Göttingen, Bürgerstrasse 42-44, 37073 Göttingen, Germany

(Received 14 March 2002; revised 12 July 2002; accepted 25 July 2002)

Sonoluminescing single bubbles driven simultaneously by two harmonic frequencies were recently reported to increase the maximum light output up to a factor of 3 with respect to single mode excitation. In this paper, experimental and numerical results on single-bubble sonoluminescence (SBSL) in an air/water system using the fundamental mode of 25 kHz and the second harmonic at 50 kHz are presented. The region of light emission is mapped in the three-dimensional parameter space spanned by the two driving pressure amplitudes and their relative phase. Good agreement was seen between measured light output, maximum bubble radius, and stability boundaries and the numerical model which is based on spherical bubble oscillations regarding diffusive and shape stability. The maximum brightness was enhanced by a factor up to 2.5 with respect to single mode SBSL. However, long-term measurements reveal great variation of the emission at fundamental mode driven SBSL and of the boost factor reached with two frequencies. The overall brightness maxima of both excitation methods within a period of several hours turn out to show little difference. [DOI: 10.1121/1.1509427]

PACS numbers: 43.35.Hl, 43.25.Yw, 43.35.Ei [AJS]

I. INTRODUCTION

The light pulse emission of a bubble acoustically trapped in a standing sound field is known as single-bubble sonoluminescence, SBSL.[1,2] This effect occurs in a regime of strongly nonlinear radial oscillations of the bubble with a violent collapse. The short-lived extreme conditions in the bubble cause significant interest in the phenomenon, and hard bubble collapses are in fact essential for many observed cavitation effects besides luminescence, for instance erosion or sonochemistry. While the detailed mechanisms of light generation are still under debate (see, e.g., Ref. 3 and references therein), the idea of enhancing the emission has been pursued in the meantime by several authors.[4,5] One hope is to increase collapse temperature and pressure and maybe even facilitate nuclear fusion in the bubble.[6]

In an SBSL experiment, many parameters influence the bubble stability, its oscillation, and the compression during collapse. Not all parameters, however, are readily and independently adjustable from the outside, as they are mainly a question of various stability aspects. In particular, the bubble size (equilibrium or ambient radius R_0, or equivalently the number of molecules in the bubble) is a diffusion-controlled quantity which is indirectly influenced by virtually all external parameters. This is also true of the compression ratio of equilibrium and minimum radius during one oscillation period, $R_0/R_{\min}$ or a similar measure, the expansion ratio described by maximum and equilibrium radius, $R_{\max}/R_0$ (which has been related to the light emission experimentally[7]). Therefore, it is not surprising, even without specific assumptions on light generation, that strong variations of emission have been reported for the change of SBSL conditions.

Let us split the parameters of a typical SBSL setup into those connected with the liquid–gas system, and those concerning the acoustic driving, both essentially independent from each other. The former comprise mainly the liquid and gas types used and their properties (like density, viscosity, surface tension, sound speed, vapor pressure, heat conduction, etc.), the dissolved gas concentrations, and the ambient pressure and temperature. The latter are usually frequency and amplitude of a standing wave mode with sinusoidal time variation. Many investigations regard the physico-chemical properties of the liquid and dissolved gases, which are quite important to infer the light-emitting mechanism, but fewer reports deal with efforts to optimize the driving, which in principle we should be able to combine with any other light enhancement strategy. Among such suggestions are a shift in driving frequency,[5] and adding spikes[6,8,9] or higher harmonics[4] to the applied alternating pressure.

The idea of a multifrequency or even a nonperiodic driving of cavitation bubbles has been investigated by several authors,[10–13] also in connection with SBSL.[4,8,14–16] The only experiments known to us that have shown an increased light signal compared to sinusoidal excitation were reported by Holzfuss *et al.*[4] In that work the relative phase shift between the fundamental (or first) and second harmonic of fixed amplitude was varied, and at certain ''stability islands'' up to a tripling of emitted photons relative to sine excitation could be measured. Ketterling and Apfel[16] interpreted the findings in the context of the dissociation hypothesis, DH.[17,18] The trends of the experiments could be explained well theoretically under consideration of diffusive, shape, and positional stability. For a prediction of the light output, an empirical formula found by Gaitan and Holt[7] was used by Ketterling and Apfel. However, a more complete and direct comparison of data and theory could not be performed because of the lack of additional measurements.

In this article we present results from a systematic study on two-frequency excitation of SBSL. We used a spherical resonator in contrast to the setup of Holzfuss *et al.*,[4] who employed a cylinder. The advantage of a sphere is that ideally all integer multiples of the fundamental mode frequency have a pressure antinode at the geometrical center. Therefore, the displacement of the bubble from the center is small in the case of multimode driving, and the influence of such a

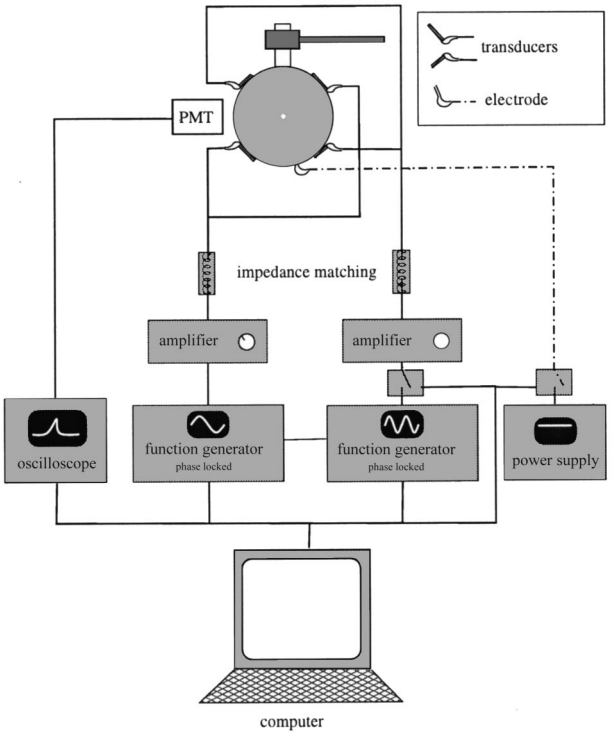


FIG. 1. Experimental arrangement for two-frequency SBSL measurements.

spatial shift on the effective driving pressure amplitude is also small. Throughout our experiments, the positional stability[19,20] was always met near the center of the flask, which is different from the reported cylindrical standing wave.[4] There, a pressure node of the higher mode was found at the cuvette center and a shift of the bubble position was essential for a two-frequency driving.[16] A side effect of this bubble motion was that only effective phase shifts of up to π were realizable in the cylinder. This limitation is overcome by a spherical container.

II. EXPERIMENTAL SETUP

Figure 1 shows a schematic diagram of the arrangement of the experiment. The spherical resonator (outer radius $L=3$ cm) consists of two polystyrene hemispheres put together with a two-component glue. The cell was fixed on top by an attached acrylic neck of diameter 10 mm and length of 2 cm, shut by a removable plastic cap. The acoustic signal was transmitted by two pairs of piezoceramic transducers bonded with epoxy to the wall on opposite sides, one pair for each frequency. The first and the second harmonic were separately generated by two phase-locked function generators (HP 33120A), boosted by two common amplifiers (SONY TA-F445R, Krohn & Hite DCA-50) and connected via impedance matching networks to the transducers.

Our liquid–gas system was a standard choice of degassed water and air at normal ambient pressure and temperature. For maintaining room temperature, a fan was positioned above the resonator. The water (demineralized and filtered) was degassed outside the resonator and then carefully filled in through a system of small latex tubes. However, the filling process always caused a certain reflux of air to the liquid. As the gas content of the water could not be measured inside the sphere, it was determined by an oximeter in the degassing container directly before and after the resonator filling procedure. The results gave an upper and lower threshold for the real value in the cuvette. Therefore, the gas content entered the calculations as an unknown parameter which could be adjusted within a certain given range.

Bubbles were seeded by electrolysis at a vanadium electrode, which was installed at the bottom of the cell and triggered by a short current pulse from a controllable power supply. The light output of the luminescing bubble was sensed by a photomultiplier (PMT, Hamamatsu R 5600U-06) and digitally recorded by an oscilloscope (LeCroy LC 574 AM). All devices were connected to a computer that allowed program-controlled change of settings and acquisition of data.

The acoustic pressure amplitude was measured by a fiber-optic probe hydrophone (FOPH 300),[21] the spatial sound-field distribution was determined using a PVDF needle hydrophone (Imotec). Dynamic bubble-size sequences were accessible via high-speed optical recordings (Imacon 468) through a long-distance microscope (Questar QM 100). By this method, the maximum radius could be extracted.

The fundamental mode frequency of the resonator, which we want to call the ''first harmonic,'' was found at $f_1=24.76$ kHz. This fits very well to the theoretical expectations for the first j_0 spherical Bessel mode. [The nth mode is ideally found by $j_0(k_nL)=0$ with $k_n=2\pi f_n/C$ and $k_nL=\xi_n$, C being the sound speed in water and ξ_n the nth zero of j_0. Because the spherical Bessel function $j_0(kL)$ can be written as $\sin(kL)/kL$, we get $\xi_n=n\pi$ and an equidistant frequency spacing $f_n=nC/2L$.] The center frequency of the second j_0 spherical Bessel mode appeared at a slightly lower value than $2f_1$, probably because of damping by surface waves of the plastic sphere. However, a driving at exactly $2f_1$ was quite possible, and only near the resonator wall were some deviations from the ideal j_0 mode shape detected. For more details of the experimental setup we refer to Ref. 22.

III. NUMERICAL MODELING

We compare the experimental results with simulations based upon a spherical bubble model. We used a Keller–Miksis model in a formulation similar to Ref. 23 with a van der Waals hard core[24] to describe the evolution of the bubble radius R vs time

$$\left(1-\frac{\dot{R}}{C}\right)R\ddot{R}+\frac{3}{2}\dot{R}^2\left(1-\frac{\dot{R}}{3C}\right)=\left(1+\frac{\dot{R}}{C}\right)\frac{p_l}{\rho}+\frac{R}{\rho C}\frac{dp_l}{dt}, \tag{1}$$

$$p_l=p_g-\frac{2\sigma}{R}-\frac{4\mu}{R}\dot{R}-p_0-p_a(t), \tag{2}$$

$$p_g=\left(p_0+\frac{2\sigma}{R_0}\right)\left(\frac{R_0^3-bR_0^3}{R^3-bR_0^3}\right)^{\gamma}. \tag{3}$$

The parameter values used correspond to air bubbles in water at 20 °C and isothermal compression (sound speed C

$=1482\,\mathrm{m/s}$, density $\rho=998\ \mathrm{kg/m^3}$, surface tension $\sigma=0.072\,75\ \mathrm{N/m}$, viscosity $\mu=0.001\ \mathrm{Ns/m^2}$, ambient pressure $p_0=100\,\mathrm{kPa}$, van der Waals coefficient $b=0.0016$, and polytropic exponent $\gamma=1$). The driving pressure is taken to be the superposition of the first two j_0 modes at the center of the sphere, i.e., $p_a=P_1\cos(2\pi f_1 t)+P_2\cos(4\pi f_1 t+\varphi)$, where the amplitudes P_1 and P_2 and the relative phase φ are externally controlled parameters. Additionally, the relative gas saturation in the liquid far from the bubble, c_∞/c_0, is set to a value determined from the degassing pressure p_d, the ambient pressure at measurement, and the relative volume fraction of inert gas in air: $c_\infty/c_0=(p_d/p_0)\times 0.01$. This expression is in accordance with Henry's law and the hypothesis of dissociation, further reaction, and solution of the noninert gases.[17,18] Given a set of parameters $(P_1,P_2,\varphi,c_\infty/c_0)$, our numerical algorithm determines a diffusively stable equilibrium radius R_0 of the bubble, if it exists. This is done by an automatic search, bracketing, and bisection to fulfill the diffusive stability requirements[16,25–29]

$$\overline{c}(R_0,P_1,P_2,\varphi)\equiv\frac{\langle p_g R^4\rangle_T}{p_0\langle R^4\rangle_T}\stackrel{!}{=}\frac{c_\infty}{c_0},\quad \frac{\partial\overline{c}}{\partial R_0}>0, \tag{4}$$

where $\langle\cdot\rangle_T$ denotes the time average over an oscillation period.

Once a stable equilibrium $R_0(P_1,P_2,\varphi,c_\infty/c_0)$ is found, it is checked for shape stability. One standard procedure is to consider the linearized motion of small surface mode amplitudes a_m perturbing the spherical shape. With inclusion of viscous damping (derived for the case of constant R), we get[30–32]

$$\ddot{a}_m+\left[3\frac{\dot{R}}{R}+2(m+2)(2m+1)\frac{\mu}{\rho R^2}\right]\dot{a}_m+(m-1)\times\left[(m+1)(m+2)\frac{\sigma}{\rho R^3}-\frac{\ddot{R}}{R}+2(m+2)\frac{\mu\dot{R}}{\rho R^3}\right]a_{\dot{m}}=0,\qquad m=2,3,\ldots. \tag{5}$$

By integration over the radial oscillation period we obtain the Floquet multipliers, and the one with the maximum absolute value, $\lambda_{\max}$, determines the parametric stability of the considered mode m. In experiments, however, the bubbles turn out to be more stable than predicted by this approach.[33,34] Extensions of the formulation have been given,[35–38] but they are difficult to implement numerically, and approximations are partly under debate.[39,40] Here, we use a pragmatic approach to keep the calculations simple and to compensate for the apparently underestimated damping mechanisms (mainly viscous effects, but also contributions from mode coupling and sound radiation). We regard only the usually most unstable $m=2$ mode and may allow the $\lambda_{\max}$ threshold for parametric stability to have a value greater than 1 (which would be the strict parametric shape stability border). Thus, the bubble destruction threshold is represented by a $\lambda_{\max}$ isoline, e.g., $\lambda_{\max}=1000$.

Finally, a certain light output is assigned to the diffusively and spherically stable bubbles. Following Gaitan and Holt,[7] the number of (detected) photons n_{ph} per (equilibrium) bubble volume scales approximately as $(R_{\max}/R_0)^{4.1}$,

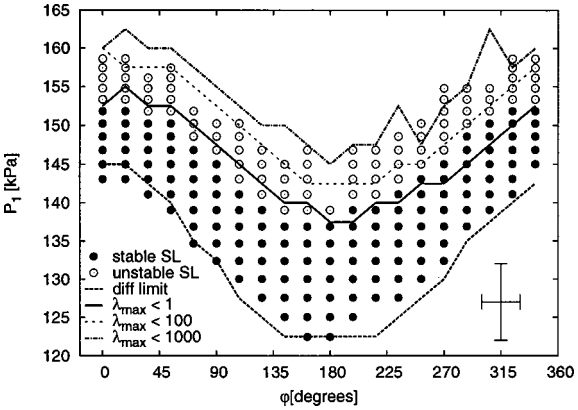


FIG. 2. Phase diagram of variation of P_1 and φ at $P_2=20\,\mathrm{kPa}$. Solid dots denote stable SBSL, open dots unstable SBSL. Different calculated stability boundaries are given for $P_2=15\,\mathrm{kPa}$ and $c_\infty/c_0=0.15\%$: Lower limit of diffusively stable (Ar-) bubble (diff limit), parametric shape stability limit ($\lambda_{\max}<1$) and limits for different maximum Floquet multiplier ($\lambda_{\max}<100$, $\lambda_{\max}<1000$). The crosshairs represent the experimental error.

where the number of photons was set proportional to the rms voltage of the photomultiplier, measured with a lock-in amplifier.[41] In our experiment, the measured quantity scales as well with the number of photons (within the detection wavelength window of 250 to 550 nm), if the light pulse width (typically 60 to 250 ps[42]) is assumed to be much shorter than the impulse response of the PMT ($\approx$1 ns). Therefore, we adopt the above relationship (with a rounded exponent and the ambient bubble volume taken out) to predict the total light output (i.e., the averaged PMT peak voltage $\overline{U}_{\mathrm{PMT}}$)

$$\overline{U}_{\mathrm{PMT}}\propto n_{ph}\propto R_0^3\left(\frac{R_{\max}}{R_0}\right)^{4.1}\approx\frac{R_{\max}^4}{R_0}. \tag{6}$$

We do not employ a "Mach criterion"[16,28] for the onset of SBSL. As we checked *a posteriori*, in our model the bubble wall in any case typically reaches the gas sound speed during collapse for a bubble in diffusive (noble gas) equilibrium.

IV. RESULTS

A. Variation of fundamental amplitude and phase difference

Due to technical simplicity and better automation of the measurements, we decided to fix the amplitude of the second harmonic ($P_2=\mathrm{const}$) and to successively increase the amplitude of the first harmonic P_1 at a given phase difference φ. The phase is altered afterwards in the "outer loop" of the parameter scanning procedure. First, we present a comparison of experiment and numerics with respect to the light emission, then we compare measured and predicted maximum radii. Finally, we refer to the comparison of two frequency with pure sinusoidal excitation in connection with results of long-term experiments.

1. Light emission

A representative result of the indicated type of (φ,P_1) scan is given for a fixed $P_2=20\,\mathrm{kPa}$. Figure 2 shows the SBSL region; Fig. 3 gives the light output for selected phase

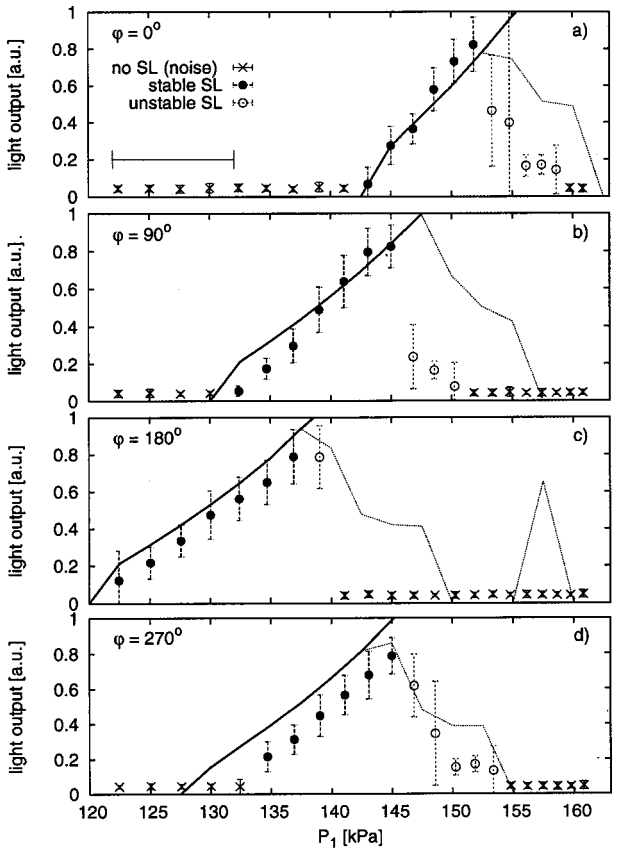


FIG. 3. Light emission at different φ from the same experiment and simulation as in Fig. 2. Experimental data (dots) are normalized to the maximum value of the whole phase scan; the simulated light output is in arbitrary units. Calculated stable SBSL emission is denoted by the solid line. Calculated emission for parametrically unstable SBSL ($1<\lambda_{max}<1000$, dashed line) is scaled by the shape instability measure λ_{max} (by dividing the light output by $\log_{10}(\lambda_{max})+1$; see the text). The isolated theoretical SBSL at $\varphi=180°$, $P_1=155$ kPa is caused by an island of low λ_{max} (see the text).

differences φ. The measured light output is the average over 30 flashes of the peak PMT voltage. It is given in arbitrary units.

In the course of raising P_1 for the phase difference held constant, a trapped bubble is typically dark at low amplitude, then starts to glow, increases its size and its light emission, and finally is destroyed after a condition of ''cyclic'' or ''unstable'' SBSL.[1,28] Thus, the region of light emission is bordered by a ''first light'' boundary at lower pressure amplitude, and an extinction threshold at higher amplitude. The required amplitude P_1 to get first SBSL is minimal at a phase of $\varphi\approx180°$, and it increases almost linearly depending on the phase until the maximum at $\varphi\approx0°$ $(=360°)$. The same holds for the onset of unstable SBSL and the bubble destruction threshold, which leads to a characteristic V-shape of the luminescence zone in the (φ,P_1) plane. The width ΔP_1 of the zone is slightly larger around $\varphi=180°$ ($\Delta P_1\approx20$ kPa) than at 0° ($\Delta P_1\approx15$ kPa).

From lower to higher P_1 the light emission mostly grows monotonically as is known from experiments with pure sinusoidal excitation. Then, a region follows where the (averaged) light output decreases with increasing pressure amplitude until the bubble is destroyed. The moment of destruction is given by the border of data values.

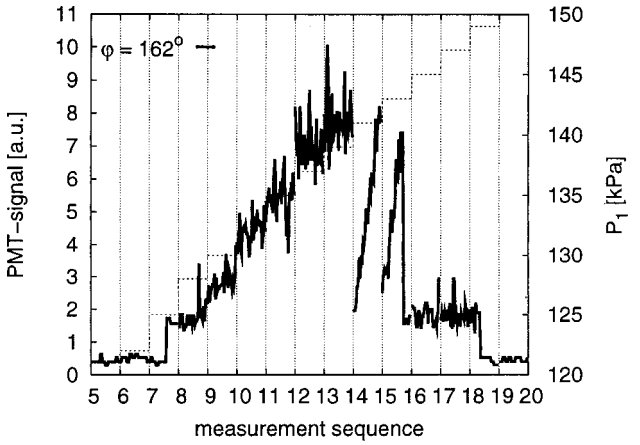


FIG. 4. Measured PMT signals at $\varphi=162°$ (solid line, from the same measurement as in Fig. 2). The amplitude P_1 is increased in steps that define the measurement sequence (dotted horizontal lines, step duration roughly 2 seconds). During each sequence, at constant parameters 30 individual measurements of SBSL flashes are sampled. Between the sequences, a switching and initialization of 0.2-s length takes place (depicted by the vertical lines).

The decreasing averaged light output is connected with so-called cyclic or (diffusive) unstable SBSL states: the bubble size (and its emission) starts growing in time because of rectified diffusion until a surface instability is met, which happens before the diffusive equilibrium size is reached. After fragmentation and splitting of microbubbles, the main remnant continues glowing, but with lower emission. Then, a growing cycle starts again until the next splitting. The mean output is therefore lowered, although the maximum brightness may be even larger than at stable SBSL states.

The changing SBSL behavior is illustrated in Fig. 4, where the light signal for a P_1 scan with fixed φ is shown. The series of PMT data represents the time evolution of the light emission. After onset in sequence 7, SBSL is observed which we call stable in the sense of relatively small variation of the signal. This behavior continues until sequence 13 while the brightness increases. Afterwards (sequences 14 and 15), alternating breakdown and increase of emission can be seen. This causes a large variation and a lower mean value of the signal. Although dramatic drops occur only beyond sequence 13, the variance of the pulses already seems to get larger earlier. This may indicate the gradual loss of spherical stability of the bubble oscillation, if one adopts results from laser-induced bubble luminescence.[43] Additionally, it can be noted that breakdowns appear most frequently at the moment of parameter switching, as especially large perturbations are caused by the stepped change of the pressure. It is noteworthy that we sometimes observed a plateau of emission beyond the typical unstable SBSL, but before total bubble destruction (as in Fig. 3, sequences 16 and 17). Our interpretation is that these data indicate the light emission of bubbles dynamically stabilized (by frequent microbubble ejection) at a relatively small size. This would correspond to unstable SBSL states of very short shrink-and-grow cycles.

The numerical simulations of bubble stability properties give reasonable results. The essential V-shape of the luminescence region can be reproduced well; see the lines in Figs. 2 and 3. The diffusion limit line in Fig. 2 indicates a

lack of diffusively stable Ar bubbles below (we assume that any such stable bubble would emit light). The upper boundaries are determined by the maximum allowed absolute value of the Floquet multipliers, λ_{max}.

Three parameters have been slightly adjusted in the numerical simulation to test their influence and to find best agreement with the observed experimental stability region: (i) the gas saturation c_∞/c_0, which was known only inaccurately; (ii) the second harmonic amplitude P_2, as the pressure amplitude measurements had a relatively large absolute error (10 kPa); and (iii) the Floquet multiplier λ_{max}, as the shape instability onset is probably estimated too low. We found that the largest effect of an increase in gas saturation was to move the whole V-shape down to lower P_1 amplitudes. An increase in P_2 basically stretches the V-shape in the vertical (P_1) direction, and an increase of λ_{max} extends the stability zone to larger P_1. The extension of the experimental stability region from Fig. 2(a) could be best reproduced with $c_\infty/c_0=0.15\%$ and by scaling the measured second harmonic amplitude P_2 by 3/4, i.e., $P_2=15$ kPa instead of $P_2=20$ kPa. The calculated stability boundary $\lambda_{max}=1$ is to be taken as the transition to unstable SBSL. If one raises λ_{max} to a value of 1000, the absolute stability boundary is best reproduced. Unstable SBSL states are not simulated as they are not stationary with respect to the bubble size. To take into account that bubbles perform unstable SBSL before they are destroyed, we have divided the calculated light output for $1<\lambda_{max}<1000$ by $\log_{10}(\lambda_{max})+1$. Bubbles beyond $\lambda_{max}=1000$ are assumed to be destroyed. Thus, the predicted average emission is reduced as the bubble oscillation becomes more shape unstable. This argument assumes that the destruction threshold of the growing bubbles (after a bubble size reduction and emission breakdown) is reduced as well. The emission characteristics of unstable SBSL seem to be captured reasonably well by this approach. Note that it is possible to find stability islands, i.e., driving parameters that allow nondestructed bubbles beyond a ''forbidden'' gap as it appears in Fig. 3(c) at $P_1=155$ kPa. The evolution of the simulated light output with P_1 is very similar to the observed emission. Most photons have been observed at the upper boundary of the left wing of the V, around $\varphi\approx 60°$. The left wing appears brighter in the simulations as well. The φ position of maximum emission is, however, not unique in long-term measurements, as will be shown later.

2. Maximum bubble radius

To get more information about the system from a different observable, we have extracted the maximum radius R_{max} of stable bubble oscillations from microscopic high-speed recordings (Imacon camera and Questar microscope). As these recordings are more costly than the PMT measurements, only one representative and more coarse (φ,P_1) parameter scan was done for a fixed second harmonic amplitude $P_2=25$ kPa. Figure 5 gives the experimental data (symbols) and numerical results (lines) at different φ. Again, the value of P_2 has been reduced in the calculations (to 20 kPa) and c_∞/c_0 has been set to 0.3%. The solid lines indicate stable bubbles with $\lambda_{max}\leqslant 1$ and dotted lines mark the the region $1<\lambda_{max}\leqslant 1000$.

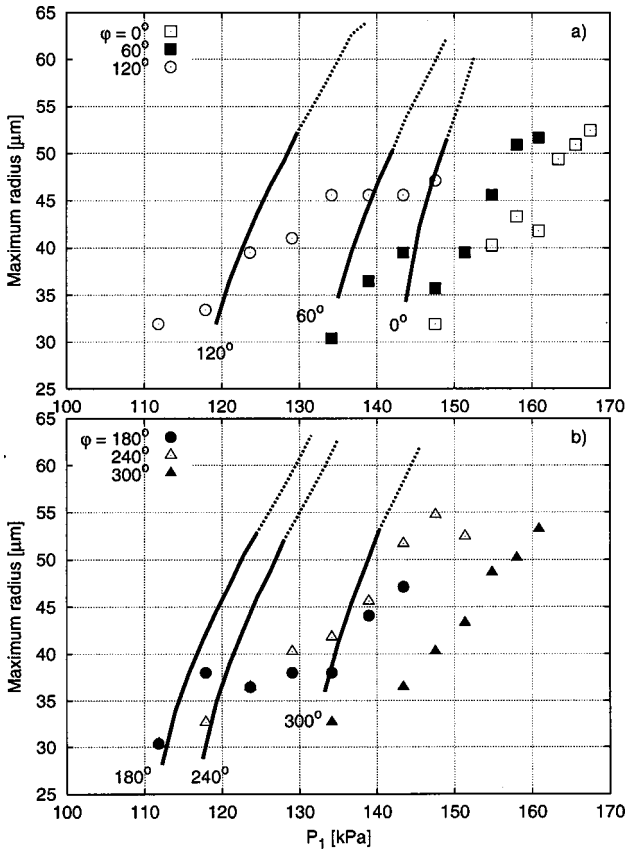


FIG. 5. Measured (symbols) and calculated (lines) maximum radius R_{max} for different P_1 and φ. $P_2=20$ kPa. At the stability border $\lambda_{max}=1$, the lines are extented by thinner segments up to the limit $\lambda_{max}=1000$.

Very similar to the pattern for light emission, the maximum radius increases with an increasing P_1 at fixed φ. This is well reflected by the calculations. The diffusion stability limits, i.e., the smallest R_{max}, are adequately reproduced by the model. The range of calculated stable R_{max} is also found to agree with the observations for $\lambda_{max}=1$. However, the slope of the R_{max} vs P_1 curve is overestimated and the simulated bubbles become shape unstable at lower P_1 values (up to a difference of 20 kPa). We note that in these experiments the P_1 measurement could only be done with a higher uncertainty than otherwise.

To estimate the bubble's inner conditions, such as the temperature at emission, the relative photon emission per molecule (or unit volume) can be employed. Unfortunately, the rest radius R_0 was not accessible in the experiment, and we can only rely on the simulations to yield the expansion ratio, which in turn predicts the relative emission with Eq. (6) after Ref. 7.

The calculations show an increase of R_0 similar to R_{max}, but an overall decrease of R_{max}/R_0 for raised P_1. This is in accordance with previous measurements and calculations[7,16] and yields maximum expansion ratios just after the onset of SBSL (at the lower boundary of the V). This implies that the brighter emission at the top border of the V-shape is due to the bubble containing more molecules, but not to elevated temperature. Interestingly, the compression ratio R_0/R_{min}, although only weakly changing, displays rather the opposite

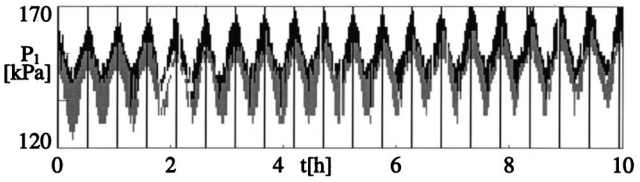


FIG. 6. SBSL phase diagram of a long-term measurement at $P_2=20$ kPa. Gray denotes stable SBSL; unstable regions are filled with black. Each V-shaped segment is a complete (φ,P_1) scan from $\varphi=0°$ to $\varphi=360°$. The horizontal axis is labeled with the elapsed time; a complete scan took slightly over 30 min.

to the expansion ratio: it is typically higher for the larger and brighter bubbles at the top of the V.

3. Long-term measurements

We performed automatic repetitions of the parameter scans to test for reproducibility of the results. The light output of complete (φ,P_1) scans was recorded up to 20 times in a row, over a period of several hours. In Fig. 6 one of several series taken is presented.

It can be seen that the V-pattern of the SBSL region is quite well reproduced by every scan. The details of the emission strength, however, may vary with time, as seen in Fig. 7. While the main figure of a darker bottom line and an increase of brightness towards higher P_1 is preserved, the absolute light output and the parameters at maximum emission change. This gives evidence for some statistical noise in the data, which is best explained by random perturbations of the bubbles influencing their destruction. Additionally, systematic effects can be observed. A pronounced feature is the decrease of the SBSL regime at about 2 h in Fig. 6; a typical evolution of light emission in this region is given in Fig. 7(c). After ca. 3 h, the pattern regains stability. A similar emission drop appeared in all our long-term parameter scans after some hours. Figures 8 and 9 show an emission recording for pure sinusoidal excitation at the fundamental mode (repeated simple P_1 upward scans at $P_2=0$). The bubbles

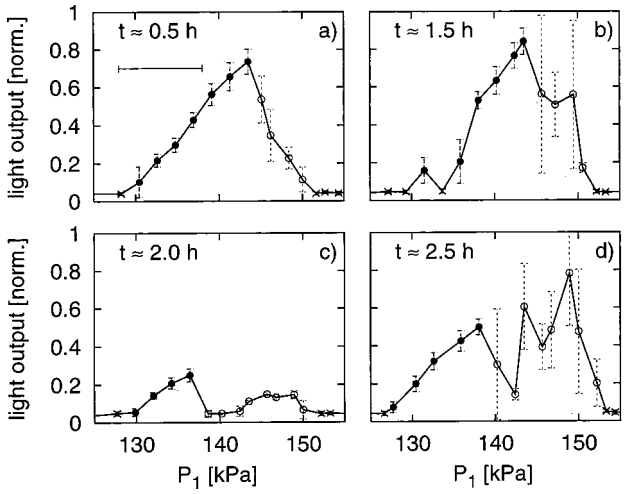


FIG. 7. Sample of light output evolution from one long-term experiment (see Fig. 6) at constant phase ($\varphi=234°$). Time t gives the time relative to the beginning of the experiment. The meaning of the symbols follows Fig. 3. The light output is normalized to the maximum emission of the entire measurement.

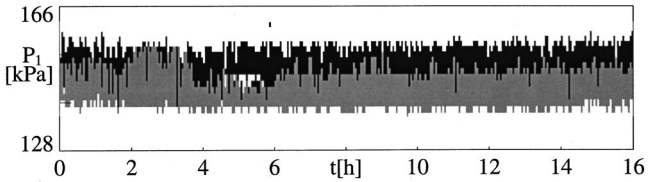


FIG. 8. Phase diagram for fundamental mode driving only ($P_2=0$). Gray denotes stable SBSL, black gives the region of unstable SBSL. The data set is part of a long-term comparative measurement of single- and two-frequency driving (compare Fig. 11).

become brighter and the measurements less scattered after 2 h, and from hours 4 to 6 an emission drop appears [Fig. 9(c)]. After 6 h, the data are more noisy, but remain regular until the end of the sequence.

As the reason for the change in SBSL behavior on the expanded time scale, we postulate a slow drift of temperature and/or gas content of the water. The experiment was cooled by a fan, but a temperature drift of about 2 °C could be observed in some runs. Typically a temperature rise would be accompanied by a shift of the stability region to larger excitation voltages, because the fixed fundamental driving frequency would no longer meet the resonance condition, and higher driving voltage at the transducers would be needed for the same effective pressure amplitude. Since the values of P_1 are derived from the voltage after a calibration in resonance, temperature rise would shift the observed emission pattern to an artificially large P_1. Such a feature can indeed be perceived to a small extent in Fig. 6, but not in Fig. 8.

The gas content of the water is increased by air leakage and by chemical reactions. Electrolysis for seeding bubbles creates O_2 and H_2, and dissociation and reaction of O_2, N_2, and H_2O during the collapse produce a further variety of molecules dissolving in water, according to the dissociation hypothesis.[17,18] The influx of air is a source of argon and would shift the relative gas saturation which is relevant for the bubble equilibrium size. Calculations show that, as a consequence, the V-shaped stability region would move slightly downwards in the (φ,P_1) plane and become thinner.

Another effect of temperature increase is the decrease of

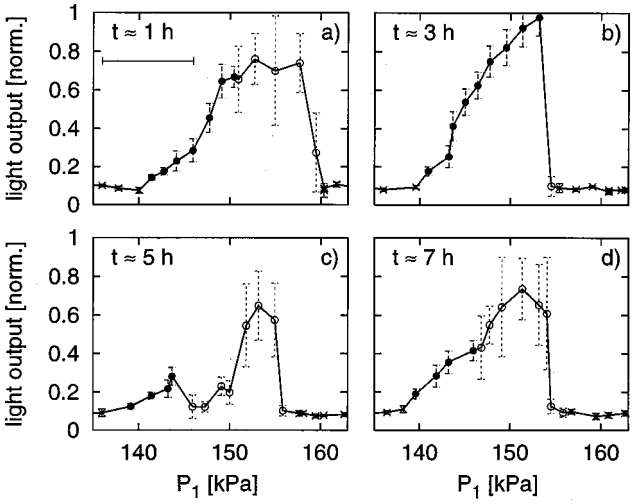


FIG. 9. Sample of light output evolution from the single mode long-term experiment in Fig. 8 (compare Fig. 7).

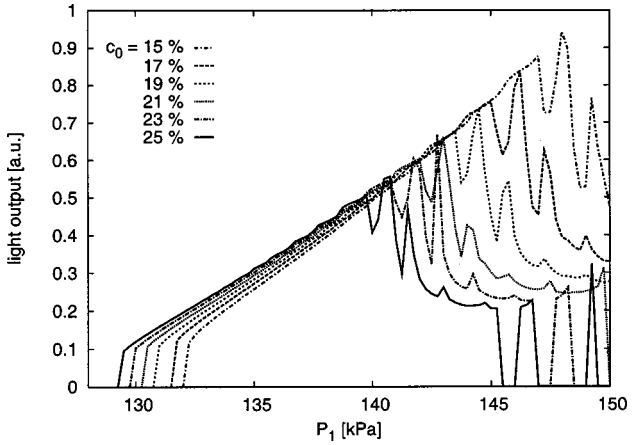


FIG. 10. Calculated light output at different air saturations of the water for fundamental mode (P_1) driving. Unstable SBSL emission is simulated as in Fig 3.

solubility of argon in water (about 5% less for a 2 K rise). This acts in the same direction as an Ar influx, because the saturation concentration c_0 goes down and the relative Ar saturation c_∞/c_0 goes up, respectively. Furthermore, the water vapor pressure increases roughly by 10% per 2 K.

It is difficult to estimate the combined effect of temperature rise and air leakage since they are partly counteracting. The influence of the addition of other gases and reaction products is probably negligible, but increased water vapor is likely to contribute to changes of the light emission.[44–46]

To demonstrate that even a small change of the relative gas saturation may lead to a nonmonotonic alteration of the observed maximum luminescence, we present a simulation for pure sinusoidal excitation. In Fig. 10 we see the calculated light output for different values of relative air saturation of the water (corresponding to $100\times c_\infty/c_0$). As the experimental value was not precisely measurable, we give a sample calculation for 15% to 25%. We modified the light output for unstable SBSL $1<\lambda_{\max}<1000$ as before.

While the lower boundary for emitting bubbles (the diffusive stability threshold) is smooth and simply decreases slightly with increasing air percentage, the upper border possesses more features because of the inclusion of the unstable SBSL effect depending on the degree of shape stability. In fact, the maximum Floquet exponent $\lambda_{\max}$ varies nonmonotonically and shows resonant behavior with valleys and hills including peaks. Consequently, an increase in air saturation can lead to the observation of ups and downs in the maximum brightness (although the general trend is a decrease of light). We conclude that the observed SBSL emission drop (and recovery) during long-term experiments could be explained by stability changes due to a high sensitivity to the relative air (Ar) concentration. Air leakage and temperature fluctuations (via change of gas solubility) can lead to this effect.

To investigate how to boost the light output by two-frequency driving, we performed long-term measurements with alternating excitation type. Again, several successive (φ,P_1) scans were completed as described before, now at $P_2=40$ kPa. Additionally, after every upward scan of P_1 at a fixed phase of the second harmonic, a P_1 upward scan with

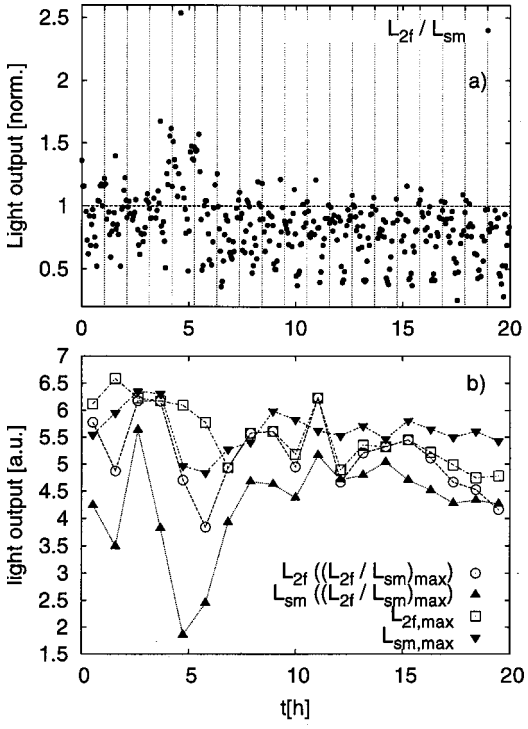


FIG. 11. (a) Maximum achieved light output of two-frequency (L_{2f}) driven SBSL at $P_2=40$ kPa, normalized to the maximum light output of SBSL driven at fundamental mode only (L_{sm}), probed directly afterwards (shown in Fig. 9). The horizontal axis gives the elapsed measurement time; vertical lines indicate completed phase scans. (b) From the same measurement, certain absolute light output levels have been extracted. Note that the data point showing maximum light output in one completed phase scan $L_{2f,\max}$ (open square) need not be the one that shows the maximum boost (open circle) relative to the subsequent single mode light emission (upward triangle). The maximum boost often appears with low single mode light output, while the maximum light output in a phase scan subset $L_{sm,\max}$ (downward triangle) exceeds $L_{2f,\max}$ in most cases.

$P_2=0$ was performed to check the maximum light emission for sinusoidal driving. The relative brightness of the maximum with second harmonic, normalized to the brightness without second harmonic measured directly afterwards, is plotted in Fig. 11(a) versus the elapsed time. The data show large scatter because all scanned phases are included, not only the ''best'' phase of a complete (φ,P_1) scan. Vertical lines mark the endings of full phase turns (overall 19 in 20 h). In almost every such (φ,P_1) scan, adding the second harmonic results in relative boosts, i.e., between nearly all vertical lines values above unity appear. Most of the time, however, the gain factor is rather small and stays well below 1.5. In the period between 4 and 6 h of measurement, a pronounced gain emerges with an optimum of 2.5. Additionally, it turns out that the optimal phase φ is not unique, but distributed over the whole circle with some preference around 75°.

A more differentiated picture of the long-term evolution of the maximum luminescence can be obtained from the absolute emission values given in Fig. 11(b). For each completed phase turn we show the following data: (i) absolute maximum at two-frequency driving (open square) over all phase differences; (ii) absolute maximum at monofrequent driving (downward triangle); (iii) the subsequent pair of two-frequency (open circles) and single-frequency (upward tri-

angles) excitation values that gave the largest relative boost within that phase turn. The vertical interval between the upward and downward triangles gives an impression of the variation of the maximum light output at single-frequency driving, which is quite large. It is found that large boost factors appear mainly because of a poor monofrequent maximum brightness (at that time), while the longer-term monofrequent maximum is seldom exceeded by the highest two-frequency brightness. Mainly around the 5-h time mark the two-mode excitation results in significantly higher absolute light output than the fundamental mode driving. This is the same time interval where the largest relative boost was measured, but also where a regime of minor shape stability and lower emission of the single mode excitation was found (compare Figs. 8 and 9). Thus, one might speak of stabilization rather than of amplification of the absolute output by the second harmonic.

B. Variation of second harmonic and phase

In analogy with Holzfuss *et al.*,[4] we undertook (P_2,φ) scans of the three-dimensional parameter space. For an overall fixed amplitude P_1, the second harmonic amplitude P_2 was increased stepwise. At each step the relative phase φ was scanned from 0° upwards to 360°. Directly afterwards (after a possible necessary bubble reinjection), it was decreased downwards to 0°. Then, the next P_2 step took place. The results are shown in Fig. 12 for $P_1=140$ kPa. We normalized the light output to the averaged value in absence of the second harmonic (at $P_2=0$). The amplitude of P_1 was chosen to get a stable and noticeable emission, but not the maximum; the measured value was $P_1=140$ kPa. Adding the second harmonic weakly we found a brightness minimum around $\varphi=0°$ (=360°) and a maximum around $\varphi=180°$ with monotonic behavior in between. This tendency increases with higher amplitude of the second harmonic. It leads to regions without detectable light output around $\varphi=0°$ and events of high light output and sometimes bubble destruction around $\varphi=180°$. The maximum in light output is shifted to slightly larger phases for larger P_2. At brighter emission the variance of the sampled PMT voltage maxima appeared to be increased, but it was lower than for a typical unstable SBSL state. The lines in Fig. 12 give the results of a numerical simulation for $P_1=140$ kPa and an air saturation of 15%. The shape of the region of luminescence is well reproduced. The absolute values are better approximated by the numerical results at lower P_2; good results are given by a reduction of 3/4 of the measured P_2, which is consistent to the result of the (φ,P_1) scan (see Fig. 3). At high P_2, the real bubbles seem more stable than the calculations predict, and the luminescence seems to achieve some saturation. The measured light output for higher P_2 does not significantly exceed the values at $P_2=20$ kPa. Unstable SBSL is not observed in parameter ranges where the calculated maximum eigenvalue is higher than 1 (visible by the drops in the numeric curves). Some bubbles indeed disappeared near $\varphi=180°$, which is an indication of high sensitivity to perturbations, but other bubbles survived the full phase shift, especially the downtuning. In comparison to amplitude steps, the phase changes correspond to smoother parameter transitions. This seems to be advantageous. It appears that gentle adjustment procedures and preferable paths through parameter space can lead to otherwise hardly reachable bubble oscillation states.

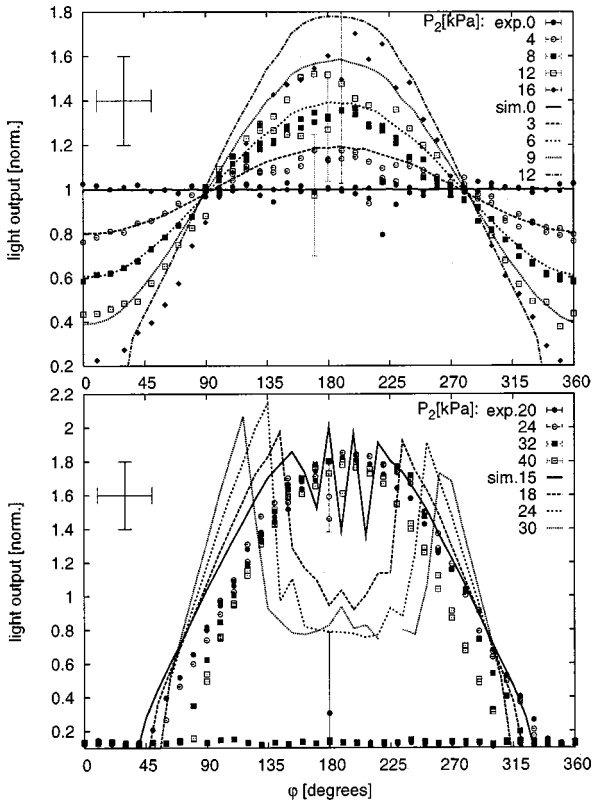


FIG. 12. Light emission of a phase scan at $P_1=140$ kPa and different P_2, normalized to the single mode SBSL emission. The variance of experimental data (symbols) is only shown when greater than 0.2. Simulated results (lines) are shown for $c_\infty/c_0=0.15\%$ using the same modification for unstable SBSL as in Fig. 3. The data are best fit by a reduction of P_2 by a factor of 3/4. From $P_2=20$ kPa up, the experimental maximum light output around $\varphi=180°$ does not increase significantly.

The calculated maximum light output at some parameter values is significantly greater than what was measured, especially at higher second harmonic excitation. According to the model and the empirical rule Eq. (6) a bubble at $\varphi=180°$ and $P_2=40$ kPa, theoretically shape unstable, should emit more than four times the light of a bubble driven with $P_1=140$ kPa alone (three times at $P_2=30$ kPa). We observed merely a doubling of the light signal. A significant reduction of the observed signal by unstable SBSL and averaging can be excluded for these data, because no typical unstable SBSL was met here (only fluctuations around a mean value, no breakdowns). Therefore, we conjecture that relation (6) might not be universally applicable.

V. SUMMARY AND DISCUSSION

An extensive investigation of SBSL driven at the fundamental mode of 25 kHz and twice that frequency has been presented. Representative parameter space sections have been scanned experimentally to measure regions of light emission and to quantify the light output and, to some extent, the maximum bubble radius. Measurements of relative phase φ and fundamental pressure amplitude P_1 at fixed second harmonic amplitude P_2 have been presented as well as a

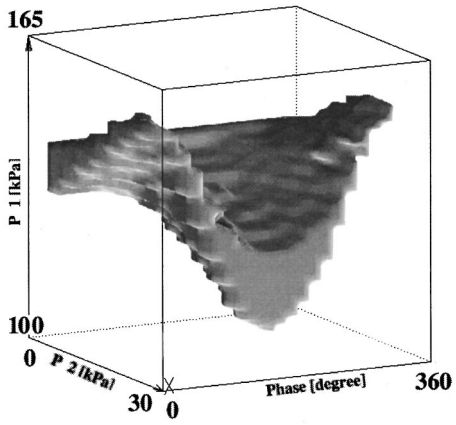


FIG. 13. Three-dimensional extension of the parameter region of stable SBSL emission: calculation for 15% air saturation and $\lambda_{max}=1$. The ''stepped'' surface is an artifact of the grid-based calculations.

survey with respect to φ and P_2 for fixed P_1. The parameter zones of SBSL have been calculated numerically by a Keller–Miksis-type bubble model taking into account shape and diffusive stability under assumption of the dissociation hypothesis.[17,18] We find mostly good agreement between experimental and theoretical SBSL regions in parameter space. Likewise, the calculated maximum radii range corresponds within 10% to the observed ones. This can be interpreted as another confirmation of the dissociation hypothesis.[47,48]

The upper stability border of the SBSL parameter region has been numerically determined in terms of the maximum absolute value of the Floquet multipliers of the $m=2$ surface mode, λ_{max}. This rather simple approach neglects several damping mechanisms that are numerically expensive to model. This shortcoming could be compensated quite reasonably by a uniform increase of the bubble destruction threshold to values of λ_{max} greater than 1. The quantity λ_{max} has also been used to describe the reduction of the averaged light output for unstable SBSL states. It has been illustrated that a monotonic parameter drift (like an increase of the gas saturation) may cause nonmonotonic emission alterations if the unstable SBSL effect is included. The reason is the parametric resonances of the surface mode. Similar changes of stability and emission have been found in our long-term experiments.

Figure 13 gives a plot of the (calculated) stability zone in three dimensions for $\lambda_{max}<1$. We note that the approximate symmetry with respect to the phase shift $\varphi=180°$ is remarkable and not evident *a priori*, because driving signal shapes are in general different for $\varphi=180°-\Delta\varphi$ and $\varphi=180°+\Delta\varphi$ with some $\Delta\varphi$. The symmetry is reflected, however, by the peak negative pressure of the driving (cf. Ref. 16). This quantity, the most negative pressure, seems to be essential for the diffusive equilibrium radius R_0, which displays a similar symmetry, at least for small contributions of the second harmonic. The bubble size in turn is the main contributing factor to the number of emitted photons (via the molecule number inside the bubble). For a larger share of the second harmonic, the symmetry becomes increasingly destroyed.

The light emission was measured in terms of maximum PMT voltage averaged over 30 SBSL flashes. In the case of pulses much shorter than the PMT impulse response time (our assumption), the signal scales with the number of emitted photons of detectable wavelength. The measurements have been compared with numerically estimated values from the bubble model and following the empirical relation (6) from Ref. 7. The experimental trends could be well reproduced, although the modeled emission appears to predict a partly stronger boost. Thus, the estimation according to relation (6) may not apply to all situations. This conclusion, however, is based on the modeled expansion ratios only, as the equilibrium bubble radius has not been monitored in the experiments.

It was possible to amplify SBSL emission relative to sinusoidal driving by adding the second harmonic. A comparison of the maximum achievable photon output between single- and double-frequency excitation showed a boost factor of up to 2.5. However, long-term measurements of this effect indicated variations of the relative gain as well as the optimal phase shift. In most cases the observed amplification factor was less than 1.5. Additionally, a large gain was seen only where pure sinusoidal driving resulted in rather low absolute emission maxima. Within a period of 16 h the observed overall maxima of both excitation methods were little different.

The investigations exhibited a relatively high sensitivity of the experiments. In particular, bubble destruction, which is based on random perturbations, and the degree of shape stability may be influenced by external noise and small parameter drifts. On the other hand, a smooth parameter change without abrupt perturbations could avoid early bubble destruction. Potentially a parameter adjustment via a smooth path may achieve preferable regimes in parameter space that are impossible to establish directly, for instance ''stability islands'' with high light emission. However, we believe that a reproducible performance of such experiments requires an excellent control of parameters and noise, which is not usually given in a standard SBSL system.

ACKNOWLEDGMENTS

We thank R. Geisler, C.-D. Ohl, O. Lindau, S. Luther, J. Segal, F. Gaitan, and G. Holt for help, discussion, and information. We also thank the mechanical and electronical workshops of the DPI.

[1] D. F. Gaitan, L. A. Crum, C. C. Church, and R. A. Roy, ''Sonoluminescence and bubble dynamics for a single, stable, cavitation bubble,'' J. Acoust. Soc. Am. **91**, 3166–3183 (1992).
[2] B. P. Barber, R. A. Hiller, R. Löfstedt, S. J. Putterman, and K. R. Weninger, ''Defining the unknowns of sonoluminescence,'' Phys. Rep. **281**, 65–143 (1997).
[3] S. Putterman, P. G. Evans, G. Vazquez, and K. Weninger, ''Is there a simple theory of sonoluminescence?'' and reply by S. Hilgenfeldt, S. Grossmann, and D. Lohse, Nature (London) **409**, 782–783 (2001).
[4] J. Holzfuss, M. Rüggeberg, and R. Mettin, ''Boosting sonoluminescence,'' Phys. Rev. Lett. **81**, 1961–1964 (1998).
[5] S. Hilgenfeldt and D. Lohse, ''Predictions for upscaling sonoluminescence,'' Phys. Rev. Lett. **82**, 1036–1039 (1999).
[6] W. C. Moss, D. B. Clarke, J. W. White, and D. A. Young, ''Sonoluminescence and the prospects for table-top micro-thermonuclear fusion,'' Phys. Lett. A **211**, 69–74 (1996).
[7] D. F. Gaitan and R. G. Holt, ''Experimental observations of bubble re-

sponse and light intensity near the threshold for single bubble sonoluminescence in an air–water system,'' Phys. Rev. E **59**, 5495–5502 (1999).
[8] K. Hargreaves and T. J. Matula, ''The radial motion of a sonoluminescence bubble driven with multiple harmonics,'' J. Acoust. Soc. Am. **107**, 1774–1776 (2000).
[9] J.-L. Thomas, Y. Forterre, and M. Fink, ''Boosting sonoluminescence with a high-intensity ultrasonic pulse focused on the bubble by an adaptive array,'' Phys. Rev. Lett. **88**, 074302 (2002).
[10] A. J. Szeri and L. G. Leal, ''The onset of chaotic oscillations and rapid growth of a spherical bubble at subcritical conditions in an incompressible liquid,'' Phys. Fluids A **3**, 551–555 (1991).
[11] R. I. Nigmatulin, V. Sh. Shagapov, N. K. Vakhitova, and R. T. Lekhi, ''A method of ultrahigh compression of gas bubbles in a liquid by the nonperiodic vibrational action of pressure of moderate magnitude,'' Dokl. Akad, Nauk. **341**, 37 (1995) [Phys. Dokl. **40**, 122–126 (1995)].
[12] M. M. Fyrillas and A. J. Szeri, ''Control of ultra- and subharmonic resonances,'' J. Nonlinear Sci. **8**, 131–159 (1998).
[13] R. I. Nigmatulin, A. A. Gubaidullin, and O. Sh. Beregova, ''Method of resonance overcompression in a bubble liquid by a moderate aperiodic action,'' Dokl. Akad. Nauk. **374**, 489–492 (2000) [Phys. Dokl. **45**, 539–542 (2000)].
[14] R. Mettin, J. Holzfuss, and W. Lauterborn, ''Optimierte Anregung von Kavitationsblasen durch anharmonische Ultraschallsignale,'' in *Fortschritte der Akustik—DAGA 95*, edited by W. Arnold and S. Hirsekorn (DEGA, Oldenburg, 1995), pp. 1147–1150.
[15] F. B. Seeley, ''Effects of higher-order modes and harmonics in single-bubble sonoluminescence,'' J. Acoust. Soc. Am. **105**, 2236–2242 (1999).
[16] J. A. Ketterling and R. E. Apfel, ''Using phase space diagrams to interpret multiple frequency drive sonoluminescence,'' J. Acoust. Soc. Am. **107**, 819–826 (2000).
[17] D. Lohse, M. P. Brenner, T. F. Dupont, S. Hilgenfeldt, and B. Johnston, ''Sonoluminescing air bubbles rectify argon,'' Phys. Rev. Lett. **78**, 1359–1362 (1997).
[18] D. Lohse and S. Hilgenfeldt, ''Inert gas accumulation in sonoluminescing bubbles,'' J. Chem. Phys. **107**, 6986–6997 (1997).
[19] I. Akhatov, R. Mettin, C. D. Ohl, U. Parlitz, and W. Lauterborn, ''Bjerknes force threshold for stable single bubble sonoluminescence,'' Phys. Rev. E **55**, 3747–3750 (1997).
[20] T. J. Matula, S. M. Cordry, R. A. Roy, and L. A. Crum, ''Bjerknes force and bubble levitation under single-bubble sonoluminescence conditions,'' J. Acoust. Soc. Am. **102**, 1522–1527 (1997).
[21] J. Staudenraus and W. Eisenmenger, ''Fiber-optic probe hydrophone for ultrasonic and shock-wave measurements in water,'' Ultrasonics **31**, 267–273 (1993).
[22] D. Krefting, ''Einzelblasensonolumineszenz bei Zweifrequenzanregung (Single-bubble sonoluminescence under two-frequency excitation),'' Diploma thesis, Göttingen (1999) (in German).
[23] U. Parlitz, V. Englisch, C. Scheffczyk, and W. Lauterborn, ''Bifurcation structure of bubble oscillators,'' J. Acoust. Soc. Am. **88**, 1061–1077 (1990).
[24] R. Löfstedt, B. P. Barber, and S. Putterman, ''Toward a hydrodynamic theory of sonoluminescence,'' Phys. Fluids A **5**, 2911–2928 (1993).
[25] A. Eller and H. G. Flynn, ''Rectified diffusion during nonlinear pulsations of cavitation bubbles,'' J. Acoust. Soc. Am. **37**, 493–503 (1965).
[26] M. M. Fyrillas and A. J. Szeri, ''Dissolution or growth of soluble spherical oscillating bubbles,'' J. Fluid Mech. **277**, 381–407 (1994).
[27] M. P. Brenner, D. Lohse, D. Oxtoby, and T. F. Dupont, ''Mechanisms for stable single bubble sonoluminescence,'' Phys. Rev. Lett. **76**, 1158–1161 (1996).
[28] S. Hilgenfeldt, D. Lohse, and M. P. Brenner, ''Phase diagrams for sonoluminescing bubbles,'' Phys. Fluids **8**, 2808–2826 (1996).
[29] I. Akhatov, N. Gumerov, C. D. Ohl, U. Parlitz, and W. Lauterborn, ''The role of surface tension in stable single-bubble sonoluminescence,'' Phys. Rev. Lett. **78**, 227–230 (1997).
[30] H. Lamb, *Hydrodynamics*, 6th ed. (Cambridge University Press, New York, 1975).
[31] M. S. Plesset, ''On the stability of fluid flows with spherical symmetry,'' J. Appl. Phys. **25**, 96–98 (1954).
[32] H. W. Strube, ''Numerische Untersuchungen zur Stabilität nichtsphärisch schwingender Blasen,'' Acustica **25**, 289–303 (1971).
[33] J. A. Ketterling and R. E. Apfel, ''Extensive experimental mapping of sonoluminescence parameter space,'' Phys. Rev. E **61**, 3832–3837 (2000).
[34] J. A. Ketterling and R. E. Apfel, ''Shape and extinction thresholds in sonoluminescence parameter space,'' J. Acoust. Soc. Am. **107**, L13–L18 (2000).
[35] A. Prosperetti, ''Viscous effects on perturbed spherical flows,'' Q. Appl. Math. **34**, 339 (1977).
[36] C. C. Wu and P. H. Roberts, ''Bubble shape instability and sonoluminescence,'' Phys. Lett. A **250**, 131–136 (1998).
[37] Y. Hao and A. Prosperetti, ''The effect of viscosity on the spherical stability of oscillating gas bubbles,'' Phys. Fluids **11**, 1309–1317 (1999).
[38] B. D. Storey, ''Shape stability of sonoluminescence bubbles: Comparison of theory to experiments,'' Phys. Rev. E **64**, 017301 (2001).
[39] M. P. Brenner, D. Lohse, and T. F. Dupont, ''Bubble shape oscillations and the onset of sonoluminescence,'' Phys. Rev. Lett. **75**, 954–957 (1995).
[40] S. J. Putterman and P. H. Roberts, Comment on ''Bubble shape oscillations and the onset of sonoluminescence,'' and Reply to the Comment by Brenner *et al.*, Phys. Rev. Lett. **80**, 3666–3669 (1998).
[41] D. F. Gaitan and R. G. Holt (private communication).
[42] B. Gompf, R. Günther, G. Nick, R. Pecha, and W. Eisenmenger, ''Resolving sonoluminescence pulse width with time-correlated single photon counting,'' Phys. Rev. Lett. **79**, 1405–1408 (1997).
[43] C. D. Ohl, ''Probing luminescence from nonspherical bubble collapse,'' Phys. Fluids **14**(8), 2700–2708 (2002).
[44] K. Yasui, ''Alternative model of single-bubble sonoluminescence,'' Phys. Rev. E **56**, 6750–6760 (1997).
[45] B. D. Storey and A. J. Szeri, ''Water vapor, sonoluminescence, and sonochemistry,'' Proc. R. Soc. London, Ser. A **456**, 1685–1709 (2000).
[46] W. C. Moss, D. A. Young, J. A. Harte, J. L. Levatin, B. F. Rozsnyai, G. B. Zimmerman, and I. H. Zimmerman, ''Computed optical emissions from a sonoluminescing bubble,'' Phys. Rev. E **59**, 2986–2992 (1999).
[47] T. J. Matula and L. A. Crum, ''Evidence for gas exchange in single-bubble sonoluminescence,'' Phys. Rev. Lett. **80**, 865–868 (1998).
[48] J. A. Ketterling and R. E. Apfel, ''Experimental validation of the dissociation hypothesis for single bubble sonoluminescence,'' Phys. Rev. Lett. **81**, 4991–4994 (1998).

The time-dependent magneto-visco-elastic behavior of a magnetostrictive fiber actuated viscoelastic polymer matrix composite

Cosmina S. Hogea
Watson School of Engineering and Applied Science, Department of Mechanical Engineering, Binghamton University, Binghamton, New York 13902

William D. Armstrong
Department of Mechanical Engineering, University of Wyoming, Laramie, Wyoming 82072

(Received 12 September 2001; revised 13 April 2002; accepted 26 July 2002)

The paper develops a one-dimensional magneto-elastic model of a magnetostrictive fiber actuated polymer matrix composite material which accounts for a strong viscoelastic response in the polymer matrix. The viscoelastic behavior of the composite polymer matrix is modeled with a three parallel Maxwell element viscoelastic model, the magnetoelastic behavior of the composite fibers is modeled with an anhysteric directional potential based domain occupation theory. Example calculations are performed to identify and explain the dynamical behavior of the composite. These calculations assume that a constant stress and the oscillating magnetic field are applied in the fiber longitudinal direction. The inclusion of matrix viscosity results in an apparent hysteresis loop in the magnetization and magnetostriction curves even though the model does not include magnetoelastic hysteresis in the fibers. The apparent hysteresis is a consequence of the interaction of the time varying fiber stress caused by matrix viscosity with a multidomain state in the fiber. The small increase in fiber longitudinal compressive stress due to matrix viscosity under increasing field inhibits the occupation of domains with magnetization orientations near the fiber longitudinal [112] direction. As a consequence, the summed longitudinal magnetization and magnetostriction is reduced as compared to the decreasing field limb. [DOI: 10.1121/1.1508790]

PACS numbers: 43.35.Mr [SLE]

I. INTRODUCTION

The purpose of this paper is to develop a magnetoelastic model of a magnetostrictive fiber actuated polymer matrix composite material, which accounts for a strong viscoelastic response in the polymer matrix. Magnetostrictive fiber actuated polymer matrix composite materials are being considered as the active element inside low cost naval sonar sources. It is necessary to include matrix viscoelastic effects in our modeling description as these materials typically operate at elevated temperatures within an encompassing solenoid heated by the large cyclic electric current responsible for the actuating field. The elevated operating temperature in turn induces a viscous response in the composite polymer matrix. We use the anhysteric directional potential based domain occupation theory[1] to describe the magnetoelastic behavior of the Terfenol-D actuating fibers, the magnetoelastic hysteresis in the composite actuating fibers is ignored in order to clearly identify the influence of the viscous effects in the polymer matrix on the composite material behavior. The composite is homogenized by assuming the simple one-dimensional stress equilibrium and uniform strain conditions appropriate for a uniaxial long fiber composite. Example calculations are performed to identify and explain the dynamical behavior of the composite.

This paper will show that the composite magnetoelastic behavior is significantly influenced by the matrix viscoelastic response. The use of a magnetostrictive fiber actuated polymer matrix composite drive element will therefore necessitate changes in all aspects of sonar sources using this type of material.

II. POLYMER MATRIX

A. Theoretical polymer matrix

We will use a one-dimensional parallel Maxwell multi-element model to describe the viscoelastic behavior of the polymer matrix. The one-dimensional Maxwell viscoelastic element consists of an elastic spring in series with a viscous damper, and it is therefore characterized by only two parameters: the elastic constant of the spring, E, and the viscosity of the damper, η. In a Maxwell element, the total stress $\sigma(t)$ is equal to the stress in the spring $\sigma^e(t)$ and equal to the stress in the viscous damper $\sigma^v(t)$ at any moment of time t. Denoting the time-dependent strain in the spring (elastic strain) by $\varepsilon^e(t)$, the time-dependent strain in the damper (viscous strain) by $\varepsilon^v(t)$, and the total time-dependent strain in the one-dimensional Maxwell element by $\varepsilon(t)$, we have the following relationships:

$$\varepsilon(t)=\varepsilon^e(t)+\varepsilon^v(t), \quad \forall t, \tag{1}$$

$$\sigma^e(t)=E\varepsilon^e(t),$$

$$\sigma^v(t)=\eta\dot{\varepsilon}^v(t), \quad \forall t, \tag{2}$$

$$\sigma(t)=\sigma^e(t)=\sigma^v(t).$$

From Eqs. (1) and (2), we get

$$\dot{\varepsilon}(t)=\frac{1}{E}\dot{\sigma}(t)+\frac{1}{\eta}\sigma(t), \quad \forall t. \tag{3}$$

The stress-relaxation law $\sigma(t)=E\varepsilon_0 e^{-(E/\eta)t}$, $\forall t\geqslant 0$ [here assuming that the total strain $\varepsilon(t)$ is instantaneously increased to ε_0 at time $t=0$] allows one to define the time-dependent relaxation modulus, $E(t)$,

$$E(t)=\frac{1}{\varepsilon_0}\sigma(t)=Ee^{-(E/\eta)t}, \quad \forall t\geqslant 0. \tag{4}$$

Boltzmann's superposition principle provides the stress as a function of time depending on the instantaneous strain ε_0, the relaxation modulus, $E(t)$, and the total strain rate history[2]

$$\sigma(t)=\varepsilon_0 E(t)+\int_0^t E(t-\tau)\dot{\varepsilon}(\tau)d\tau, \quad \forall t>0, \tag{5}$$

assuming that the strain history is continuous for all time $t>0$.

In the case of an oscillating strain history

$$\varepsilon(t)=\varepsilon_a e^{i\omega t}, \quad \forall t>0,$$

where we denoted by ε_a the oscillation amplitude and by ω the oscillation frequency, it is useful to introduce the following frequency dependent steady state material response functions:[1]

the storage modulus $E'(\omega)=\dfrac{E\omega^2\eta^2}{E^2+\omega^2\eta^2}$;

and the loss modulus $E''(\omega)=\dfrac{E^2\omega\eta}{E^2+\omega^2\eta^2}$.

Consider now N parallel one-dimensional Maxwell elements, each characterized by the two parameters E_k, η_k for $k=\overline{1,N}$. The relaxation modulus of the element ensemble is the sum of the individual element relaxation moduli

$$E(t)=\sum_{k=1}^{N} E_k e^{-(E_k/\eta_k)t}, \quad \forall t\geqslant 0. \tag{6}$$

The corresponding stress is then computed using Eq. (5), with $E(t)$ given by Eq. (6).

Similarly, for an oscillating strain history, the storage modulus of the element ensemble is the sum of the individual element storage moduli

$$E'(\omega)=\sum_{k=1}^{N}\frac{E_k\omega^2\eta_k^2}{E_k^2+\omega^2\eta_k^2} \tag{7}$$

and the loss modulus for the element ensemble is the sum of the individual element loss moduli

$$E''(\omega)=\sum_{k=1}^{N}\frac{E_k^2\omega\eta_k}{E_k^2+\omega^2\eta_k^2}. \tag{8}$$

The stress in the parallel Maxwell element model is evaluated as a function of time depending on the instantaneous strain ε_0, the relaxation modulus, $E(t)$, and the total strain history by substituting Eq. (6) into Eq. (5).

B. Experimental epoxy polymer matrix

The polymer matrix of experimental interest in the present study is a commercial version of low viscosity epoxy (Epo-Thin Buehler Ltd.). The polymer matrix was obtained by mixing epoxy resin with an amine-based hardener, curing for 18 hours at room temperature, followed by a 90 °C final curing in an oven for 3 hours. This process ensured a complete cure of the specimen.

A thin rectangular polymer test sample was machined from the cured epoxy material with a longitudinal dimension/transverse dimension ratio of 24:1 (the average thickness around 2 mm and a length of 48.5 mm). The flexural storage and flexural loss moduli of the specimen were then obtained from 5 to 100 Hz during a thermal DMA analysis performed by TA Instruments using a DMA 2980 in an oscillating frequency mode. The experimental results showed that, as expected, at 30 °C and 50 °C the epoxy polymer matrix exhibits little viscous response, with a Young modulus of approximately 2.7 GPa at 30 °C and 2.3 GPa at 50 °C. At 70 °C, the behavior of the polymer matrix very clearly becomes viscoelastic and we will therefore concentrate on the 70 °C temperature case for the remainder of this paper.

Table I shows the viscoelastic parameters of our three parallel Maxwell elements model for the polymer matrix. The fit of a three parallel Maxwell element model through the existing experimental points has been done by minimizing the function:[3]

$$f(E_1,\eta_1,E_2,\eta_2,E_3,\eta_3)=\sum_{j=1}^{P}\left[\frac{(E'(\omega_j)-E'_{\text{exp}}(\omega_j))^2}{E'_{\text{exp}}(\omega_j)^2}+\frac{(E''(\omega_j)-E''_{\text{exp}}(\omega_j))^2}{E''_{\text{exp}}(\omega_j)^2}\right],$$

where P stands for the total number of experimental points taken into account, ω_j is the jth experimental frequency, $E'_{\text{exp}}(\omega_j)$, $E''_{\text{exp}}(\omega_j)$ are the corresponding experimental (flexural) storage and loss moduli, respectively, and $E'(\omega_j)$, $E''(\omega_j)$ are the theoretical storage and loss moduli for a three parallel Maxwell elements as given by Eqs. (7) and (8). We choose the numerical values of the model parameter set E_1, η_1, E_2, η_2, E_3, η_3 that best fit the experimental existing points for *both* the (flexural) storage modulus and the (flexural) loss modulus for frequencies higher than 5 Hz. We notice that the three obtained viscosity parameters corresponding to our three parallel Maxwell elements have numerical values significantly different from one to another; thus, the first element has a viscosity one order of magnitude larger than the second one, which in turn is one order of magnitude larger than the third one. Therefore, the frequency dependence of the polymer behavior is properly described at different frequencies by the three Maxwell elements listed in Table I. At lower frequencies dominated by the element with the highest viscosity (element 1 in Table I), and at higher frequencies dominated by the element with the lowest viscosity (element 3 in Table I). None of the three elements can be removed without significantly moving off the experimental curves; three parallel Maxwell elements was the mini-

TABLE I. The viscoelastic parameters of our three parallel Maxwell elements model for the polymer matrix.

$E_1 = 973.647$ MPa	$\eta_1 = 677.682$ MPa$\times$s
$E_2 = 319.679$ MPa	$\eta_2 = 24.2998$ MPa$\times$s
$E_3 = 370.954$ MPa	$\eta_3 = 4.852$ MPa$\times$s

TABLE II. The material properties of Terfenol-D (Ref. 4).

M_S, saturation magnetization	0.8×10^6 A/m
K_1, magnetocrystalline anisotropy	-0.06×10^6 J/m^3
λ_{111}	1640×10^{-6} mm/mm
λ_{100}	90×10^{-6} mm/mm

mum number that allowed us a reasonable fit of the model with the experimental result.

Figure 1 shows a comparison between experiment and our three parallel Maxwell element model for the storage modulus, and loss modulus, respectively, at the fixed temperature of 70 °C. Although plotted in Fig. 1 for the sake of better illustration of the theoretical concepts, we will not use our model for frequencies lower than 5 Hz. However, we could apply the same procedure to investigate the possibility of fitting a N parallel Maxwell element model through additional low frequency experimental points.

III. TERFENOL-D MAGNETOSTRICTIVE FIBERS

We will describe the complex magnetoelastic behavior of the magnetostrictive fibers with the directional potential based domain occupation model of Armstrong.[1] This model assigns a directional potential to each of the eight ⟨111⟩ type crystallographic directions. The directional potential is the sum of applied magnetic field, magnetocrystalline and magnetoelastic directional potentials which we may evaluate provided we have sufficient knowledge of the orientation of the crystal, the intensity and direction of the magnetic field, the stress state, and the magnetic properties of our material. Table II gives the material properties of Terfenol-D.[4] The four properties in the table are hard values, i.e., they were measured unambiguously and not the result of a fitting procedure.

A very large number of individual domain wall translation events combine to produce each measurable domain transformation in a macroscopic sized sample of ferromagnetic material. Each individual domain wall may be expected to suffer some level of translation inhibition, however the severity of translation inhibition will spatially vary. Severe domain wall translation inhibition may be expected to be far less prevalent than minor domain wall translation inhibition, this is because the annihilation of the large defects responsible for the severe changes in local preferred moment orientation is energetically preferred during high temperature processing of the magnetostrictive material. We therefore assume that the probability of the occupation of magnetization potential follows a parameter selected, inverse exponential distribution familiar from the study of statistical thermodynamics. The probability, P, that magnetization occupies a spherical polar direction $[\theta_m, \varphi_m]$ is therefore strongly dependant on the total directional magnetization potential in that direction and on the values of the total directional magnetization potential at the other ⟨111⟩ type directions,

$$P_{(\theta_m,\varphi_m)} = K^* \operatorname{Exp}[-(E_{\text{field}} + E_{\text{an}} + E_\sigma)/\omega] \operatorname{Sin}[\theta] d\theta\, d\varphi, \tag{9}$$

where K is a normalization factor which must be evaluated at each applied field and stress condition, and ω is the directional potential distribution parameter. A reduction in ω narrows the magnetization potential range over which a given change in P occurs, thus sharpening the magnetization directional distribution.

The net magnetization, M, in the direction $[\theta_M, \varphi_M]$ is obtained by summing probability weighted magnetization vector projections from each of the eight ⟨111⟩ type domain orientations.

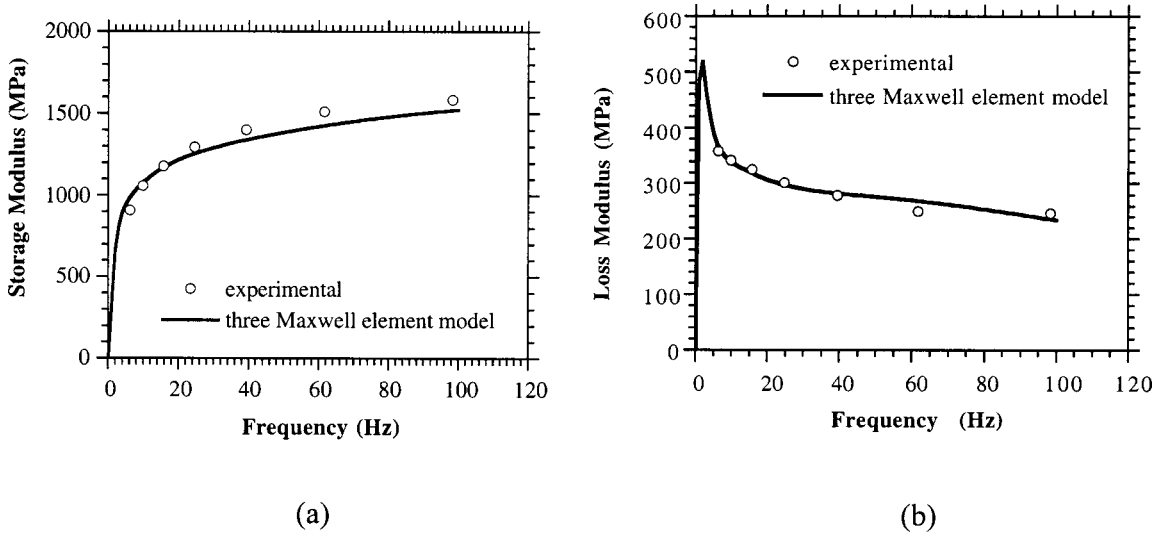


FIG. 1. Comparison between experiment and our three parallel Maxwell element model for (a) the storage modulus and (b) the loss modulus, respectively, at the fixed temperature of 70 °C.

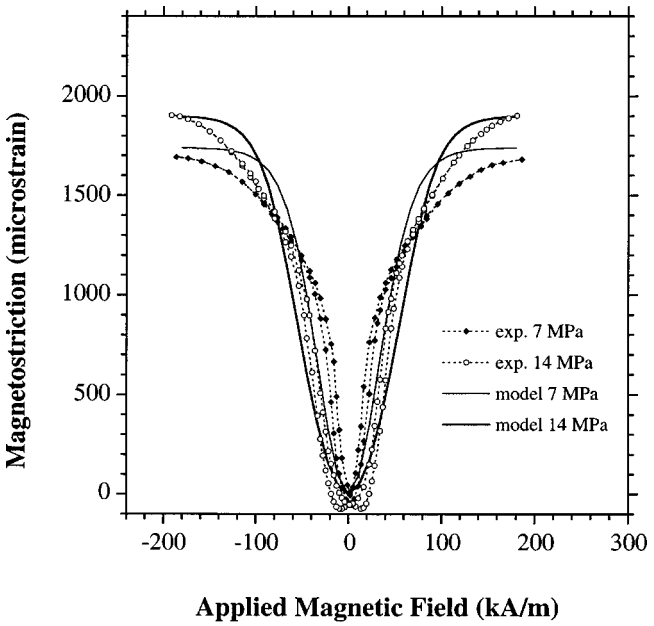


FIG. 2. Comparison between experimental and model magnetostriction in a solid rod of Terfenol-D. The specimen longitudinal axis, constant uniaxial applied compressive stress, and applied magnetic field were each parallel to the [112] crystallographic direction.

$$M = M_s \sum_{i=1}^{8} P_i \left\{ \begin{array}{c} \mathrm{Sin}[\theta_m]\mathrm{Cos}[\varphi_m]\mathrm{Sin}[\theta_M]\mathrm{Cos}[\varphi_M] \\ +\mathrm{Sin}[\theta_m]\mathrm{Sin}[\varphi_m]\mathrm{Sin}[\theta_M]\mathrm{Sin}[\varphi_M] \\ +\mathrm{Cos}[\theta_m]\mathrm{Cos}[\theta_M] \end{array} \right\}. \quad (10)$$

In a similar manner, the net magnetostriction tensor, λ, is obtained by summing probability weighted transformed magnetostriction tensor contributions from each of the eight $\langle 111 \rangle$ type domain orientations.

$$\lambda = \left\{ \begin{array}{c} \lambda_{11} \\ \lambda_{22} \\ \lambda_{33} \\ \lambda_{12} \\ \lambda_{23} \\ \lambda_{31} \end{array} \right\} = \sum_{i=1}^{8} P_i \left\{ \begin{array}{c} \lambda_{100}(\mathrm{Sin}[\theta_m]*\mathrm{Cos}[\varphi_m])^2 - \frac{1}{3} \\ \lambda_{100}(\mathrm{Sin}[\theta_m]*\mathrm{Sin}[\varphi_m])^2 - \frac{1}{3} \\ \lambda_{100}(\mathrm{Cos}[\varphi_m])^2 - \frac{1}{3} \\ \lambda_{111}(\mathrm{Sin}[\theta_m]*\mathrm{Cos}[\varphi_m]*\mathrm{Sin}[\theta_m]*\mathrm{Sin}[\varphi_m]) \\ \lambda_{111}(\mathrm{Sin}[\theta_m]*\mathrm{Sin}[\varphi_m]*\mathrm{Cos}[\theta_m]) \\ \lambda_{111}(\mathrm{Sin}[\theta_m]*\mathrm{Cos}[\varphi_m]*\mathrm{Cos}[\theta_m]) \end{array} \right\}. \quad (11)$$

In the present paper, the total magnetization potential will be calculated as a function of applied magnetic field in each of the eight $\langle 111 \rangle$ type directions for a magnetic field process corresponding to an applied oscillating (sinusoidal) magnetic field parallel to the [112] crystallographic direction. The calculations include a constant uniaxial compressive stress applied parallel to the [112] direction. Figure 2 shows a comparison between measured and calculated magnetostriction in a solid rod of Terfenol-D. The specimen longitudinal axis, constant uniaxial applied compressive stress, and applied magnetic field were each parallel to the [112] crystallographic direction. The figure shows that the model provides a useful level of agreement to experiment. The present modeling description of the Terfenol fiber is anhysteric. It is possible to include magnetoelastic hysteresis effects, however we choose not to do so in the present paper in order to most clearly present the consequences of including a viscous response in the polymer matrix.

IV. ONE-DIMENSIONAL COMPOSITE MODEL

This section is concerned with the construction of a one-dimensional model of long magnetostrictive fiber actuated viscoelastic polymer matrix composites given the individual fiber and matrix models of the preceding sections. The model we propose is noninertial, in the sense that the mass of the composite specimen does not result in a dynamic stress component. Therefore, the present model is restricted to relatively short sections and/or relatively low frequencies.

Generally, for a two-phase composite, consisting of a matrix reinforced by long uniaxial fibers, the overall longitudinal average stress is given by

$$\sigma = c_f \sigma_f + c_m \sigma_m, \quad (12)$$

where σ_f, σ_m are the longitudinal stresses in the fibers and matrix, respectively, and c_f, c_m are the volume fractions of the fibers and matrix, respectively, $c_f + c_m = 1$.

Due to the long-fiber geometry of the composite, we make the assumption that the overall strain in the composite is supported by the fibers. Denoting by ε_f, ε_m the strains in the fibers and matrix, respectively, and by ε the overall average strain in the composite, we then have

$$\varepsilon = \varepsilon_f = \varepsilon_m. \quad (13)$$

The strain in the Terfenol-D fibers ε_f is actually the Terfenol-D fiber magnetostriction, properly normalized to ensure a zero overall average strain within the composite at the initial moment of time $t = 0$,

$$\varepsilon_f = \lambda - \lambda_0, \quad (14)$$

where λ denotes the Terfenol-D fibers magnetostriction (as described in Sec. II) and λ_0 is the magnetostriction at time $t = 0$: $\lambda_0 = \lambda(t = 0)$. We also assume that at the initial moment of time $t = 0$ the composite matrix is completely relaxed, so $\sigma_m(t = 0) = 0$.

In solving the composite model the strain in the matrix ε_m must be computed at each moment of time between the initial moment of time 0 to the present specified moment of time. Each of the past matrix strain states influences the present stress in the fiber through Boltzmann's superposition. Once the stress in the fibers $\sigma_f(t_j)$ and the applied magnetic field $H(t_j)$ at time t_j have been determined we may calculate the magnetization in the fibers, $M_{\mathrm{fiber}}(t_j)$, via Eq. (10) and the magnetostriction in the fibers via Eq. (11). The magnetization in the composite $M_{\mathrm{composite}}(t_j)$ at time t_j is then directly proportional to the fibers volume fraction: $M_{\mathrm{composite}}(t_j) = c_f M_{\mathrm{fiber}}(t_j)$. We note that due to the strongly nonlinear character of both the matrix and the fibers, the

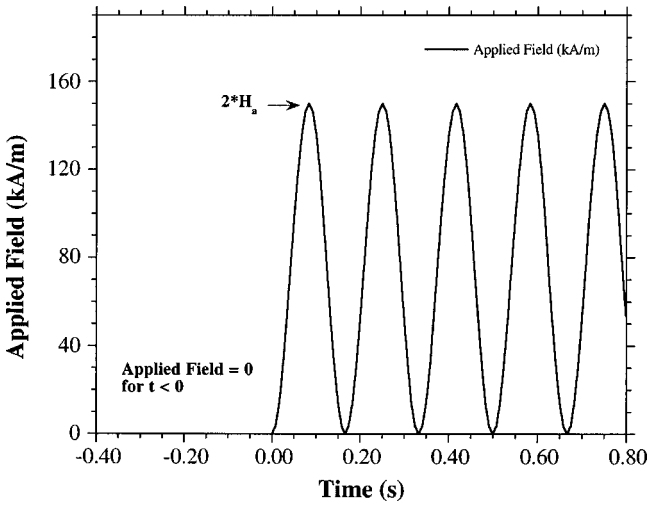


FIG. 3. Time development of the oscillating applied magnetic field considered in our computations, for ω equal to 6 Hz, and H_a equal to 75 kA/m. The applied field intensity is zero for any time $t<0$.

iteration convergence at each discrete moment of time is slow, even for very small time increments Δt.

V. RESULTS AND DISCUSSION

The present model predicts the fiber stress, the matrix stress, the magnetization and magnetostriction of a continuous Terfenol-D long, uniaxial fiber composite material for general orientation and magnitude of the applied stress and magnetic field; however, in this paper we will focus on the case where a constant compressive stress and an oscillating magnetic field are applied parallel to the fiber direction. Additionally, the fiber longitudinal direction will be in the [112] crystallographic direction. Finally, any model predictions of composite magnetization and magnetostriction will be in the fiber direction.

Table I in Sec. II and Table II in Sec. III show the material properties of the polymer matrix and the fibers used for the present calculations. The energy distribution parameter ω is the same considered for the calculations in Sec. III.

The computational assumptions are a constant longitudinal compressive stress and an oscillating magnetic field with a maximum intensity of $2H_a$ at a frequency ω

$$H(t)=H_a \sin\left(\omega t-\frac{\pi}{2}\right)+H_a, \quad \forall t\geqslant 0,$$
$$H(t)=0, \quad \forall t<0. \tag{15}$$

Figure 3 shows the time development of such an oscillating field, for ω equal to 6 Hz, and H_a equal to 75 kA/m. The applied field intensity is zero for any time $t<0$. The magnetic field strength increases smoothly from zero at time $t=0$, and is of the form of a series of positive haversine pulses with a minimum field intensity of zero.

The present calculations further specify that at the initial moment of time $t=0$ the composite matrix is completely relaxed, so $\sigma_m(t=0)=0$, as would be the case if one were to first apply the constant compressive load then wait until the composite matrix has completely relaxed, and as a consequence the composite fibers sustain the entire applied load, before applying the oscillating magnetic field.

We will present a number of example calculations in order to highlight the dynamic behavior of the present model. We will apply the present model to the analysis of a composite with a fiber volume fraction of 30%, subjected to an constant compressive longitudinal stress of 3 MPa, and a oscillating magnetic field of $H_a=75$ kA/m (maximum intensity equal to 150 kA/m) with a frequency of 6 Hz. Results from the model are plotted to clarify the effects of matrix viscosity on the time development of internal stress, magnetization, and magnetostriction. We will then present comparative calculations for conditions of lower magnetic field amplitude ($H_a=30$ kA/m) and higher magnetic field frequency (40 Hz) while keeping other variables constant.

Figure 4 shows the calculated behavior of a continuous long fiber composite specimen with a fiber volume fraction of 30%, subjected to a constant compressive load of 3 MPa in the composite longitudinal direction and to an oscillating magnetic field at a frequency $\omega=6$ Hz, with $H_a=75$ kA/m. Figure 4(a) shows the longitudinal stress in the composite fiber and matrix versus time. The stress in the fibers remains compressive, while the longitudinal stress in the matrix oscillates between compression and tension. Both the matrix and fiber cyclic stress histories approach a steady state condition after the fourth cycle. The figure further shows that the viscous elongation of the polymer matrix results in a compressive trend in the polymer matrix stress and, as a consequence, partial unloading of the fibers. Figure 4(b) shows the calculated fiber longitudinal stress versus applied field. The fiber stress trajectory is rounded at points of field reversal due to the changing stress, the figure further shows the fiber are partially unloaded with increasing cycle number. Figure 4(c) shows the longitudinal magnetization in the composite versus applied magnetic field. Figure 4(d) shows the longitudinal magnetostriction in the composite versus applied magnetic field. The figures show the development of hysteresis loops in the magnetization and magnetostriction curves. The zero field and high field magnetization and magnetostriction endpoints are stable because at both of these terminal points the Terfenol-D is in very nearly a single domain state. The increasing and decreasing limbs of the magnetization and magnetostriction loops are offset at middle levels of applied field (approximately 50 kA/m). This offset is a consequence of the interaction of the time varying fiber stress caused by the matrix viscosity with a multidomain state in the fiber. The small increase in fiber longitudinal compressive stress under increasing field inhibits the occupation of domains with magnetization orientations near the fiber longitudinal [112] direction. As a consequence, the summed longitudinal magnetization and magnetostriction is reduced as compared to the decreasing field limb. Furthermore, the fiber unloading trend mentioned in our discussion of Figs. 4(a) and (b) results in a slight magnetic softening by the same mechanism.

The level of hysteresis in the composite model is roughly similar to the level of nearly static hysteresis in the homogeneous Terfenol-D rod of Fig. 2. We might therefore expect that a real composite would exhibit roughly twice the

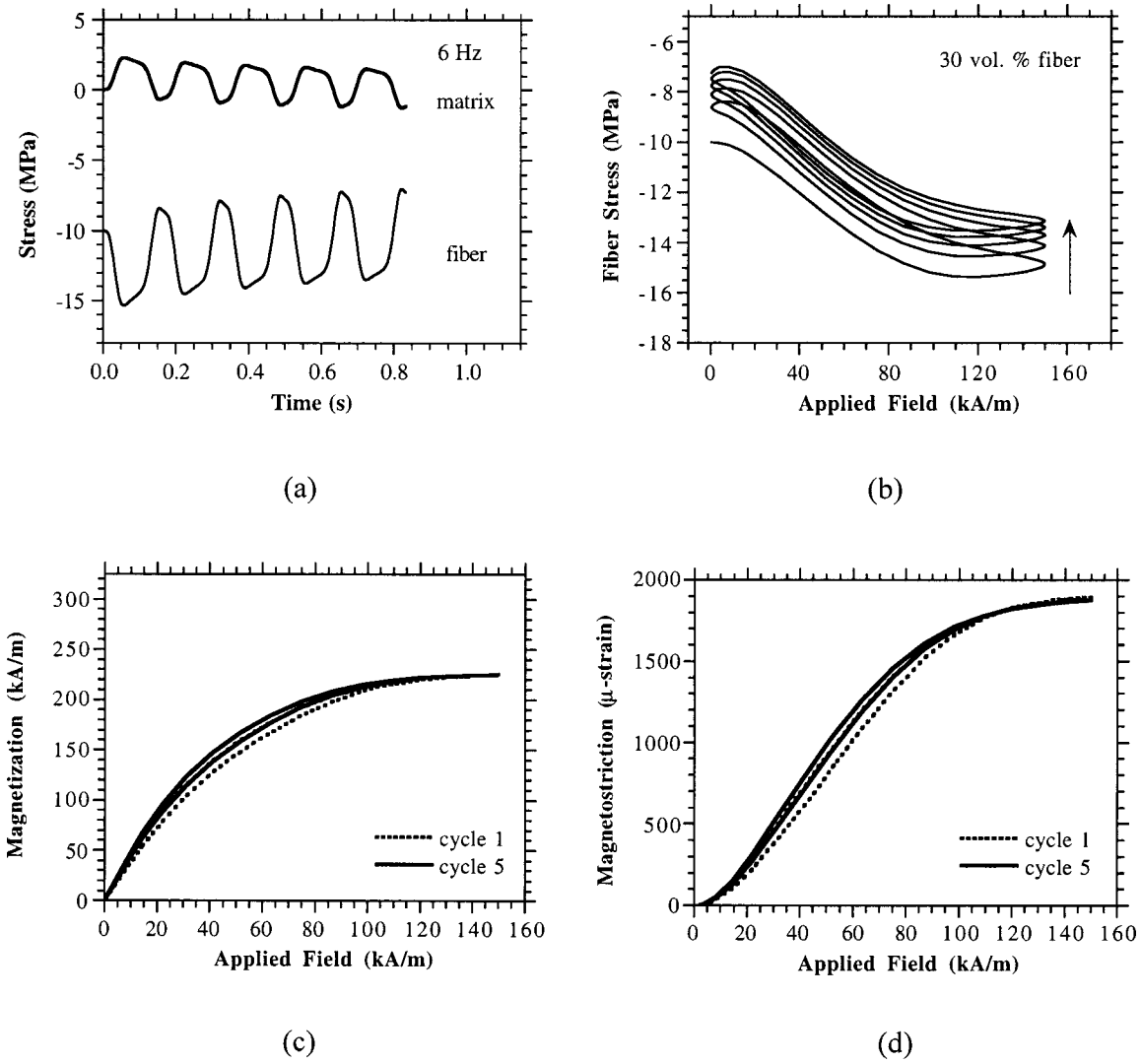


FIG. 4. The calculated behavior of a continuous long fiber composite specimen with a fiber volume fraction of 30%, subjected to a constant compressive load of 3 MPa in the composite longitudinal direction and to an oscillating magnetic field at a frequency $\omega = 6$ Hz, with $H_a = 75$ kA/m: (a) The longitudinal stress in the composite fiber and matrix versus time; (b) the calculated fiber longitudinal stress versus applied field, showing partial unloading with the increasing of the cycles number; (c) the longitudinal magnetization in the composite versus applied magnetic field, showing a hysteresis effect; (d) the longitudinal magnetostriction in the composite versus applied magnetic field, also showing a hysteresis effect.

hysteresis predicted by our model. However, Terfenol-D exhibits a frequency dependant hysteresis, one which until now has not been modeled accurately. It is therefore presently impossible to mathematically predict how the hysteresis due to matrix viscosity and the hysteresis due to domain pinning in the Terfenol-D fibers will interact to determine the frequency dependant hysteresis properties of real composite materials. Clearly more study is needed to address this important point.

Figure 5 shows model results calculated for parameters nearly identical to that of Fig. 4, the only difference being the combined field amplitude mean field value H_a has been reduced to 30 kA/m. Figures 5(a) and (b) show that the reduction of the applied field amplitude, H_a, has resulted in lower cyclic variations in matrix and fiber longitudinal stresses and a significantly faster approach to a cyclic steady state. The amount of fiber unloading is also significantly less. Figure 5(c) shows the longitudinal magnetization in the composite versus applied magnetic field. Figure 5(d) shows the longitudinal magnetostriction in the composite versus applied magnetic field. The figures again show the development of hysteresis loops in the magnetization and magnetostriction curves due to the previously discussed interaction of the time varying fiber stress caused by the matrix viscosity with a multidomain state in the fiber. However in this case the high field magnetization and magnetostriction endpoints are no longer stable. The drift of the high field magnetization and magnetostriction endpoints is a consequence of the slight fiber unloading due to the matrix polymer viscosity.

Figure 6 shows model results calculated for parameters nearly identical to that of Fig. 4, the only difference now being the applied field frequency, ω, has been increased to 40 Hz. Figures 6(a) and (b) show that the increase in applied field frequency has significantly reduced amount of drift in composite internal stress. Figure 6(c) shows the longitudinal magnetization in the composite versus applied magnetic field. Figure 6(d) shows the longitudinal magnetostriction in the composite versus applied magnetic field. The figures

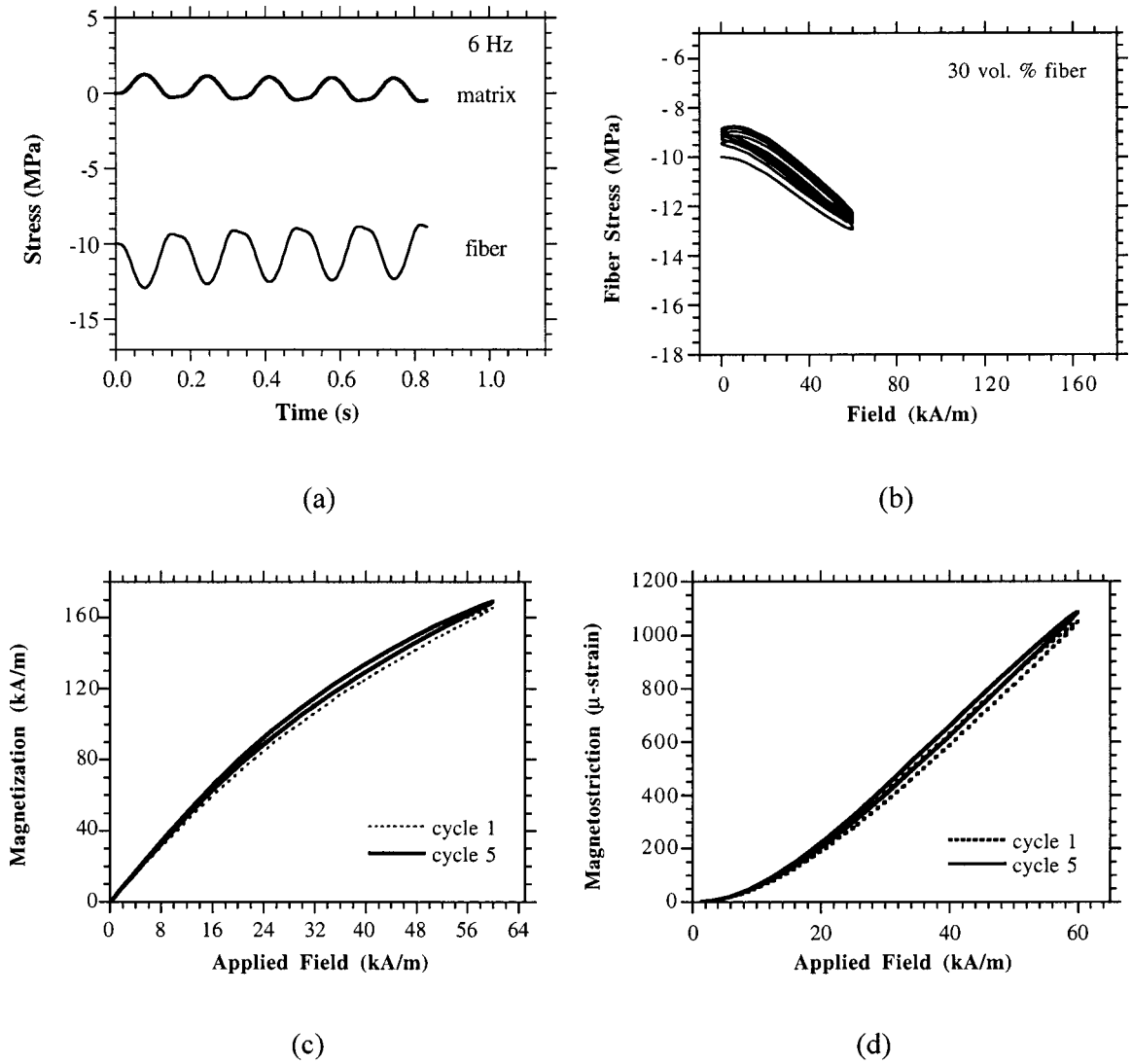


FIG. 5. Model results calculated for parameters nearly identical to that of Fig. 4, the only difference being the combined field amplitude mean field value H_a has been reduced to 30 kA/m: (a) and (b) The reduction of the applied field amplitude, H_a, has resulted in lower cyclic variations in matrix and fiber longitudinal stresses and a significantly faster approach to a cyclic steady state. The amount of fiber unloading is also significantly less; (c) the longitudinal magnetization in the composite versus applied magnetic field; (d) the longitudinal magnetostriction in the composite versus applied magnetic field.

again show the development of hysteresis loops in the magnetization and magnetostriction curves due to the previously discussed interaction of the time varying fiber stress caused by the matrix viscosity with a multidomain state in the fiber. However, in this case offset between increasing and decreasing applied field limbs is far less since the polymer matrix exhibits a significantly less viscous response at higher frequencies.

Additional calculations were performed for composites with increased fiber volume fraction conditions. Increasing the fiber volume fraction significantly reduced the cyclic variation of fiber longitudinal stress, while the cyclic variation of matrix longitudinal stress remains nearly the same. This is due to the matrix stress being primarily dependant on the composite strain, while the fiber stress is that necessary to equilibrate the matrix and constant applied stresses. The model again exhibits the development of hysteresis loops in the magnetization and magnetostriction curves due to the previously discussed interaction of the time varying fiber stress caused by the matrix viscosity with a multidomain state in the fiber. However in this case offset between increasing and decreasing applied field limbs is far less since less cyclic variation in fiber stress occurs, resulting in less variation in domain configuration. Finally, the higher fiber volume fraction composite is predicted to exhibit an elevated magnetization curve, while exhibiting a nearly similar magnetostriction curve. This is a consequence of the volumetric nature of total magnetization.

VI. SUMMARY AND CONCLUSIONS

The purpose of this paper was to develop a magnetoelastic model of a magnetostrictive fiber actuated polymer matrix composite material, which accounts for a strong viscoelastic response in the polymer matrix. It is necessary to include matrix viscoelastic effects in our modeling description as these materials typically operate at elevated temperatures within an encompassing solenoid heated by the large cyclic

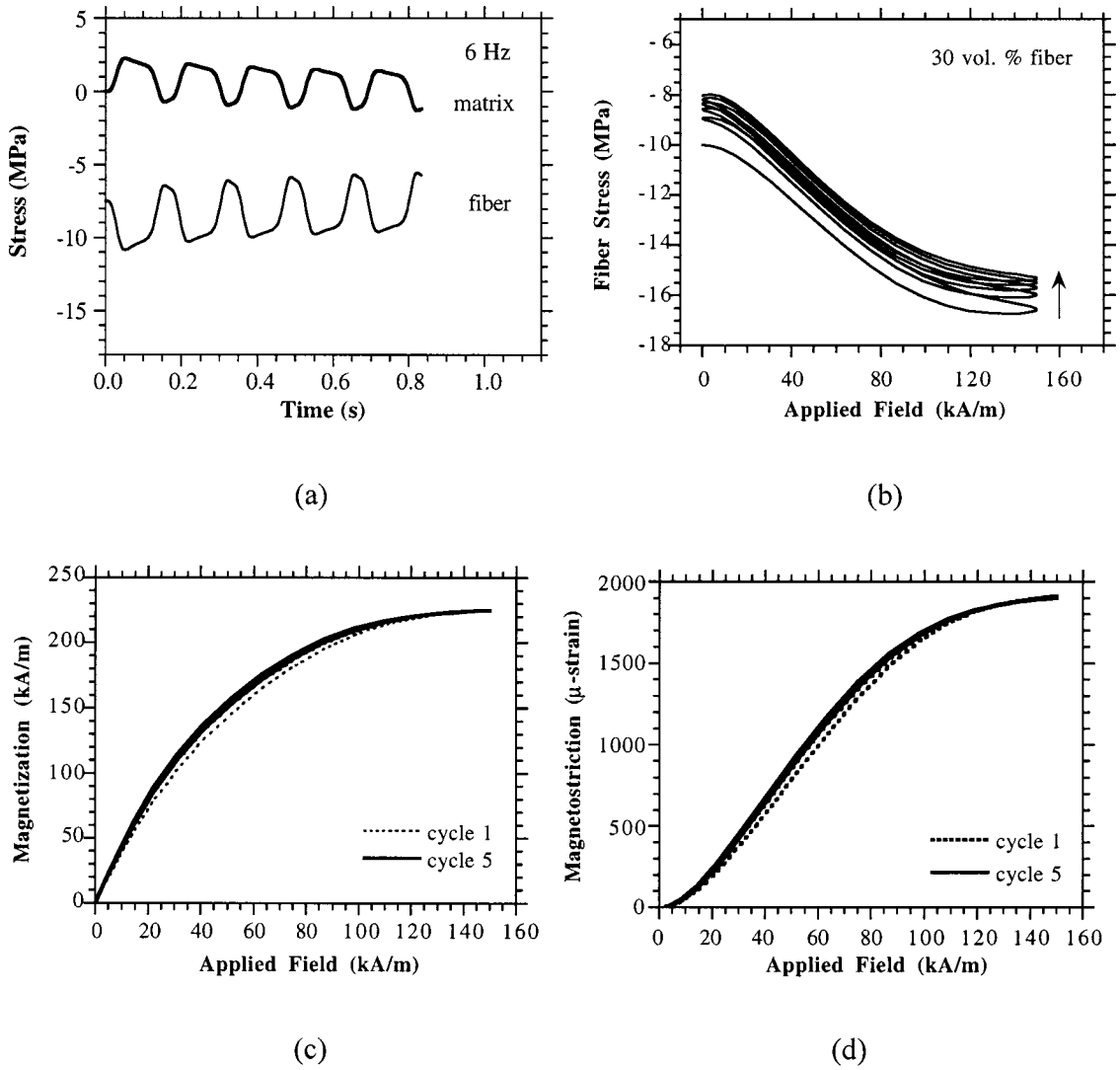


FIG. 6. Model results calculated for parameters nearly identical to that of Fig. 4, the only difference now being the applied field frequency, ω, has been increased to 40 Hz: (a) and (b) The increase in applied field frequency has significantly reduced amount of drift in composite internal stress; (c) the longitudinal magnetization in the composite versus applied magnetic field; (d) the longitudinal magnetostriction in the composite versus applied magnetic field. Offset between increasing and decreasing applied field limbs is far less since the polymer matrix exhibits a significantly less viscous response at higher frequencies.

electric current responsible for the actuating field. The elevated operating temperature in turn induces a viscous response in the composite polymer matrix.

We developed our composite model by first selecting a three parallel Maxwell element viscoelastic model to describe our polymer matrix, this model was shown to exhibit good agreement to experimental storage and loss moduli measurements on a representative commercial epoxy material over the range frequencies we were interested. We selected the anhysteric directional potential based domain occupation theory to describe the magnetoelastic behavior of the Terfenol-D actuating fibers. The absence of magnetoelastic hysteresis in the composite actuating fibers clarified the identification of viscous effects in the polymer matrix. We then homogenized the composite by assuming the simple one dimensional stress equilibrium and uniform strain conditions appropriate for a uniaxial long fiber composite.

Example calculations were performed to identify and explain the dynamical behavior of the model. Four different specific calculations were then performed, the base calculation was for a 30% fiber volume fraction composite subjected to a 6 Hz oscillating magnetic field with a combined mean field and sine amplitude value of 75 kA/m. We then considered reduced field amplitude, increased field frequency and increased fiber volume fraction variations from the base conditions. These results support the following conclusions.

(1) Under all of the conditions considered in this paper the stress in the fibers remains compressive, while the longitudinal stress in the matrix oscillates between compression and tension. Both the matrix and fiber cyclic stress histories approach a steady state condition after the fourth cycle. The viscous elongation of the polymer matrix results in a compressive trend in the polymer matrix stress and, as a consequence, partial unloading of the fibers. The fiber stress trajectory is rounded at points of field reversal due to changing stress conditions.

(2) Under the base condition the zero field and high field magnetization and magnetostriction endpoints are stable be-

tween cycles because at both of these terminal points the Terfenol-D is in essentially a single domain state. We observe that the increasing and decreasing limbs of the magnetization and magnetostriction loops are offset at middle levels of applied field (approximately 50 kA/m). This offset is a consequence of the interaction of the time varying fiber stress caused by matrix viscosity with a multidomain state in the fiber. The small increase in fiber longitudinal compressive stress under increasing field inhibits the occupation of domains with magnetization orientations near the fiber longitudinal [112] direction. As a consequence, the summed longitudinal magnetization and magnetostriction is reduced as compared to the decreasing field limb. This results in an observable hysteresis loop in the magnetization and magnetostriction curves even though the model does not include magnetoelastic hysteresis in the fibers.

(3) The increased field frequency conditions significantly reduced the amount of drift in composite internal stress. The model again exhibits the development of hysteresis loops in the magnetization and magnetostriction curves due to the interaction of the time varying fiber stress caused by the matrix viscosity with a multidomain state in the fiber. However in this case offset between increasing and decreasing applied field limbs is far less since the polymer matrix exhibits a significantly less viscous response at higher frequencies.

ACKNOWLEDGMENTS

The support of the Department of the Navy, Office of Naval Research under Contract No. N00014-99-1-0276 and the direction of Program Officer Jan Lindberg is gratefully acknowledged.

[1] W. D. Armstrong, "A directional magnetization potential based model of magneto-elastic hysteresis," J. Appl. Phys. **91**, 2202–2022 (2002).
[2] A. S. Wineman and K. R. Rajagopal, *Mechanical Response of Polymers* (Cambridge University Press, Cambridge, UK, 2000).
[3] Faith A. Morrison, *Understanding Rheology* (Oxford University Press, Oxford, 2001).
[4] D. C. Jiles and J. B. Thoelke, "Theoretical modeling of the effects of anisotropy and stress on the magnetization and magnetostriction of $Tb_{0.3}Dy_{0.7}Fe_2$," J. Magn. Magn. Mater. **134**, 143–160 (1994).

Micromechanical modeling of viscoelastic voided composites in the low-frequency approximation

Michael R. Haberman,[a] Yves H. Berthelot,[b] and J. Jarzynski
Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

Mohammed Cherkaoui
Laboratoire de Physique et Mécanique des Matériaux, Institut Supérieur de Génie Mécanique et Productique, URA CNRS 1215, Université de Metz, 57045 Metz, France

(Received 26 February 2002; revised 18 July 2002; accepted 1 August 2002)

The self-consistent model of Cherkaoui *et al.* [J. Eng. Mater. Technol. **116**, 274–278 (1994)] is used to compute the effective material moduli of a viscoelastic material containing coated spherical inclusions. Losses are taken into account by introducing the frequency-dependent, complex shear modulus of the viscoelastic matrix. Mode conversion appears through the localization tensors that govern the micromechanical behavior near the inclusions. The results are compared with the scattering model and the data of Baird *et al.* [J. Acoust. Soc. Am. **105**, 1527–1538 (1999)]. The two models are in good agreement. the advantage of the self-consistent model is that it is applicable to the case of nonspherical inclusions embedded in anisotropic materials. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1509424]

PACS numbers: 43.35.Mr, 43.20.Fn, 43.40.Fz [ANN]

I. INTRODUCTION AND BACKGROUND

A common way to attenuate the transmission of sound from water to a submerged elastic structure is to add a layer of viscoelastic material containing microinclusions, which are typically microvoids separated from the viscoelastic matrix by a thin glass shell. The presence of inclusions greatly increases losses by enhancing shear deformation in the vicinity of the compliant microscatterer. The shear modulus of the viscoelastic material is complex, leading to significant dissipation within the material despite the fact that the wavelength is considerably larger than the size of the inclusions. There is a vast body of literature dealing with the problem of low-frequency scattering of elastic waves in materials containing inclusions. Chaban[1] proposed a model in which the field from many noninteracting scatterers is assumed to be equivalent to the field created by a single scatterer whose volume includes all of the individual scatterers. It is known as a self-consistent model because it assumes that the field at any inclusion inside the volume is the field scattered by other inclusions in the volume minus its own field. An alternative approach is the scattering technique of Kuster and Toksöz.[2] Because it is a nonresonant scattering technique, it does not predict any frequency dependence of the effective properties of the composite material, a dependence that is observed in experiments. By including resonant scattering, Gaunaurd and Überall[3] (GU) capture the frequency dependence of the effective material properties. These models are based on single-scattering theory in the long-wavelength regime and they are therefore limited to low void fractions, i.e., dilute, noninteracting spherical inclusions. Waterman and Truell[4] (WT) consider first-order effects in multiple scattering, and Varadan *et al.*[5] develop a more complete multiple-scattering approach that is computationally intensive. Scharnhorst[6] and Kerr[7] provide an in-depth discussion of the GU and WT models and their limitations. We also note Norris's approach[8] that treats both shear and compressional incidence in a unified manner, and where losses are calculated by virtue of the optical theorem. These models are all restricted to composite materials containing single-phase inclusions, and therefore they do not take into account the thin coating on each inclusion. Recently, Baird, Kerr, and Townend[9] extended the GU model to include the effect of the thin coating. They obtained explicit expressions (subsequently referred to as the BKT model) for the effective bulk and shear moduli of composite materials with coated microinclusions, and obtained good agreement between predicted and measured transmission loss through a slab of material suspended in water. They also include the effect of a distribution of spheres of various sizes on the observed effective moduli.

The idea of using effective medium theory to model the *static* properties of anisotropic composites was introduced by Hornby *et al.*[10] in their study of shales, which are complex porous media composed of interpenetrating fluid and clay mineral phases occurring in most sedimentary basins. Their theory is based on a combination of the self-consistent (SC) model for anisotropic medium and the differential effective medium approximation. Independently, Cherkaoui *et al.*[11] proposed a self-consistent micromechanical model that predicts the effective moduli of an anisotropic viscoelastic material filled with ellipsoidal, coated, microinclusions. The originality of Cherkaoui's model is that it is valid for a medium containing *coated* inclusions. It is a purely static model and therefore, in its current form, it cannot be used to predict acoustical attenuation of elastic waves. The main idea of the present paper is to extend Cherkaoui's model to predict losses under dynamic loading. Cherkaoui's model belongs to

[a]Current address: Beltone Electronics Corporation, 4201 West Victoria St., Chicago, IL 60646.
[b]Author to whom correspondence should be addressed. Electronic mail: yves.berthelot@me.gatech.edu

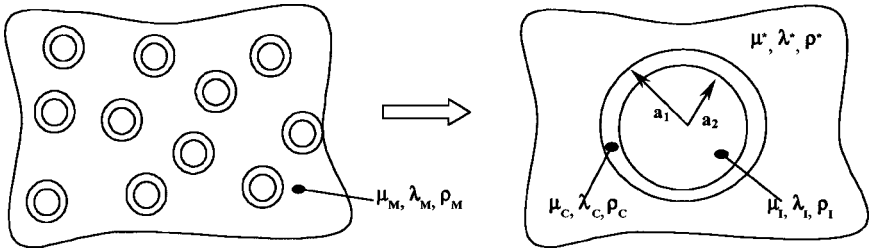


FIG. 1. The self-consistent approach: the matrix containing multiple coated inclusions is equivalent to a homogeneous effective matrix containing a single coated inclusion. λ and μ are the Lamé constants, and ρ is the density.

the class of static, self-consistent, micromechanical models that initiated with the seminal works of Kerner,[12] Hill,[13] and Budianski.[14] See, for instance, Christensen's classic review article.[15] Christensen and Lo[16] used the generalized self-consistent approach to model two-phase (uncoated) isotropic composites with a three-phase representation. Berryman[17] used the self-consistent approach to model the effective properties of two-phase composites containing uncoated ellipsoidal viscoelastic inclusions. Datta[18] and Gubernatis[19] also consider the case of uncoated ellipsoidal inclusions in the quasistatic, long-wavelength approximation. It appears that Walpole[20] was the first to study the effect of a thin coating on the local field just around the inclusion. The effect of coated particles on the effective moduli of the composite is also found in Huang *et al.*,[21] and in Garboczi and Berryman.[22] Huang *et al.* use a generalized self-consistent formulation, whereas Garboczi and Berryman choose a new type of differential scheme that enables calculations for a composite containing a range of particle size and obtain excellent agreement when comparing their results with finite element models. A critical evaluation of effective medium theories can be found in Christensen,[15] and also in Berryman.[23] A well-documented criticism of self-consistent models of composites is the appearance of a threshold of rigidity at a finite concentration, and also a difficulty in dealing with constituents with very large contrast in material properties. As pointed out in Ref. 23, effective medium theories are inherently approximate compared to the exact bounding methods such as those of Hashin–Shtrikman[24] or the more stringent Miller bounds.[25] We also note the recent derivation by Gibiansky and Milton[26] of rigorous bounds for two-phase anisotropic media in the dynamic regime, where the bounds become areas in the complex plane. The bounding methods are currently limited to two-phase media; they cannot be used in their current form to validate self-consistent models of three-phase composites, such as coated inclusions in a viscoelastic matrix considered in the present article. A comparison of the results presented herein with exact bounding technique is therefore beyond the scope of the present article. Nevertheless, the exact bounds are often not close to each other, and effective medium theories, despite their limitations, give useful physical insight in the micromechanics of complex composites.

Cherkaoui's self-consistent model is capable of computing the effective elastic constants of an anisotropic viscoelastic material containing thinly coated ellipsoidal inclusions, but it is limited to the static case. The objective of this paper is to extend the Cherkaoui self-consistent micromechanical model to the quasistatic regime and compare the results with those predicted by the single-scattering model of Baird, Kerr, and Townend. The paper is organized as follows. In Sec. II of the paper, the micromechanical static self-consistent model is presented. In Sec. III, the frequency dependence of the elastic constants is introduced to model dynamic loading in the quasistatic approximation. The transmission loss of a slab of such a composite material is computed as a function of frequency and the results are compared, for various volume fractions, with the BKT scattering model and experimental data.

II. STATIC MICROMECHANICAL MODEL

Micromechanics attempts to approximate the global behavior of the composite material by writing the constitutive stress and strain relationships for an element of heterogeneous material on the smallest scale that entirely describes the material. This element is known as the representative volume element. The basic assumption behind the self-consistent approach introduced by Hill[13] and Budianski[14] is that the representative volume element is assumed to be a homogeneous material that has an effective shear modulus, μ_{eff}, and an effective bulk modulus, K_{eff}, surrounding the inhomogeneity, as shown in Fig. 1. In other words, the effect of neighboring phases is taken into account by considering that the inhomogeneity is no longer surrounded by a matrix of known properties but instead by the effective medium with unknown properties. This method results in implicit relationships for the effective material properties that can be solved using numerical techniques. The self-consistent (SC) model of Cherkaoui *et al.*[11] also implies that the thickness, Δa, of the coating is much smaller than the typical radius, a, of the inclusion. So, as is done in Walpole's approach,[20] Green's function techniques and Hill's interfacial operators[27] are used simultaneously. The material is assumed to be a mixture of three phases. The first phase is constituted by several ellipsoidal inclusions with elastic moduli C^I_{ijkl}. Each inclusion is surrounded by a thin layer of another phase with elastic moduli C^C_{ijkl}. The coated inclusions are bonded to the matrix defined by the elastic moduli C^M_{ijkl}. The effective properties of the composite material are determined by a self-consistent scheme where each coated inclusion is embedded in an infinitely extended medium having the (unknown) properties of the composite. The effective moduli C^{eff}_{ijkl} are therefore found from the implicit solution of an equation [Eq. (45) of Ref. 11] that is given below for convenience

$$C^{\text{eff}}_{ijmn} = C^M_{ijmn} + f_I (C^I_{ijkl} - C^M_{ijkl}) A^I_{klmn} + f_C (C^C_{ijkl} - C^M_{ijkl}) A^C_{klmn}, \tag{1}$$

where the volume fractions of the inclusion and the coating are f_I and f_C, respectively, and where A^I_{ijrs} and A^C_{ijpq} are known as the strain localization tensors of the inclusion and coating, respectively. The tensors $\boldsymbol{A}$ are defined by Eqs. (2a) and (2b) below, where it is assumed that the thickness of the coating is very small compared to the size of the inclusion

$$A^I_{ijrs}=\left\{\frac{f_I}{f_I+F_C}(I_{ijrs}+T^I_{ijkl}(C^{\text{eff}})\Delta C^I_{klrs}) + \frac{f_C}{f_I+F_C}(I_{ijmn}+T^I_{ijkl}(C^{\text{eff}})\Delta C^C_{klmn})(I_{mnrs} + T^I_{ijkl}(C^C)\Delta C^{IC}_{klrs})\right\}^{-1}, \tag{2a}$$

$$A^C_{ijpq}=(I_{ijmn}+T^I_{ijkl}(C^C)\Delta C^{IC}_{klmn})A^I_{mnpq}, \tag{2b}$$

where $\boldsymbol{I}$ is the identify tensor defined by $I_{ijkl}=\frac{1}{2}(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})$, and where the contrast tensors $\boldsymbol{\Delta C}$ are defined by Eqs. (3a)–(3c) below

$$\Delta C^I_{ijkl}=C^I_{ijkl}-C^{\text{eff}}_{ijkl}, \tag{3a}$$

$$\Delta C^C_{ijkl}=C^C_{ijkl}-C^{\text{eff}}_{ijkl} \tag{3b}$$

$$\Delta C^{IC}_{ijkl}=C^I_{ijkl}-C^C_{ijkl}. \tag{3c}$$

In Eq. (2), the tensor $\boldsymbol{T}$ is the integral of the Green's tensor, $\boldsymbol{\Gamma}$, according to

$$T^I_{ijkl}=\int_{V_I}\Gamma_{ijkl}(r-r')dr, \quad r'\in V_I, \tag{4}$$

which can be calculated for the cases of spheres, ellipsoidal inclusions, or elliptic cylinders, according to the results of Eshelby[28] and Mura.[29] (See Appendix A for the case of an isotropic matrix containing coated spherical inclusions.)

Equations (1)–(4) are general in that they allow for the resolution of the effective modulus of the composite material in the case of *anisotropic constituent phases and coated ellipsoidal inclusions*. The solution for C^{eff}_{ijkl} is implicit and it can be obtained using a Newton–Raphson numerical scheme, keeping in mind that ill-conditioning effects may occur when calculating the inclusion localization tensor, A^I_{ijkl}. (Recently, Lanhong *et al.*[30] have proposed an explicit, self-consistent scheme to analyze the problem of coated spheres in an isotropic medium.) The self-consistent approach described by Eqs. (1)–(4) is a powerful technique because it is not limited to spherical inclusions and isotropic media, and it requires relatively little computation compared to finite element analysis.

Figure 2 shows the dependence of the Young's modulus on volume fraction for the case of the composite material studied by Huang and Gibson.[31] The material is a polyester resin containing Scotchlite, glass-coated, spherical, voided microsphere with average nondimensional thickness $\Delta a/a=0.002$. The material properties of the composite are shown in Table I. In Fig. 2, the solid line is the approximation based on Eqs. (1)–(4), the self-consistent model (SC). The symbols (+) indicate experimental data reported by Huang and Gibson. For comparison, predictions based on the much simpler Kerner model (see Appendix B) are shown by a dotted line.

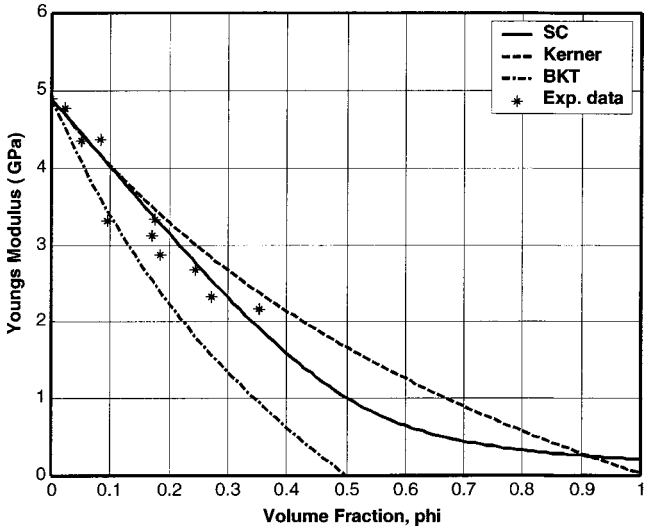


FIG. 2. Effective Young's modulus as calculated in the static case by BKT (Ref. 9), SC (Ref. 11), and Kerner (Ref. 12) models for the composite material given in Ref. 9.

Although Kerner's model should apply only at low concentrations, Madigosky *et al.*[32] have shown that it agrees well with experimental data up to 45% void content. Figure 2 also includes predictions, in the static limit, based on the Bair–Kerr–Townend[9] (BKT) dynamic model that will be discussed in the next section.

Both the SC model and the Kerner model are in good agreement with the data of Huang and Gibson. The BKT model, in this static case, seems less accurate. The advantage of the BKT model is that it is a single-scattering approach that predicts effective material parameters in the dynamic case (low *ka*). The advantage of the SC model is its ability to handle nonspherical inclusions and anisotropic media.

III. QUASISTATIC REGIME

When a longitudinal harmonic plane wave is incident on a scatterer of characteristic length scale *a* much smaller than the wavelength, the scatterer undergoes a combination of radial motion and translational motion,[33] both of which contribute to the conversion of strain energy. Strong conversion into extensional strain occurs in the vicinity of the scatterer, in the near field. Because Cherkaoui's self-consistent model takes into account this extensional strain via the localization tensors [Eq. (2)], it captures the fundamental mechanism by which losses occur in the composite. Therefore, to extend the SC model to the quasistatic (low *ka*) regime, one only has to introduce a complex modulus for the viscoelastic matrix. The frequency dependence of the complex matrix modulus can be found from either the Havriliak–Negami model[34] or from measurements based on time–temperature equivalence and

TABLE I. Constituent properties for composite tested by Huang and Gibson (Ref. 31).

	Matrix	Coating	Inclusion
μ (GPa)	1.75	28.5	0
ν	0.40	0.23	0

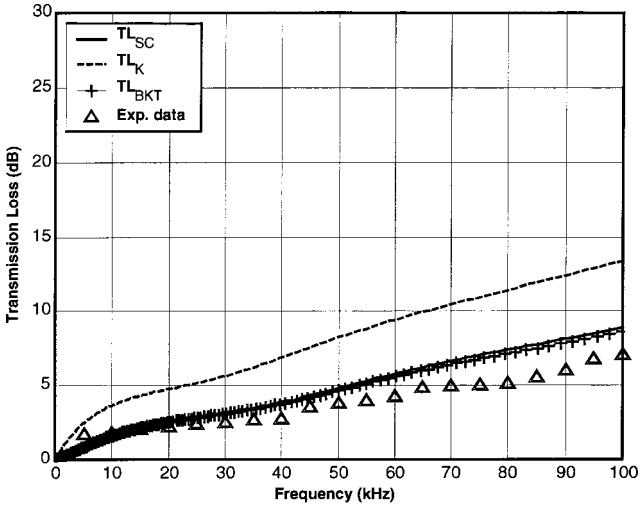


FIG. 3. Transmission loss calculated from BKT (Ref. 9), SC (Ref. 11), and Kerner (Ref. 12) models for 1-cm-thick stiff matrix material composite layer with experimental data for $\phi=5\%$.

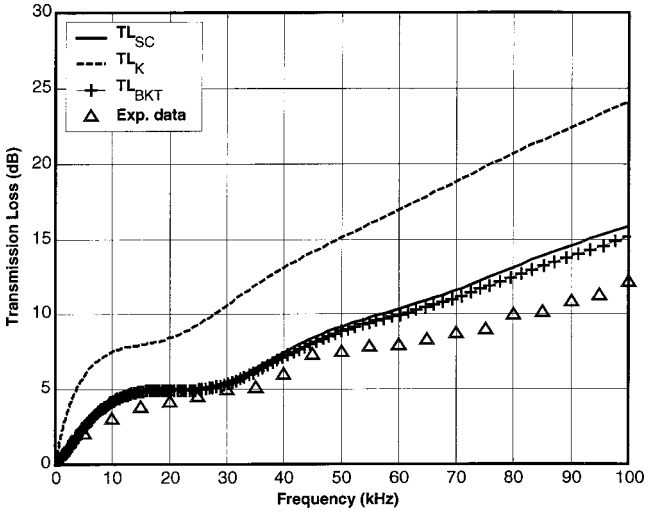


FIG. 5. Transmission loss calculated from BKT (Ref. 9), SC (Ref. 11), and Kerner (Ref. 12) models for 1-cm-thick stiff matrix material composite layer with experimental data for $\phi=13\%$.

Kramers–Kronig relations.[35] The SC model is therefore directly amenable to predicting effective moduli of anisotropic composites containing coated inclusions with morphologies that are not spherical, in the low *ka* regime.

It is instructive to compare the predictions based on the SC micromechanical model [Eqs. (1)–(4)] with predictions and data obtained by the scattering model of Baird, Kerr, and Townend.[9] The BKT model starts with the exact expression for scattering of a plane wave by a single spherical coated inclusion. The scattering coefficients derived in this manner capture the near-field physics around the scatterer. The single-scatterer approach is then extended to the case of multiple scatterers using the Gaunaurd–Überall approach. However, at this stage the effects of multiple scattering are neglected. In other words, the scattered contribution to the field around an inclusion is not rescattered by another inclusion. The analysis is restricted to far-field scattering, which is consistent with low concentrations of scatterers. Figures 3–7 represent the frequency dependence of the transmission loss, in decibels, of a plane wave traveling through a slab of composite material containing coated spherical inclusions, submerged in water. The constituent material properties are given in Table I of Ref. 9. The complex, frequency-dependent shear modulus of the viscoelastic matrix material is defined as $\mu_M(f)=\mu_M^R(1-i\tan\delta_M)$, where μ_M^R and δ_M are given by Eqs. (70) and (71) of Ref. 9. The transmission loss is defined as $\mathrm{TL}=-10\log_{10}(TT^*)$ where T is the plane-wave amplitude transmission coefficient and where the asterisk denotes complex conjugate. The expression for T is given, for instance, in Ref. 36. With the $e^{-i\omega t}$ convention, it is written as

$$T=2\left[2\cos(k_{\mathrm{eff}}l)-i\left(\frac{Z_{\mathrm{eff}}}{\rho c}+\frac{\rho c}{Z_{\mathrm{eff}}}\right)\sin(k_{\mathrm{eff}}l)\right]^{-1}, \tag{5}$$

where ρc is the characteristic impedance of water, and where $k_{\mathrm{eff}}=\omega/c_{\mathrm{eff}}$ is the effective wave number in the slab which has complex impedance

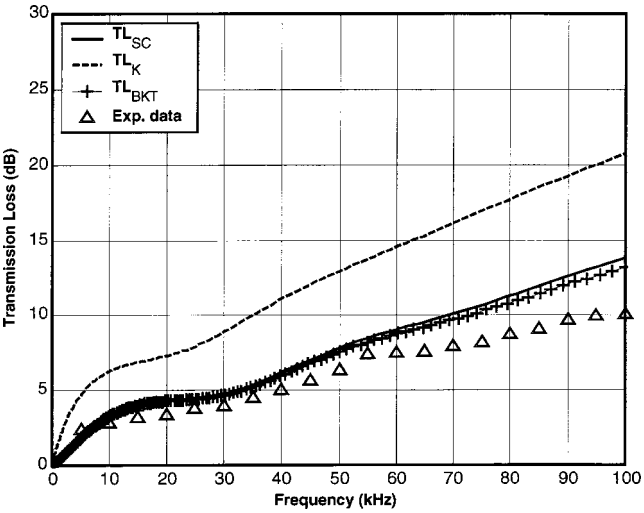


FIG. 4. Transmission loss calculated from BKT (Ref. 9), SC (Ref. 11), and Kerner (Ref. 12) models for 1-cm-thick stiff matrix material composite layer with experimental data for $\phi=9\%$.

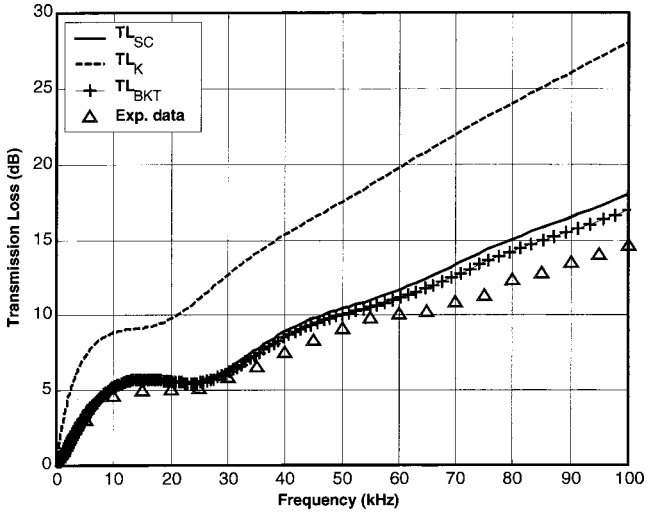


FIG. 6. Transmission loss calculated from BKT (Ref. 9), SC (Ref. 11), and Kerner (Ref. 12) models for 1-cm-thick stiff matrix material composite layer with experimental data for $\phi=17\%$.

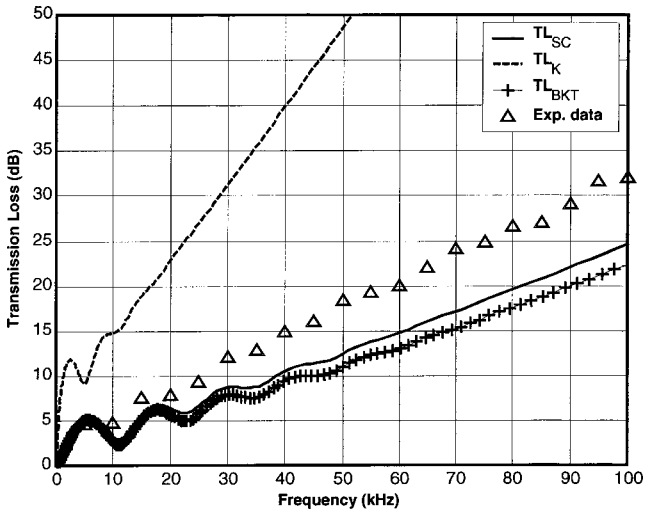


FIG. 7. Transmission loss calculated from BKT (Ref. 9), SC (Ref. 11), and Kerner (Ref. 12) models for 2.5-cm-thick soft matrix material composite layer with experimental data for $\phi=10\%$.

$$Z_{\text{eff}}=\rho_{\text{eff}}c_{\text{eff}}=\sqrt{M_{\text{eff}}\rho_{\text{eff}}}. \tag{6}$$

Here, ρ_{eff} is the effective density of the composite and $M_{\text{eff}}=K_{\text{eff}}+\frac{4}{3}\mu_{\text{eff}}$ is the plane-wave modulus. Figures 3–6 show the transmission loss through a 1-cm-thick ''stiff'' viscoelastic material with volume fractions of coated inclusions of $\phi=5\%$, 9%, 13%, and 17%, respectively, and Fig. 7 is for a 2.5-cm-thick ''soft'' viscoelastic material with volume fraction $\phi=10\%$. Here, $\phi=f_I+f_C$ denotes the combined volume fractions of the inclusion and the coating. The BKT model and experimental data[9] are compared with the SC model and with the Kerner model, where the frequency dependence of the composite is introduced by simply considering the frequency dependence of the shear modulus of the viscoelastic matrix material.

The results shown in Figs. 3–7 indicate that both the micromechanical model (SC) and the BKT model predict very similar transmission losses, despite the fact that the two approaches are very different. The SC model is an *ad hoc* extension to the quasistatic regime of a static model where mode conversion is taken into account by means of localization stress tensors in the vicinity of the scatterers. The BKT approach, on the other hand, takes losses into account by evaluating far-field single scattering. In the case of the stiff matrix, both models slightly overestimate the measured transmission loss (by up to 3 dB at 100 kHz), whereas in the case of a soft matrix, both models underestimate the measured transmission loss by up to 10 dB at 100 kHz. Several factors contribute to this discrepancy. First, Baird *et al.*[9] have shown that the error is considerably reduced if one considers a size distribution of microspheres. Second, the estimated properties of the matrix have been determined by using a fit based on DMTA[37] data and WLF time–temperature shifts.[38] Finally, experimental errors associated with the data itself are unavoidable.

Nevertheless, both the BKT and the SC model outperform the Kerner model, which significantly overpredicts the transmission loss in all cases. The Kerner model places too little emphasis on the reinforcement of the coating material, thus predicting a very compliant effective material, which, in turn, greatly overestimates the transmission losses of the composite. The advantage of the BKT model is the straightforward computations. The advantage of the SC model is that it can be used for nonspherical scatterers embedded in anisotropic materials.

IV. CONCLUSIONS

The self-consistent model can be used to predict the effective moduli of composites in the low-frequency regime. Losses are taken into account by introducing the frequency-dependent, complex shear modulus of the viscoelastic matrix. Mode conversion is taken into account through the localization tensors that govern the micromechanical behavior near the inclusions. The model is in good agreement with prediction and data of Baird *et al.*[9] The advantage of the self-consistent model is that it is applicable to the case of nonspherical inclusions embedded in anisotropic materials. Although the calculations shown in the figures of this paper are only valid for the case of spherical coated inclusions, the model may also give insight to scattering effects that could occur when the composite is exposed to large hydrostatic pressures, when the coated inclusions become spheroidal. For such a case, the resolution of the scattering problem using the Gaunaurd–Überall approach becomes very difficult, if not impossible; however, no complexity is added to the SC model. The SC model is also directly applicable to the case of anisotropy caused by long coated fibers in unidirectional composites. These applications are currently under study.

ACKNOWLEDGMENT

This research was partly supported by the Office of Naval Research, Code 334.

APPENDIX A: SPHERICAL INCLUSIONS IN AN ISOTROPIC MATERIAL

For spherical coated inclusions embedded in an isotropic host, the tensor $\boldsymbol{T}$ reduces to

$$T^I_{ijkl}=\frac{1}{30\mu(1-\nu)}\{(4-5\nu)(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})-\delta_{ij}\delta_{kl}\}, \tag{A1}$$

where μ and ν are the shear modulus and Poisson's ratio of the suspending medium, respectively. The Lamé constants of the effective material are given by Eqs. (A2) and (A3) below.

$$\lambda^{\text{eff}}=\lambda^M+f^I\left\{\left[(\lambda^I-\lambda^M)M^I+\frac{2}{3}(M^I-N^I)(\mu^I-\mu^M)\right]+3\frac{\Delta a}{a}\left[(\lambda^C-\lambda^M)M^I\phi^C+\frac{2}{3}(M^I-N^I\psi^C)\times(\mu^C[0]-\mu^M)\right]\right\}, \tag{A2}$$

$$\mu^{\text{eff}}=\mu^M+f^IN^I\left\{(\mu^I-\mu^M)+3\frac{\Delta a}{a}(\mu^C-\mu^M)\psi^C\right\}, \tag{A3}$$

where

$$M^I=\frac{K^{\text{eff}}}{K^{\text{eff}}+\alpha^{\text{eff}}(K^I-K^{\text{eff}})+3\dfrac{\Delta a}{a}\left\{\dfrac{a^C K^{\text{eff}}(1-\alpha^{\text{eff}})-\alpha^{\text{eff}}K^C(1-\alpha^C)}{K^C}\right\}(K^I-K^C)}, \tag{A4}$$

$$N^I=\frac{\mu^{\text{eff}}}{\mu^{\text{eff}}+\beta^{\text{eff}}(\mu^I-\mu^{\text{eff}})+3\dfrac{\Delta a}{a}\left\{\dfrac{\beta^C\mu^{\text{eff}}(1-\beta^{\text{eff}})-\beta^{\text{eff}}\mu^C(1-\beta^C)}{\mu^C}\right\}(\mu^I-\mu^C)}, \tag{A5}$$

$$\phi^C=\frac{K^C+\alpha^C(K^I-K^C)}{K^C}, \tag{A6}$$

$$\psi^C=\frac{\mu^C+\beta^C(\mu^I-\mu^C)}{\mu^C}, \tag{A7}$$

$$\alpha^x=\frac{1}{3}\left(\frac{1+\nu^x}{1-\nu^x}\right), \tag{A8}$$

and

$$\beta^x=\frac{2}{15}\left(\frac{4-5\nu^x}{1-\nu^x}\right). \tag{A9}$$

APPENDIX B: KERNER MODEL

The Kerner model[12,32] is often used because of its simplicity. The effective bulk and shear moduli are given by Eqs. (A1) and (A2), respectively.

$$K_{\text{eff}}=\frac{\dfrac{(1-\phi)K}{3K+4\mu}+\displaystyle\sum_i\frac{\phi_i K_i}{3K_i+4\mu}}{\dfrac{(1-\phi)}{3K+4\mu}+\displaystyle\sum_i\frac{\phi_i}{3K_i+4\mu}} \tag{B1}$$

$$\mu_{\text{eff}}=\mu\frac{\displaystyle\sum_i\frac{\phi_i\mu_i}{(7-5\nu)\mu+(8-10\nu)\mu_i}+\frac{1-\phi}{15(1-\nu)}}{\displaystyle\sum_i\frac{\phi_i\mu}{(7-5\nu)\mu+(8-10\nu)\mu_i}+\frac{1-\phi}{15(1-\nu)}}, \tag{B2}$$

where ϕ is the void fraction, K and μ are the bulk and shear moduli of the host medium, respectively, and the subscript i refers to the ith inclusion.

[1] I. A. Chaban, "Self-consistent field approach to calculation of the effective parameters of microinhomogeneous media," Sov. Phys. Acoust. **10**, 298–304 (1965).
[2] G. T. Kuster and M. N. Toksoz, "Velocity and attenuation of seismic waves in two-phase media. I. Theoretical formulations, Geophysics **39**, 587–606 (1974).
[3] G. C. Gaunaurd and H. Überall, "Resonance theory of the effective properties of perforated solids," J. Acoust. Soc. Am. **71**, 282–295 (1982).
[4] P. C. Waterman and R. Truell, "Multiple scattering of waves," J. Math. Phys. **2**, 513–537 (1961).
[5] V. K. Varadan, Y. Ma, and V. V. Varadan, "A multiple scattering theory for elastic wave propagation in discrete random media," J. Acoust. Soc. Am. **77**, 375–385 (1985).
[6] K. P. Scharnhorst, "Comments on the applicability of the Kuster–Toksoz method to the derivation of the dynamic material parameters of inhomogeneous media," J. Acoust. Soc. Am. **82**, 692–695 (1987).
[7] F. Kerr, "The Scattering of a Plane Elastic Wave by Spherical Elastic Inclusions," Int. J. Eng. Sci. **30**, 169–186 (1992).
[8] A. N. Norris, "Scattering of elastic waves by spherical inclusions with applications to low frequency wave propagation in composites," Int. J. Eng. Sci. **24**, 1271–1282 (1986).
[9] A. M. Baird, F. H. Kerr, and D. J. Townend, "Wave propagation in a viscoelastic medium having fluid-filled microspheres," J. Acoust. Soc. Am. **105**, 1527–1538 (1999).
[10] B. E. Hornby, L. M. Schwartz, and J. A. Hudson, "Anisotropic effective modeling of the elastic properties of shales," Geophysics **59**, 1570–1583 (1994).
[11] M. Cherkaoui, H. Sabar, and M. Berveiller, "Micromechanical approach of the coated inclusion problem and applications to composite materials," J. Eng. Mater. Technol. **116**, 274–278 (1994).
[12] E. H. Kerner, "The elastic and thermo-elastic properties of composite media," Proc. Phys. Soc. London **69**, 808–813 (1956).
[13] R. Hill, "A self-consistent mechanics of composite materials," J. Mech. Phys. Solids **13**, 213–222 (1965).
[14] B. Budianski, "On the elastic properties of some heterogeneous media," J. Mech. Phys. Solids **13**, 223–227 (1965).
[15] R. M. Christensen, "A critical evaluation for a class of micromechanics models," J. Mech. Phys. Solids **38**, 379–404 (1990).
[16] R. M. Christensen and K. H. Lo, "Solutions for effective shear properties of three-phase sphere and cylinder models," J. Mech. Phys. Solids **27**, 315–330 (1979).
[17] J. G. Berryman, "Long-wavelength propagation in composite elastic media. II. Ellipsoidal inclusions," J. Acoust. Soc. Am. **68**, 1820–1831 (1980).
[18] S. K. Datta, "Diffraction of plane elastic waves by ellipsoidal inclusions," J. Acoust. Soc. Am. **61**, 1432–1437 (1977).
[19] J. E. Gubernatis, "Long-wave approximations for the scattering of elastic waves from flaws with applications to ellipsoidal voids and inclusions," J. Appl. Phys. **50**, 4046–4058 (1979).
[20] L. J. Walpole, "A coated inclusion in an elastic medium," Math. Proc. Cambridge Philos. Soc. **83**, 495–506 (1978).
[21] Y. Huang, K. X. Hu, X. Wei, and A. Chandra, "A generalized self-consistent mechanics method for composite materials with multiphase inclusions," J. Mech. Phys. Solids **42**, 491–504 (1994).
[22] E. J. Garboczi and J. G. Berryman, "Elastic moduli of a material containing composite inclusions: Effective medium theory and finite element computations," Mech. Mater. **33**, 455–470 (2001).
[23] J. G. Berryman, "Long-wavelength propagation in composite elastic media. I. Spherical inclusions," J. Acoust. Soc. Am. **68**, 1809–1819 (1980).
[24] Z. Hashin and S. Shtrikman, "A variational approach to the theory of the elastic behavior of multiphase materials," J. Mech. Phys. Solids **11**, 127–140 (1963).
[25] M. N. Miller, "Bounds for effective electrical, thermal, and magnetic properties of heterogeneous materials," J. Math. Phys. **10**, 1988–2005 (1969).
[26] L. V. Gibiansky and G. W. Milton, "On the effective viscoelastic moduli of two-phase media. I. Rigorous bounds on the complex bulk modulus," Proc. R. Soc. London, Ser. A **440**, 163–188 (1993).
[27] R. Hill, "Interfacial operators in the mechanics of composite media," J. Mech. Phys. Solids **31**, 347–357 (1983).
[28] J. D. Eshelby, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," in Proceedings of the Royal Society, A Vol. 241, 1957, pp. 376–396.
[29] T. Mura, *Micromechanics of Defects in Solids*, 2nd ed. (Kluwer Academic, 1987).

[30] D. Lanhong, H. Zhuping, and W. Ren, ''An explicit expression of the effective moduli for composite materials filled with coated inclusions,'' Acta Mech. Sin. **14**, 37–52 (1998).
[31] J. S. Huang and L. J. Gibson, ''Elastic moduli of a composite hollow spheres in a matrix,'' J. Mech. Phys. Solids **25**, 55–75 (1993).
[32] W. M. Madigosky, R. W. Harrison, and K. P. Scharnhorst, ''Effect of air inclusions on absorption and velocity of sound in viscoelastic materials,'' Polym. Mater. Sci. Eng. **60**, 489–496 (1989).
[33] J. Jarzynski, ''Mechanisms of Sound Attenuation in Materials,'' in *Sound and Vibration Damping with Polymers*, edited by R. D. Corsaro and L. H. Sperling (The American Chemical Society, Washington D.C., 1990), pp. 167–207.
[34] S. Havriliak and S. Negami, ''A complex plane analysis of α-dispersion in some polymer systems,'' J. Polym. Sci., Part C: Polym. Symp. **14**, 99–117 (1966).
[35] N. Lagakos, J. Jarzynski, J. H. Cole, and J. A. Bucaro, ''Frequency and temperature dependence of elastic moduli polymers,'' J. Appl. Phys. **59**, 4017–4031 (1986).
[36] D. T. Blackstock, *Fundamentals of Physical Acoustics* (Wiley, New York, 2000), pp. 163–164.
[37] Dynamic Mechanical Thermal Analyzer, Rheometrics Scientific, Inc.
[38] J. D. Ferry, *Viscoelastic Properties of Polymers* (Wiley, New York, 1970), pp. 292–351.

Optimal design of panel speaker array with omnidirectional characteristics

Mingsian R. Bai[a] and Kuochan Chung
Department of Mechanical Engineering, National Chiao-Tung University, 1001 Ta-Hsueh Road, Hsin-Chu 300, Taiwan, Republic of China

(Received 31 August 2001; accepted for publication 31 July 2002)

A panel speaker system intended for a projection screen is developed. Like other sound sources with large dimension, the panel speaker has a beaming problem in high frequencies. To alleviate the problem, panel speakers are integrated into an array, with optimal electronic compensation for omnidirectional response and array efficiency. The heart of the design procedure is a three-stage optimization scheme involving two nonlinear and nonconvex objectives. The process is interactive, allowing the array coefficients to be tailored so that the specifications of directional response can be met. The optimal design of panel speaker array is then implemented by using a multichannel digital signal processor. In addition, a Hilbert transformer is required to produce the quadrature components of the array coefficients. A small array and a large matrix were constructed to validate the implemented array signal processing system. The experimental results indicate that, without degradation of efficiency, the proposed optimization technique in conjunction with electronic compensation is effective in attaining omnidirectional radiation property. [DOI: 10.1121/1.1509435]

PACS numbers: 43.38.Ar [SLE]

I. INTRODUCTION

This work focuses on the development of a projection screen which is composed of panel speakers. This system is intended for applications such as oral presentation, public addressing, or home theater. The system integrates both the audio and video functions into one unit, which may provide certain advantages over conventional systems. The main reason for using panel speakers lies in the flatness and compactness, which makes them well suited for the application as a projection screen. In general, a large and properly designed panel speaker is less directional in high frequencies than conventional cone speakers. However, a detailed electroacoustic analysis[1] revealed that this desirable property of panel speaker comes at the expense of efficiency. Furthermore, like other sound sources with large dimension, the panel speaker will still suffer from a peculiar beaming problem at coincidence angles in high frequencies if the radiating area is large.[2] As a solution to the above problem, this paper presents a speaker array approach, using an idea that contradicts the original distributed mode concept of panel speakers. We use small panels which are as light and stiff as possible to produce coherent but directional sound beams. Then, using electronic compensation and digital signal processing, we seek to achieve simultaneously omnidirectional response and array efficiency. If individual elements are identical, then the proposed beamforming technique also is applicable to arrays of conventional loudspeakers. The analysis and design of conventional loudspeaker arrays can be found in the literature.[3–6] It is also pointed out by the reviewer that a sound and light spectacle in Mexico used an array of panel speakers as the projection screen. This technology, invented by Sound Advance, is similar to a conventional cone speaker in as much as it uses a voice coil assembly.[7] However, it uses a flat diaphragm molded from expanded polystyrene, which allows the user to flush mount it inside a wall or ceiling in ''invisible'' fashion.

The design of an omnidirectional array boils down to finding a set of array coefficients that gives rise to a ''flat spectrum'' in the wave number space. For a linear array with real coefficients, direct inversion of an all-pass flat spectrum will apparently lead to the trivial solution: only one single element is active at the origin. This case corresponds to an array with very poor efficiency. Hence, two approaches have been proposed to avoid running into the dilemma between flat spectrum and array efficiency. One approach is to introduce a phase function into the directional response. The Bessel array and the quadratic phase array (QPA) are based on this idea.[6] The array gains in the QPA are purely phase compensation in quadratic forms. Another approach[8] is to choose a white-noise-like sequence with low correlation property, e.g., the Barker code, the Huffman code, and the maximum flatness sequence to produce a flat spectrum of array radiation pattern.

Different from the earlier approaches, an optimization technique is proposed in this paper to find array coefficients that maximize two cost functions: flatness and efficiency. This problem turns out to be a nonlinear and nonconvex problem for which it is generally very difficult to locate the global optimum. Hence, instead of finding the global optimum, we are content with the solution of a three-stage suboptimal problem. The design process is interactive, allowing us to tailor the array coefficients so that the specifications of directional response can be met. The array coefficients thus found are generally complex numbers, which entail the implementation of a Hilbert transformer.[9–11] The Hilbert

[a]Electronic mail: msbai@cc.nctu.edu.tw

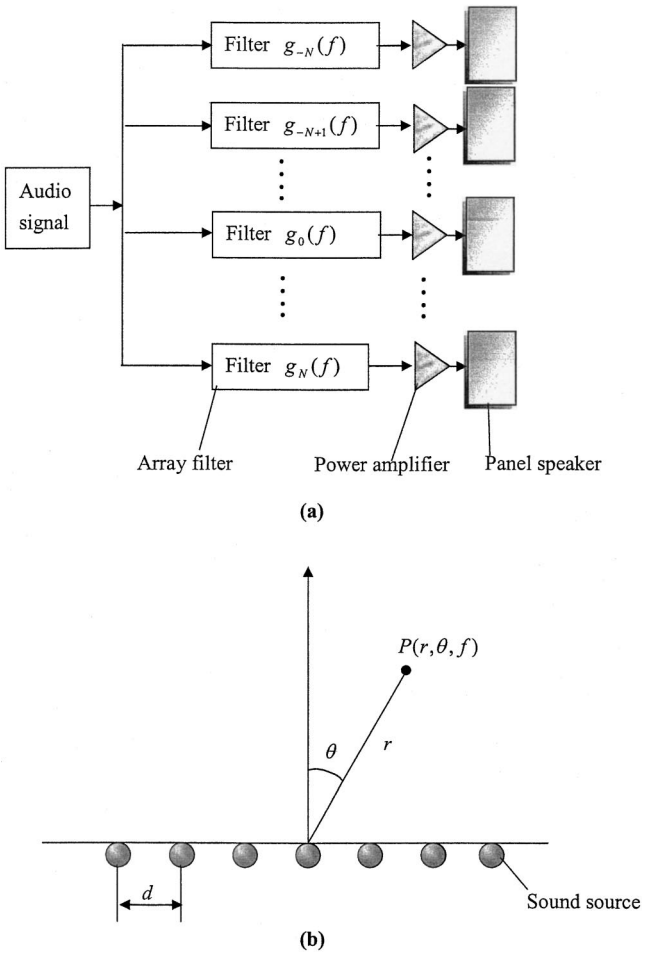


FIG. 1. A uniform linear array. (a) The schematic of a panel speaker array; (b) the array geometry.

transformer can be implemented in either IIR or FIR filter;[9] we will only discuss the latter approach.

A 5×1 panel speaker array and a 3×3 panel speaker array were constructed for experimental verification. Signal processing and electronic compensation are carried out by using a multichannel digital signal processor (DSP). Results will be compared and discussed with regard to an uncompensated array and the array obtained using the proposed optimization technique.

II. FAR-FIELD MODEL FOR UNIFORM LINEAR ARRAYS

A panel speaker array is schematically shown in Fig. 1(a). Audio signals are processed, often digitally, by a bank of filters before feeding to the power amplifiers and panel speakers. With reference to the geometry of Fig. 1(b), the far-field pressure radiated by a source array with $2N+1$ equally spaced elements can be expressed as[6,12]

$$P(r,\theta,f)=A(f,\theta)R(f,r)B(f,\theta), \tag{1}$$

where d is the spacing between two adjacent speakers, $r \gg d$ is the distance between the array center and a far-field observation point, θ is the angle measured from the normal to the array, f is the frequency, c is the sound speed, $A(f,\theta)$ is the radiation pattern of each source, $R(f,r)=r^{-1}\exp(j2\pi fr/c)$ represents the spherical spreading, $B(f,\theta)$ is the *array pattern*, defined as

$$B(f,\theta)=\sum_{n=-N}^{N} g_n(f)e^{j(2\pi nfd\sin\theta/c)}, \tag{2}$$

with $g_n(f)$ being the array coefficient of the nth element. Using the array filters, one is able to manipulate the array pattern to obtain the desired directional response.

Hereafter, we further restrict the array coefficients $g_n(f)$ to be frequency-independent, complex constants, g_n. The array pattern can then be written as

$$B(u)=\sum_{n=-N}^{N} g_n e^{jnu}, \tag{3}$$

where $u=2\pi fd\sin\theta/c$ is a dimensionless angle. Inspection of Eq. (3) reveals that the array pattern is essentially the frequency response of an FIR filter with coefficients g_n. That is, the design problem of an omnidirectional array can be regarded as the design of an FIR all-pass filter. For latter use, define the angular spectrum

$$S(u)=\|B(u)\|_2^2, \tag{4}$$

where $\|\ \|_2$ denotes the 2-norm, and the autocorrelation

$$R(k)=\sum_{n=-\infty}^{\infty} g_n g_{k-n}^{*}, \tag{5}$$

where k is the array index, g_n has a compact support within $[-N,N]$, "*" denotes complex conjugate. It can be shown that the angular spectrum is the Fourier transform of the autocorrelation, i.e.,

$$S(u)=\sum_{k=-N}^{N} R(k)e^{-jku}. \tag{6}$$

III. OBJECTIVES AND CONSTRAINTS

As mentioned earlier, the design goal of our problem is to find an array with omnidirectional characteristics and good efficiency. In this section, these design objectives will be formulated as two performance indices: *spectral flatness* and *array efficiency*.

A. Array efficiency

The array efficiency of a $(2N+1)\times 1$ array is defined as

$$\eta=\frac{R(0)}{(2N+1)|\mathbf{g}|_\infty}, \tag{7}$$

where $\mathbf{g}=\{g_n|-N\leqslant n\leqslant N, n\in N\}$, $|\mathbf{g}|_\infty=\max\{g_n|-N\leqslant n\leqslant N\}$ is the infinity norm of $\mathbf{g}$, and

$$R(0)=\sum_{n=-N}^{N}|g_n|^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}S(u)du, \tag{8}$$

where the Parseval theorem has been invoked. Array efficiency is thus the mean-squared array gains normalized by $|\mathbf{g}|_\infty$. The physical meaning of the array efficiency can be interpreted as the degree of participation of active array elements. The efficiency will be close to unity if most array elements are active with full power (large gain values g_n).

B. Spectral flatness

For the same array, the merit factor, or the spectral flatness, is defined as[12,13]

$$F=\frac{R^2(0)}{\Sigma_{k\neq 0}|R(k)|^2}. \tag{9}$$

Using Eq. (6), we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}S(u)du=\int_{-\pi}^{\pi}\sum_{k=-N}^{N}R(k)e^{jku}\,du=R(0). \tag{10}$$

According to the Parseval's relation, we have

$$\sum_{k=-N}^{N}|R(k)|^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}S^2(u)du. \tag{11}$$

Substituting Eqs. (10) and (11) into Eq. (9) gives

$$F=\frac{(\int_{-\pi}^{\pi}S(u)du)^2}{2\pi\int_{-\pi}^{\pi}[S^2(u)-R^2(0)]du}. \tag{12}$$

The denominator of F can also be written as

$$\int\ [S^2(u)-R^2(0)]du=\int_{-\pi}^{\pi}[S(u)-R(0)]^2du. \tag{13}$$

From Eqs. (12) and (13), the spectral flatness F is the ratio of the mean-square spectrum over the spectral variance. An array with omnidirectional response tends to have large spectral flatness.

C. Constraints

Unfortunately, the aforementioned objective functions are not sufficient to reach a unique solution because any array coefficients differing within a scaling and/or a rotation will lead to identical F and η. This fact will be detailed in the following analysis.

Let $\mathbf{C}$ and $\mathbf{R}$ be the sets of real numbers and imaginary numbers, respectively. Suppose an array coefficient set $\mathbf{g}_1$ can be obtained from another set $\mathbf{g}=\{g_n|-N\leqslant n\leqslant N\}$ by a magnitude scaling r and a finite rotation u_0, i.e.,

$$\mathbf{g}_1=T(r,u_0,\mathbf{g})=\{rg_ne^{-jn\delta}|-N\leqslant n\leqslant N, r\in\mathbf{C}, u_0\in\mathbf{R}\}. \tag{14}$$

Then, it is not difficult to verify the spectrum of the new coefficients

$$S_1(u)=|r|^2S(u+u_0). \tag{15}$$

The spectrum is scaled by a factor $|r|^2$ and shifted in u axis by u_0. Furthermore, the spectral flatness and array efficiency remain invariant under the transformation $T(r,u_0,g)$, i.e.,

$$F_1=F,\quad \eta_1=\eta, \tag{16}$$

and both coefficient sets are considered "equivalent." Consequently, additional constraints are needed to resolve this nonuniqueness problem. These constraints are derived from the following theorem. *Theorem*: For a $(2N+1)\times 1$ array, consider an array coefficient set $\mathbf{g}=\{g_n|-N\leqslant n\leqslant N\}$, with

$$|\mathbf{g}|_\infty=|g_0|. \tag{17}$$

Using the equivalent transformation of Eq. (14), the array coefficient set $\mathbf{g}$ can always be transformed into a "reduced" array coefficient set $\mathbf{J}$ given by

$$\mathbf{J}=\{J_n|J_n\in\mathbf{C},\ |J_n|\leqslant 1,\ -N\leqslant n\leqslant N,\ J_0=1\ \text{ and }\ \angle J_0=\angle J_1=0\} \tag{18}$$

Proof:

Let $r_0=e^{-j\angle g_0}/|\mathbf{g}|_\infty$, $\delta_0=\angle g_0-\angle g_1$, and

$$\mathbf{J}=T(r_0,\delta_0,g)=\left\{J_n=\frac{1}{|g|_\infty}g_ne^{j(\angle g_n-\angle g_0+n(\angle g_0-\angle g_1))}|J_n\in\mathbf{C},\ -N\leqslant n\leqslant N\right\}. \tag{19}$$

Substituting Eq. (17) into (19) leads to: $J_0=1$, $J_1=|g_1/g_0|$, and $|J|=|g_n/g_0|\leqslant 1$. Hence,

$$\mathbf{J}=\{J_n|J_n\in\mathbf{C},\ |J_n|\leqslant 1,\ -N\leqslant n\leqslant N,\ J_0=1\ \text{ and }\ \angle J_0=\angle J_1=0\},$$

which is Eq. (18). Q.E.D.

Thus, on the basis of the theorem, the feasible solution set of the array design problem can be dramatically reduced by imposing the following "fundamental constraints" in optimization:

$$g_0=1,\ \angle g_0=\angle g_1=0\ \text{ and }\ |g_n|\leqslant 1,\ -N\leqslant n\leqslant N. \tag{20}$$

Under this constraint, the maximum magnitude of array gain never exceeds unity.

IV. THE THREE-STAGE OPTIMIZATION PROCEDURE

In this section, an optimization technique is proposed to find array coefficients that maximize the aforementioned two cost functions: flatness and efficiency. Unfortunately, this problem turns out to be a nonlinear, nonconvex, and multiobjective problem for which it is generally very difficult and time-consuming to locate the global optimum. Hence, instead of finding the global optimum, we seek to find a suboptimal solution by using a three-stage optimization scheme—two stages for single-objective optimization and one stage for weighting two objectives.

A. The first stage: Phase optimization

To reiterate, the goal for the array design is to find array coefficients with high array efficiency and flat spectrum. For high efficiency, the magnitude of each array coefficient should be close to unity. For flat spectrum, the array coefficient set should be a random sequence with low correlation property. Putting these two statements together, one may conclude that the desired array coefficient set should be a random sequence satisfying two conditions: $|g_n|\approx 1$ and $\angle g_n$ is a random number over $[0,2\pi]$. Thus, in the first stage of optimization, we restrict the magnitude of each array coefficient at a "full efficiency" state, i.e., $|g_n|=1$, and adapt the phases as randomly as possible. The optimization problem in this stage can be stated as

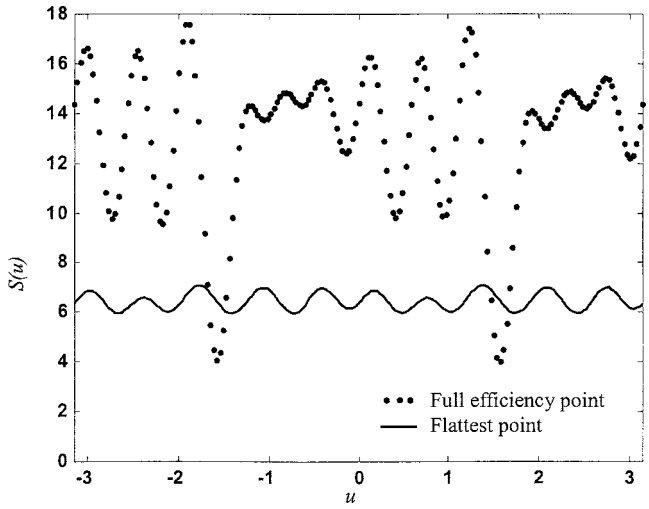


FIG. 2. The angular spectra at the full efficiency point and the flattest point for the 13×1 optimal array.

$$\underset{\theta_k \in R, -N \leqslant k \leqslant N}{\text{Max}} \quad F = \frac{R^2(0)}{\Sigma_{k \neq 0} R^2(k)}, \tag{21}$$

subject to the aforementioned fundamental constraints, and the ''full efficiency constraint,'' defined as

$$|g_k| = 1, \quad -N \leqslant k \leqslant N. \tag{22}$$

In this setting, the optimization in the first stage has only a single objective, spectral flatness. The *constrained steepest descent* (CSD)[14] method is employed to find the local maximum, or the ''full efficiency point.''

However, the result of the search is quite sensitive to initial conditions because the problem is nonlinear and nonconvex. As motivated by the characteristics of the optimal array, we thus adopted a heuristic but efficient approach and assigned random numbers to the phases as the initial guess. A simulation result obtained using the first-stage optimization for a 13×1 array is shown in Fig. 2. This simulation took approximately 1 min on a personal computer. The result was found by performing the first-stage optimization 20 times and selecting the flattest pattern as the full efficiency point. The step size of optimal search starts at 0.3 and decreases for convergence. Each optimization converged within 100 iterations. The calculated array coefficients are listed in Table I. It can be seen from the result that the angular spectrum fluctuates randomly since spectral flatness is not optimized in this stage.

B. The second stage: Magnitude and phase optimization

In this stage, both magnitude and phase of array coefficients are adjusted to further improve spectral flatness. After phase adapting, the only way to obtain more flatness is to adjust the magnitude of array coefficients. In effect, such optimization procedure is a ''tapering'' process of magnitude to get a broader and flatter spectrum, at the cost of array efficiency. The optimal problem in this stage is formulated as follows:

TABLE I. The array coefficients for the 13×1 optimal array at (a) the full efficiency point and (b) the flattest point.

Array index	(a)	(b)
−6	$\exp(j0.06)$	$0.21\exp(-j1.27)$
−5	$\exp(-j1.05)$	$0.51\exp(-j2.10)$
−4	$\exp(-j1.81)$	$0.58\exp(-j2.10)$
−3	$\exp(-j0.63)$	$0.72\exp(-j1.67)$
−2	$\exp(-j1.41)$	$1.00\exp(-j2.40)$
−1	$\exp(j3.07)$	$0.76\exp(j3.13)$
0	1	1
1	1	0.76
2	$\exp(j1.43)$	$1.00\exp(j2.40)$
3	$\exp(-j2.55)$	$0.73\exp(-j1.48)$
4	$\exp(j1.78)$	$0.58\exp(j2.08)$
5	$\exp(-j2.05)$	$0.50\exp(-j1.05)$
6	$\exp(-j0.10)$	$0.21\exp(j1.25)$
Flatness	18.74	360
Efficiency	1	0.50

$$\underset{\theta_k \in R, -N \leqslant k \leqslant N}{\text{Max}} \quad F = \frac{R^2(0)}{\Sigma_{k \neq 0} R^2(k)}, \tag{23}$$

subject to the fundamental constraints. The initial guess is the full efficiency point found in the last stage. The optimization in this stage is a nonlinear and nonconvex problem with one objective, flatness. CSD is used to find the suboptimal solution that is called the ''flattest point.''

The example of the 13×1 array is used again for the second-stage optimization. The result corresponding to the flattest point is also shown in Fig. 2. The array coefficients are listed in Table I. If the status of each iteration in this optimization stage is recorded, a trade-off can be observed between the flatness and efficiency. The search path for the 13×1 array is shown in Fig. 3. It can be seen in the plot that flatness is a monotonically increasing function of efficiency. The curve is terminated at both ends by the flattest point and the full efficiency point, respectively. The search path is

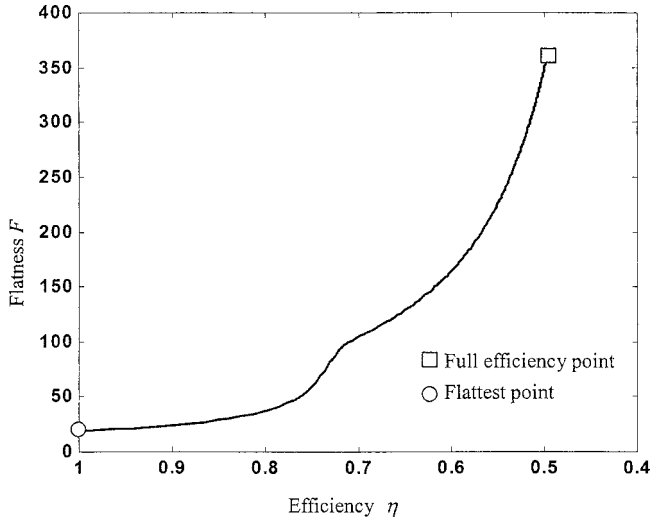


FIG. 3. The search path of efficiency and flatness for the 13×1 optimal array.

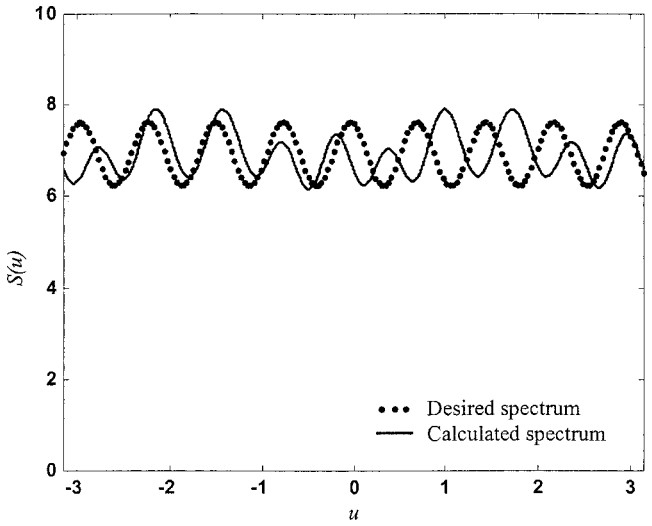


FIG. 4. The desired spectrum vs calculated spectrum for the 13×1 optimal array ($F=200$) found in the third stage.

stored in the computer for final tuning of the efficiency and flatness.

C. The third stage: Tuning of efficiency and flatness

As shown in Fig. 3, the search path for the array lies within the window

$$0.50 \leqslant \eta \leqslant 1, \tag{24}$$

and

$$19.68 \leqslant F \leqslant 360, \tag{25}$$

which renders the reachable performance limit in the design. In the third stage, the following template is employed as the desired spectrum:

$$S_d(u) = R_d(0) + A_{\text{ripple}} \sin u/T, \tag{26}$$

where $R_d(0)$ and A_{ripple} represent, respectively, the mean and the ripple size of the desired spectrum. For example, let $R_d(0)/A_{\text{ripple}}=10$. Substituting Eq. (26) into Eq. (12), the corresponding flatness should be

$$F_d = 2(R_d/A_{\text{ripple}})^2 = 200. \tag{27}$$

Using the search path in Fig. 3, one can find the corresponding efficiency to be 0.56. Recall the definition of array efficiency

$$\eta = R_d^2(0)/2N + 1 = 0.56, \tag{28}$$

where $N=6$ for a 13×1 array. Solving Eqs. (27) and (28) yields $R_d(0)=7.30$ and $A_{\text{ripple}}=0.73$, based on which the desired spectrum is given by

$$S_d(u) = 7.28 + 0.728 \sin \frac{u}{T}. \tag{29}$$

If the result is acceptable, the corresponding array coefficients will be retrieved from the computer and the actual spectrum will be calculated. Otherwise, one should select another $R_d(0)/A_{\text{ripple}}$ and repeat this optimization stage. The desired spectrum and the actual spectrum for the example are compared in Fig. 4. Two spectra have similar ripple size and spectrum mean. The flow chart of the optimization with three stages is shown in Fig. 5.

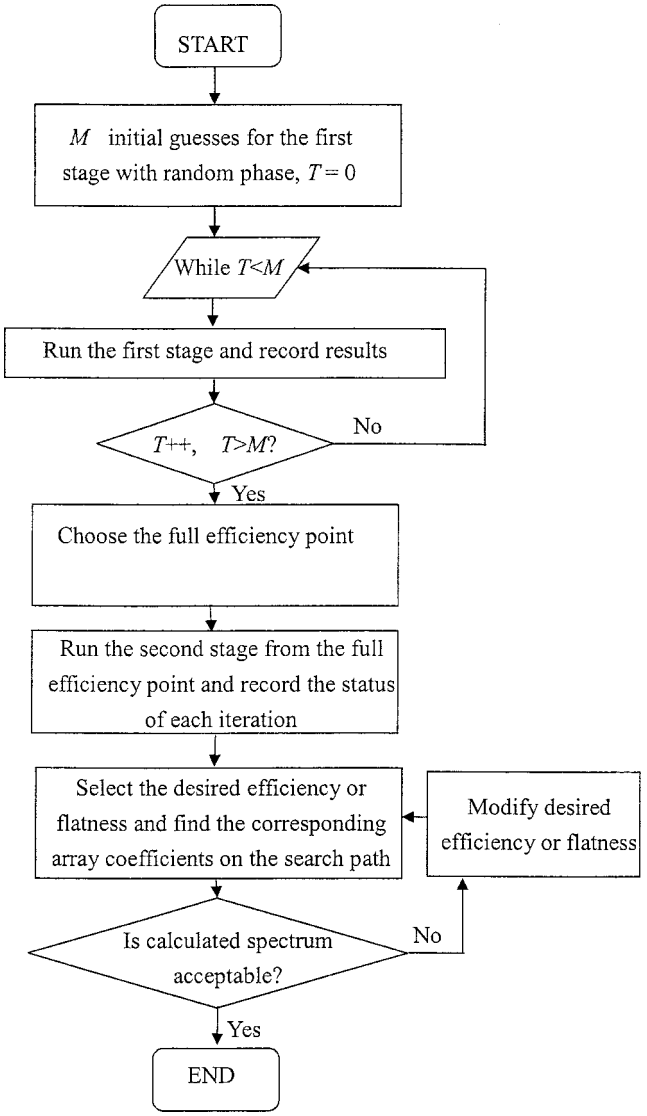


FIG. 5. The flow chart of the optimization procedure.

D. QPA array versus the optimal array

To justify the proposed technique, the 13×1 array designed using our optimization method is compared to the QPA array ($z=18$) where z is the shape factor. The array gains in the QPA are purely phase compensation in quadratic forms, $z(1-|\theta|/\pi)\theta/\pi$. The details of the QPA array can be found in Ref. 6. The array coefficients used in this simulation are listed in Table II. The angular spectra in Fig. 6 show that the spectrum of our array appears flatter than that of the QPA array. This is also reflected in Table II, where the calculated flatness is 11.5 vs 114.9 for the QPA array and our array, respectively, with identical efficiency.

V. EXPERIMENTAL INVESTIGATIONS

Two panel speaker arrays are constructed for experimental verification of the proposed array signal-processing techniques. In this section, a technical, but critical, issue in

TABLE II. The comparison of 13×1 QPA ($z=18$) and the 13×1 optimal array when the optimal array and QPA has the same efficiency.

Array index	QPA ($z=18$)	Optimal array
−6	0.74	0.35 exp(−j0.78)
−5	−0.78	0.61 exp(−j1.78)
−4	0.96	0.66 exp(−j1.92)
−3	−1.00	0.90 exp(−j1.22)
−2	0.45	1.00 exp(−j2.09)
−1	0.67	0.90 exp(j3.09)
0	−0.86	1
1	−0.67	0.94
2	0.45	1.00 exp(j2.03)
3	1.00	0.88 exp(−j2.02)
4	0.96	0.62 exp(j1.79)
5	0.78	0.59 exp(−j1.52)
6	0.74	0.37 exp(j0.57)
Flatness	11.53	114.9
Efficiency	0.63	0.63

implementing complex array coefficients shall be addressed, followed by hardware implementation of a 5×1 small array and 3×3 large array.

A. The implementation of complex array coefficients

As noted earlier, the resulting array coefficients g_n obtained using the proposed optimization techniques are generally complex. Although they are constants in nature, the approximation of which calls for the use of frequency-dependent filters

$$g_n(\omega)\approx \mathrm{Re}\{g_n\}+j\ \mathrm{Im}\{g_n\}, \tag{30}$$

where ω is the digital frequency, n is the array index, and $j=\sqrt{-1}$. The implementation of the filters is schematically shown in Fig. 7(a). The relationship between the input a_n and output b_n is given as

$$A(e^{j\omega})=[\mathrm{Re}\{g_n\}+H(e^{j\omega})\mathrm{Im}\{g_n\}]B(e^{j\omega}), \tag{31}$$

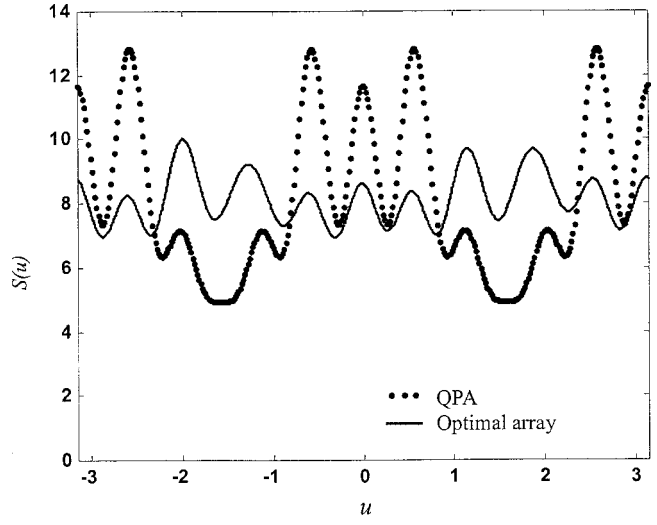


FIG. 6. The spectra of the QPA ($z=18$) and the optimal array (13×1), where both arrays have equal efficiency.

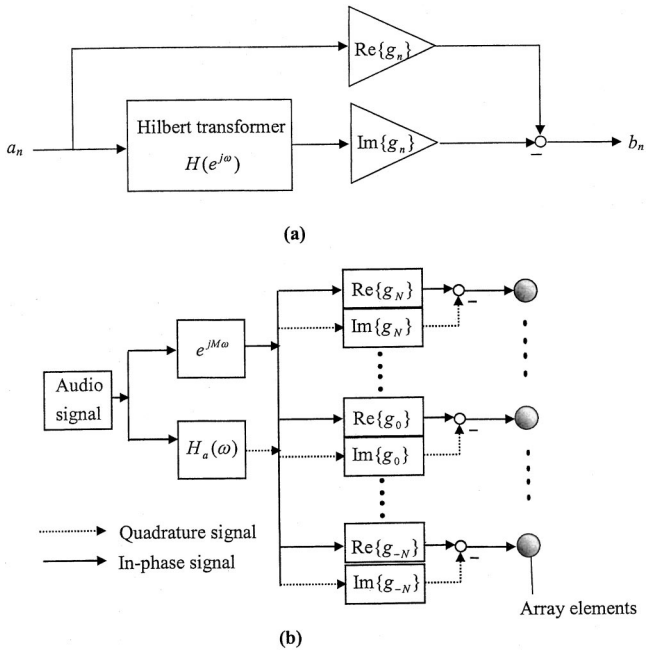


FIG. 7. The implementation diagram of complex array coefficients. (a) Single-channel array filter and the Hilbert transformer; (b) multichannel implementation.

where $A(e^{j\omega})$ and $B(e^{j\omega})$ are discrete Fourier transforms of a_n and b_n, respectively. The *Hilbert transformer*[9] $H(e^{j\omega})$ serves to generate a quadrature component. The frequency response of an ideal Hilbert transformer is

$$H(e^{j\omega})=\begin{cases} -j, & 0\leqslant\omega\leqslant\pi \\ j, & -\pi\leqslant\omega<0, \end{cases} \tag{32}$$

and the associated impulse response $h[n]$ is

$$h[n]=\begin{cases} \dfrac{2}{\pi}\dfrac{\sin^2(\pi n/2)}{n}, & n\neq 0 \\ 0, & n=0, \end{cases} \tag{33}$$

which is apparently noncausal. Hence, a delay of M samples with truncation should be introduced to approximate the ideal Hilbert transformer

$$H_a(e^{j\omega})=\begin{cases} e^{-j(M\omega+\pi/2)}, & 0\leqslant\omega\leqslant\pi \\ e^{-j(M\omega-\pi/2)}, & -\pi\leqslant\omega<0. \end{cases} \tag{34}$$

The frequency response of the modified array filter then becomes

$$g_n(\omega)\approx e^{-jM\omega}\{\mathrm{Re}\{g_n\}+j\ \mathrm{Im}\{g_n\}\}=e^{-jM\omega}g_n. \tag{35}$$

Note that no waveform distortion will arise due to the pure delay. Figure 7(b) shows the implementation in more detail. Using the implementation shown in Fig. 7(b) and substituting Eq. (35) into (2), we have

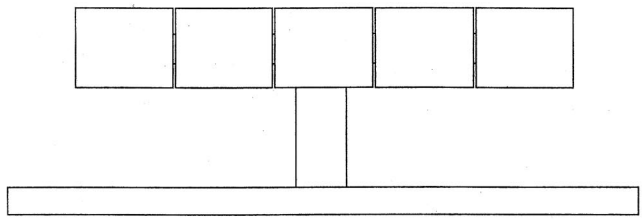
$$B(\theta,\omega)=e^{-jM\omega}\sum g_n e^{j2\pi fd\sin\theta/c}, \tag{36}$$

where $f=f_s\omega/2\pi$, f_s is the sampling frequency, f is analog frequency, and ω is digital frequency. Hence, the system in Fig. 7(b) yields an array pattern with the same magnitude response as the desired pattern. The Hilbert transformer

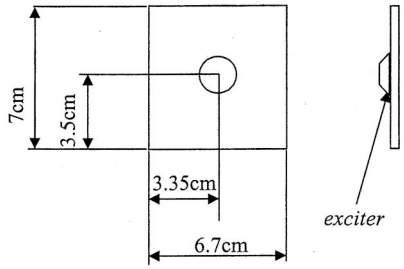
can be implemented by either an FIR[9] filter or an IIR[10,11] filter. In this work, we chose to use the FIR implementation. The Kaiser window approximation for a Hilbert transformer of order M takes the form[9]

$$h[n]=\begin{cases}\dfrac{I_o\{\beta(1-[(n-n_d)/n_d]^2)^{1/2}\}}{I_o(\beta)} \\ \quad\times\left\{\dfrac{2}{\pi}\dfrac{\sin^2[\pi(n-n_d)/2]}{n-n_d}\right\}, & 0\leqslant n\leqslant M \\ 0, \quad \text{otherwise.}\end{cases} \tag{37}$$

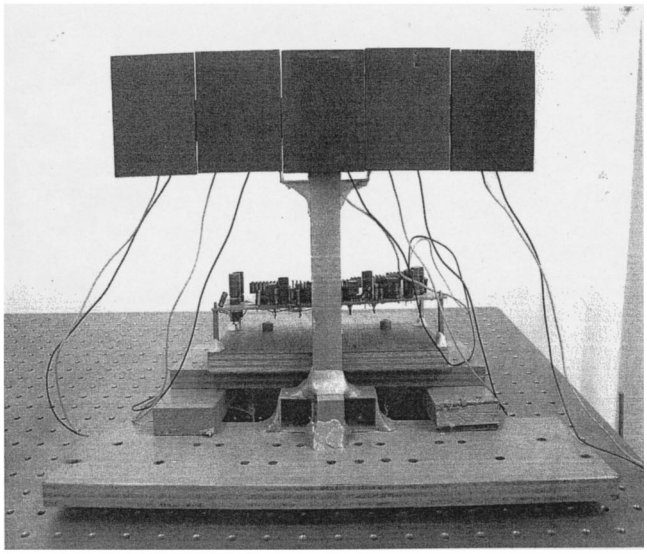
In the equation, $n_d=M/2$. It is noted that Hilbert transform will introduce frequency-dependent delay and result in some waveform distortion.



(a)



(b)



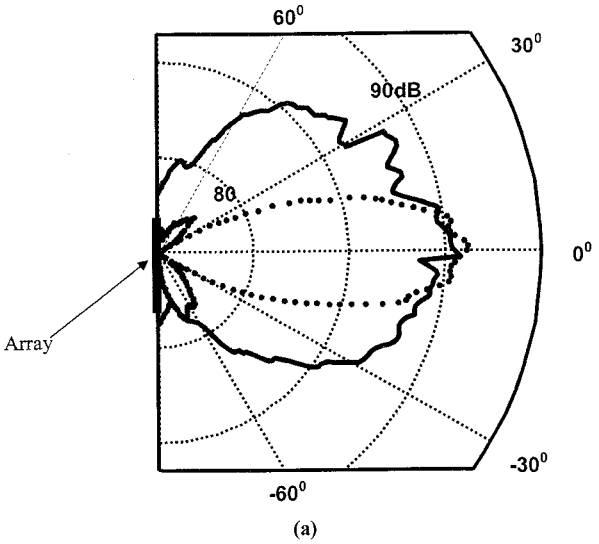
(c)

FIG. 8. The small panel speaker array. (a) Configuration of the 5×1 panel speaker array; (b) dimensions of panel speakers and the location of exciter; (c) the photo of the 5×1 panel speaker array.

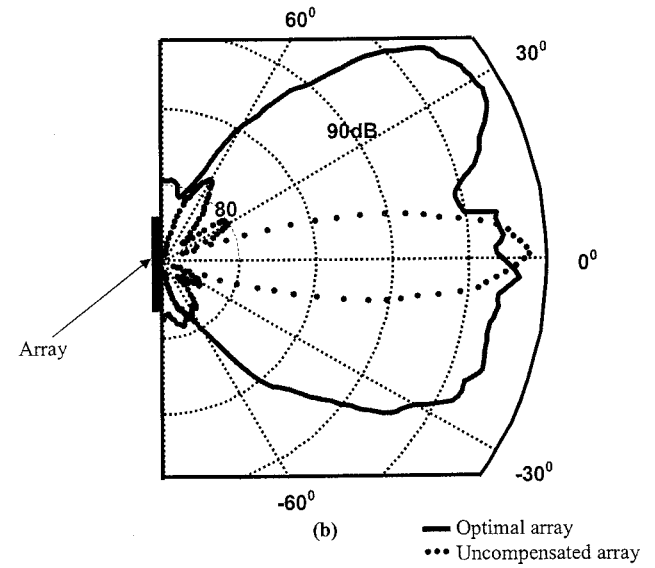
B. The experimental result of a 5×1 linear panel speaker array

Although the ultimate goal of this work was to develop the large panel speaker array, we use a small array to verify the far-field behavior because of the limitation of current measuring environment. A 5×1 panel speaker array is constructed for experimental verification. The system consists of using PU-foam panel speakers, array signal-processing unit, a monitoring microphone, data acquisition unit, and a stepping motor unit. The dimensions and structure of the 5×1 panel speakers are shown in Figs. 8(a) and (b). The size of each rectangular panel is $7\times 6.7\ \text{cm}^2$ and the spacing between adjacent speakers, $d=6.7$ cm. Each panel is driven by an electromagnetic exciter mounted on an aluminum frame. The photo of the array hardware is shown in Fig. 8(c).

The array signal processing is carried out by a floating-point DSP, TSM320C32 in conjunction with a multiple-channel IO module. Audio signals are fed to the array signal-processing unit, through AD conversion and power amplifi-



(a)



(b)

FIG. 9. The experimental results of directional response for the 5×1 panel speaker array at (a) 2 kHz and (b) 3 kHz.

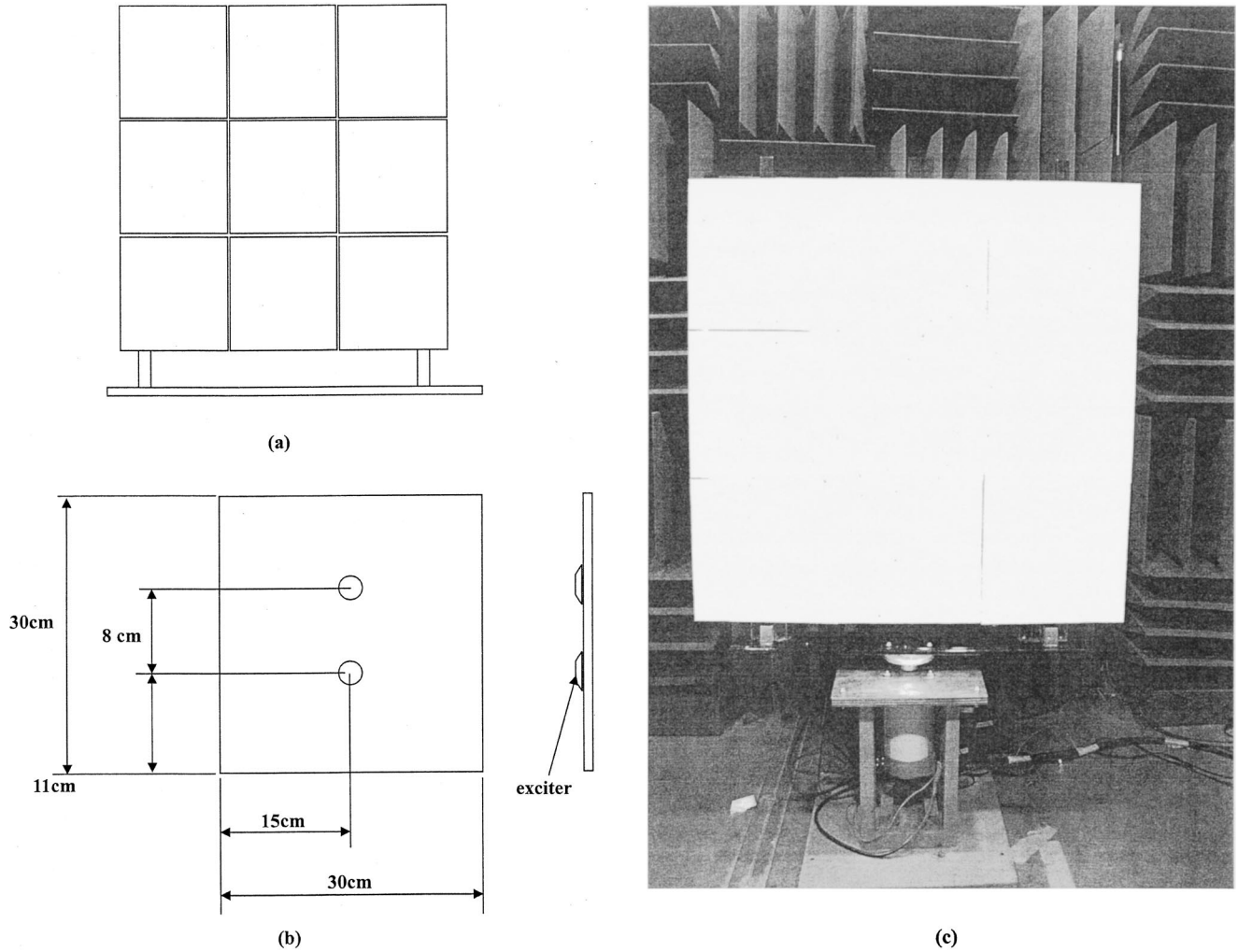


FIG. 10. The large panel speaker array of a projection screen. (a) Configuration of the 3×3 panel speaker array; (b) dimensions of panel speaker and the location of the exciter; (c) The photo of the 3×3 panel speaker array of a projection screen.

cation, and generate compensated signals for each channel of speaker. The monitoring microphone is situated 2 m from the center line of the array. The array is mounted on a turntable driven by the stepping motor so that directional responses of the array within $[-90°, 90°]$ can be recorded automatically, with every 1° increment. Necessary data acquisition/processing and motor control are all handled by the DSP as well. The experiments are conducted inside an anechoic chamber.

An experiment was conducted to compare the optimal panel speaker array to an uncompensated array ($g_n = 1, \forall n$). The optimal set of coefficients of the 5×1 array was found using the proposed optimization technique

$$\mathbf{g}_{\text{opt}} = \{-0.45, j, 1, j, -0.45\}. \tag{38}$$

The Hilbert transformer was implemented by a 100-tap FIR filter with a 50-sample delay, and sampling rate is 20 kHz. The directional response of the panel speaker array is measured at 2 and 3 kHz, respectively, on the horizontal plane of the array. Drastic differences can be observed in the experimental results of Fig. 9. The array is indicated in the figure. The uncompensated array indeed radiates a rather directional pattern, whereas the optimal array exhibits a relatively omnidirectional behavior.

C. The experimental result of a 3×3 panel speaker matrix

To further justify the panel speaker array, a large 3×3 matrix of the size of a projection screen is constructed using PU-foam panels, covered with glassy face skins. The dimensions and structure of the 3×3 panel speakers are shown in Fig. 10(a). The size of each panel is $30 \times 30\ \text{cm}^2$ and the spacing between adjacent speakers, $d = 30$ cm. The photo of the array hardware is shown in Fig. 10(b). Despite the matrix configuration, the array is based on one-dimensional compensation in the horizontal direction, which in our application is considered more important than the vertical direction. Each panel is driven by two electromagnetic exciters mounted on an aluminum frame. The three panels at each column are wired together to the same DSP output and six exciters are all in parallel connection. The rest of the details of experimental arrangement are identical to those of the 5×1 array.

An experiment was conducted to compare the optimal panel speaker array to an uncompensated array. The optimal

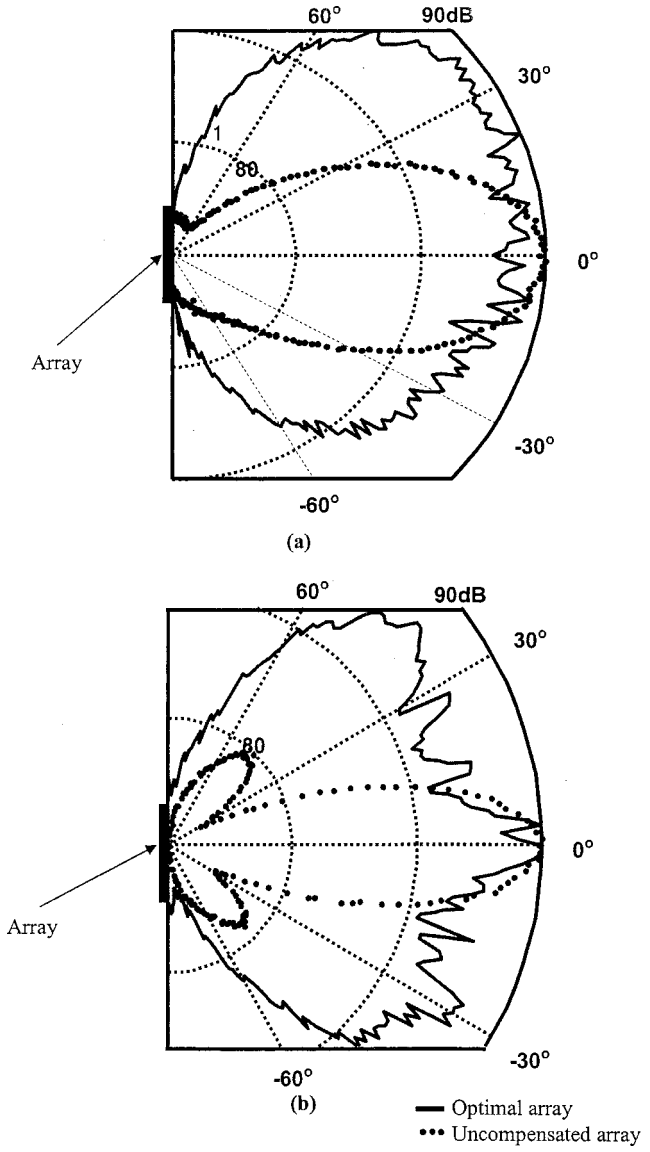


FIG. 11. The experimental results of directional response for the 3×3 panel speaker array at (a) 488 Hz and (b) 780 Hz.

set of coefficients of the 3×3 array was found using the proposed optimization technique

$$\mathbf{g}_{\text{opt}}=\{-0.45j,1,-045j\}. \tag{39}$$

The Hilbert transformer was implemented by a 150-tap FIR filter with a 75-sample delay, and sampling rate is 10 kHz. The directional responses of the panel speaker array is measured at 488 and 780 Hz, respectively, on the horizontal plane of the array. It is noted that in this large array experiment, we were able to investigate only near-field radiation pattern in our anechoic room. Similar to the case of small array, drastic difference can be observed in the experimental results of Fig. 11. The uncompensated array radiates a quite directive pattern, whereas the optimal array exhibits a relatively omnidirectional behavior.

VI. CONCLUSIONS

An important feature of this paper is the array optimization method, and the projection screen is a practical example of its use. Electronic compensation is employed to achieve omnidirectional response and array efficiency. Optimal design of array coefficients is computed by a three-stage optimization procedure that effectively solves the nonlinear and nonconvex problem. The process is interactive, allowing us to tailor the array coefficients to meet the design specifications. The array design obtained using the optimization technique has been implemented by using a multichannel DSP, where a Hilbert transformer is required to produce the quadrature components of the array coefficients. A small array and a large matrix were constructed to validate the implemented array signal-processing system. The experimental results obtained from the DSP-based system indicate that the optimally compensated panel speaker array exhibits omnidirectional radiation pattern without degradation of efficiency.

With regard to the use of panel speakers as projection screens, there are a few technical points to consider. These include the added brightness that can be achieved with a nonperforated screen, as well as the degree to which the response of large speaker arrays suffers from time smearing and poor stereo imaging.

Although the ultimate goal of this work was to develop the large array, we were unable to verify its far-field behavior due to the limitation of current measuring environment. Much work is continuing in improving the implementation as well as measurement of the large array for future research.

ACKNOWLEDGMENTS

Special thanks are due to the illuminating discussions with *NXT, New Transducers Limited*, UK. The work was supported by the National Science Council in Taiwan, Republic of China, under the project number NSC 89AFA06000714.

[1] M. R. Bai and T. Huang, "Development of Panel Loudspeaker System: Design, Evaluation and Enhancement," J. Acoust. Soc. Am. **109**, 2751–2761 (2001).
[2] L. E. Kinsler, A. R. Ffrey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982).
[3] D. L. Smith, "Discrete-Element Line Arrays—Their Modeling and Optimization," J. Audio Eng. Soc. **45**, 949–964 (1997).
[4] G. L. Augspurger, "Near-Field and Far-Field Performance of Large Woofer Arrays," J. Audio Eng. Soc. **38**, 231–236 (1990).
[5] D. G. Meyer, "Digital Control of Loudspeaker Array Directivity," J. Audio Eng. Soc. **32**, 747–754 (1984).
[6] R. M. Aarts and A. J. E. M. Janssen, "On Analytic Design of Loudspeaker Arrays with Uniform Radiation Characteristics," J. Acoust. Soc. Am. **107**, 287–292 (2000).
[7] *Sound Advance*, http://soundadvance.com/.
[8] G. F. M. Beenker, T. A. C. M. Claasen, and P. W. C. Hermens, "Binary Sequences with a Maximally Flat Amplitude Spectrum," Philips J. Res. **40**, 289–304 (1985).
[9] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1989).
[10] Z. G. Jing, "A New Method for Digital All-Pass Filter Design," IEEE Trans. Acoust., Speech, Signal Process. **35**, 1557–1564 (1987).
[11] B. Gold, A. V. Oppenheim, and C. M. Radar, "Theory and Implementation of the Discrete Hilbert Transformer," in *Proc. Symp. Computer Processing in Communications*, Vol. 19 (Polytechnic, New York, 1970).
[12] D. F. Johnson and D. F. Dudgeon, *Array Signal Processing Concepts and Techniques* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
[13] M. J. E. Golay, "The Merit Factor of Long, Low Autocorrelation Binary Sequences," IEEE Trans. Inf. Theory **28**, 543 (1982).
[14] J. S. Arora, *Introduction to Optimum Design* (McGraw-Hill, Singapore, 1989).

Analysis of axially polarized piezoelectric ceramic cylindrical shells of finite length with internal losses

D. D. Ebenezer[a] and Pushpa Abraham
Naval Physical and Oceanographic Laboratory, Kochi 682 021, India

(Received 30 November 2001; revised 13 July 2002; accepted 22 July 2002)

A thin shell analytical model of axially polarized piezoelectric ceramic cylinders with internal losses is presented. The Flugge assumptions for strain–displacement relations, Hamilton's principle extended to piezoelectric shells, and the assumption that electric potential has a quadratic variation between the curved surfaces, are used to derive displacement–potential relations that are similar to equations of motion of elastic shells. A solution, with 12 coefficients, to these relations is then derived. The coefficients are complex when the shell has internal losses and are determined by using three mechanical and three electrical boundary conditions at each end—on the flat surfaces. Computed values of input electrical admittance are presented for shells with and without internal losses, and for thin shells as well as shells with wall thickness comparable to the length. They are also compared with results obtained using the finite element program—ATILA. It is shown that the analytical values of resonance frequencies, the maximum value of input electrical conductance, and the maximum and minimum values of input electrical susceptance of thin shells are in excellent agreement with finite element results. The dependence of the maxima and minima in the complex input electrical admittance on the dimensions of the shell is inferred from the numerical results.
 [DOI: 10.1121/1.1506685]

PACS numbers: 43.38.Ar, 43.40.Ey [SLE]

I. INTRODUCTION

Axially polarized piezoelectric ceramic hollow cylinders are often used in underwater electroacoustic Tonpilz[1,2] projectors. In most cases, the wall thickness is comparable to the length, and the radius is not much greater than the length. Therefore, they can neither be classified as thin shells nor as thin disks. Even when the radius is much greater than the length, the resonance frequencies of the cylinder that are strongly dependent on the length are the ones of interest and these cannot be determined by modeling the cylinder as a thin disk. Therefore, the cylinder is modeled as a shell.

Studies on piezoelectric shells can be classified into two areas:[3] derivation of a new shell theory and solutions to the equations of motions and boundary conditions. The extensive reviews of literature in the works of Dokmeci[3,4] and Rogacheva[5] show that there are several studies in the first area until 1990 and there have been more later.[6] In the second area, only very few solutions have been presented for thin shell models. Drumheller and Kalnins,[7] Rogacheva,[5] and Ebenezer and Abraham[8] presented solutions for cylinders of finite length in which the bending stiffness of the shell was considered. They solved the governing equations by using a numerical method, an approximate method, and eigenfunction analysis, respectively. Numerical results were presented in Refs. 7 and 8 for no-loss cylinders. The results for thin shells were compared with membrane–solution results and experimental data in Ref. 8. Internal losses have been included in a membrane model[9] but not in the thin shell models.

In this article, solutions are presented for electrically excited, thin, axially polarized piezoelectric shells of finite length with internal losses. For completeness, governing equations for a thin shell are first derived by using the Flugge approximations, assuming that the electric potential has a quadratic variation between the curved surfaces and using Hamilton's principle extended to piezoelectric media. Then, a solution to the displacement–potential equations is presented. This is possible because the electrodes completely cover the flat ends of the shell. Numerical values of the input electrical admittance of free–free shells with and without internal losses are also presented. They are compared with values obtained using ATILA[10]—a finite element program.

II. THEORY

Consider a thin axially polarized piezoelectric ceramic cylindrical shell of length L, mean radius a, and thickness h where h is much less than the length and radius, as shown in Fig. 1. The shell has internal losses. The displacement, stress and electric field in the shell, and the input electrical admittance (IEA) of the shell are of interest when a voltage is applied across the electrodes that completely cover the flat surfaces.

The approximations made by Flugge[11] for thin elastic shells are also made here. Therefore, the strain–displacement relations of a thin piezoelectric cylindrical shell are expressed in cylindrical coordinates as

$$S_\theta(r,z,\omega) = W(a,z,\omega)/(a+x) \tag{1a}$$

and

$$S_z(r,z,\omega) = \partial U(a,z,\omega)/\partial z - x\ \partial^2 W(a,z,\omega)/\partial z^2, \tag{1b}$$

where U and W are the axial and radial displacements, respectively, at $r=a$, $x=r-a$, S_θ and S_z are the normal

[a]Electronic mail: tsonpol@vsnl.com

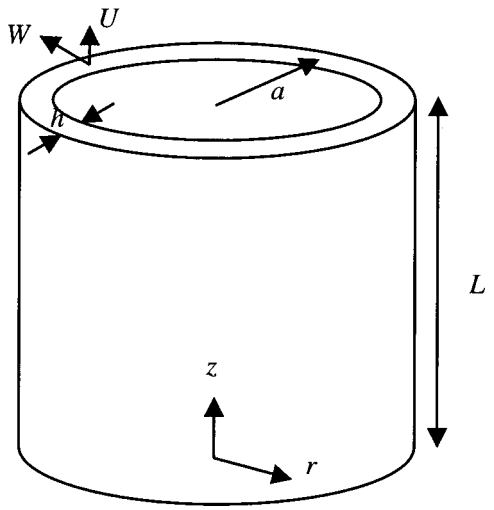


FIG. 1. A thin axially polarized piezoelectric ceramic cylindrical shell of length L, radius a, and thickness h. The top and bottom flat surfaces are electroded.

strains in the θ and z directions, respectively, and ω denotes angular frequency. When the membrane approximation or the Kirchhoff–Love hypothesis is used, $a+x$ in the denominator of (1a) is approximately expressed as a. The term containing x in (1b) is used to account for bending strain. It is neglected when using the membrane approximation but retained when the Kirchhoff–Love hypothesis is used. $\exp(j\omega t)$ has suppressed in all the equations for convenience.

Because the shell is thin, and the Flugge approximations are being used, the stresses that are zero on the curved surfaces are assumed to be zero everywhere within the cylinder. Therefore, the piezoelectric equations of state are expressed as

$$\begin{Bmatrix} T_\theta(r,z,\omega) \\ T_z(r,z,\omega) \\ D_z(r,z,\omega) \\ D_r(r,z,\omega) \end{Bmatrix} = [\mathbf{C}] \begin{Bmatrix} S_\theta(r,z,\omega) \\ S_z(r,z,\omega) \\ \partial\Phi(r,z,\omega)/\partial z \\ \partial\Phi(r,z,\omega)/\partial r \end{Bmatrix}, \tag{2a}$$

where

$$[C]^{-1} = \begin{bmatrix} s_{11}^D & s_{13}^D & g_{31} & 0 \\ s_{13}^D & s_{33}^D & g_{33} & 0 \\ g_{31} & g_{33} & -\beta_{33}^T & 0 \\ 0 & 0 & 0 & -\beta_{11}^T \end{bmatrix} \tag{2b}$$

is a symmetric matrix, the electroelastic coefficients s_{11}^D, s_{13}^D, s_{33}^D, g_{31}, g_{33}, β_{33}^T and β_{11}^T have their usual definitions, T_θ and T_z are the normal stresses in the θ and z directions, respectively, D_z and D_r are the components of electric displacement in the axial and radial directions respectively, and Φ is electric potential. When internal losses are present, each of the coefficients in (2b) is complex and must satisfy the conditions derived by Holland.[12]

The electric potential is approximately expressed as

$$\Phi(r,z,\omega) = \Phi_0(z,\omega) + x\Phi_1(z,\omega) + x^2\Phi_2(z,\omega), \tag{3a}$$

where

$$\Phi_1(z,\omega) = \Phi_2(z,\omega) = 0, \quad \text{at } z=0 \text{ and } L \tag{3b}$$

because Φ is independent of r on the equipotential flat surfaces. The electric field is now expressed as

$$E_z(r,z,\omega) = -\partial\Phi(r,z,\omega)/\partial z \tag{4a}$$

and

$$E_r(r,z,\omega) = -\partial\Phi(r,z,\omega)/\partial r. \tag{4b}$$

The electric potential across the electrodes is expressed as

$$V(\omega) = -\int_0^L \frac{\partial\Phi(r,z,\omega)}{\partial z} dz = \Phi(r,0,\omega) - \Phi(r,L,\omega), \tag{5a}$$

and the applied potentials are

$$\Phi(r,z,\omega) = 0, \quad \text{at } z=0 \tag{5b}$$

and

$$\Phi(r,z,\omega) = V(\omega), \quad \text{at } z=L, \tag{5c}$$

i.e., the voltage applied across the electrodes is $V(\omega)$.

After defining a Fourier transform pair,

$$\hat{F}(t) \leftrightarrow F(\omega), \tag{6}$$

where t denotes time and $\hat{}$ denotes a time-dependent function, Hamilton's variational principle,[13]

$$\delta\int_{t_1}^{t_2} [\hat{\mathsf{T}}(t) - \hat{\mathsf{V}}(t) + \hat{\mathsf{E}}(t)]dt = 0, \tag{7a}$$

where δ denotes variation, t_1 and t_2 are an initial time and a final time, respectively, at which the displacements are fixed,

$$\hat{\mathsf{T}}(t) = 2\rho\pi ah\int_{t_1}^{t_2}\int_0^L (\hat{U}^2(a,z,t) + \hat{W}^2(a,z,t))dz\ dt \tag{7b}$$

is the total kinetic energy of the shell when rotary inertia is neglected,[14]

$$\hat{\mathsf{V}}(t) = \pi\int_{t_1}^{t_2}\int_0^L\int_{-h/2}^{h/2} (\hat{T}_z(r,z,t)\hat{S}_z(r,z,t) + \hat{T}_\theta(r,z,t)\hat{S}_\theta(r,z,t))(a+x)dx\ dz\ dt \tag{7c}$$

is the total work energy, and

$$\hat{\mathsf{E}}(t) = \pi\int_{t_1}^{t_2}\int_0^L\int_{-h/2}^{h/2} (\hat{D}_z(r,z,t)\hat{E}_z(r,z,t) + \hat{D}_r(r,z,t)\hat{E}_r(r,z,t))(a+x)dx\ dz\ dt \tag{7d}$$

is the total electric energy, is used to derive the displacement–potential relations and the boundary condi-

tions. It can be shown[8] by using the Fourier transforms of Eqs. (1)–(4) in Eq. (7) and integrating by parts that the resulting equation can be expressed as the sum of double integrals and single integrals.

The displacement–potential relations are obtained by equating the coefficients of δu, δw, $\delta\phi_0$, $\delta\phi_1$, and $\partial\delta w/\partial z$ in the double integrals to zero. They are expressed after transforming to frequency domain as

$$[\Lambda]\begin{Bmatrix} U(a,z,\omega) \\ W(a,z,\omega) \\ \Phi_0(z,\psi) \\ \Phi_1(z,\omega) \\ \Phi_2(z,\omega) \end{Bmatrix} - \omega^2\rho\begin{Bmatrix} U(a,z,\omega) \\ W(a,z,\omega) \\ 0 \\ 0 \\ 0 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix}, \tag{8a}$$

where

$$[\Lambda] = \begin{bmatrix} -c_{22}\frac{\partial^2}{\partial z^2} & -\frac{c_{21}}{a}\frac{\partial}{\partial z} + \frac{\beta}{a}c_{22}\frac{\partial^3}{\partial z^3} & -c_{23}\frac{\partial^2}{\partial z^2} & -\frac{c_{23}\beta}{a}\frac{\partial^2}{\partial z^2} & -c_{32}\beta\frac{\partial^2}{\partial z^2} \\ \frac{c_{12}}{a}\frac{\partial}{\partial z} - \frac{c_{22}\beta}{a}\frac{\partial^3}{\partial z^3} & \frac{c_{11}}{ah}\mu + c_{22}\beta\frac{\partial^4}{\partial z^4} & -\frac{c_{23}\beta}{a}\frac{\partial^3}{\partial z^3} + \frac{c_{31}}{a}\frac{\partial}{\partial z} & -c_{32}\beta\frac{\partial^3}{\partial z^3} & \frac{c_{13}\beta}{a}\frac{\partial}{\partial z} - \frac{c_{23}h^4}{80a}\frac{\partial^3}{\partial z^3} \\ c_{32}\frac{\partial^2}{\partial z^2} & \frac{c_{31}}{a}\frac{\partial}{\partial z} - \frac{c_{32}\beta}{a}\frac{\partial^3}{\partial z^3} & c_{33}\frac{\partial^2}{\partial z^2} & \frac{c_{33}\beta}{a}\frac{\partial^2}{\partial z^2} & c_{33}\beta\frac{\partial^2}{\partial z^2} \\ \frac{c_{23}\beta}{a}\frac{\partial^2}{\partial z^2} & -c_{32}\beta\frac{\partial^3}{\partial z^3} & \frac{c_{33}}{a}\beta\frac{\partial^2}{\partial z^2} & c_{33}\beta\frac{\partial^2}{\partial z^2} - c_{44} & \frac{c_{33}h^4}{80a}\frac{\partial^2}{\partial z^2} - \frac{c_{44}h^2}{6a} \\ c_{32}\beta\frac{\partial^2}{\partial z^2} & \frac{c_{31}}{a}\beta\frac{\partial}{\partial z} - \frac{c_{32}h^4}{80a}\frac{\partial^3}{\partial z^3} & c_{33}\beta\frac{\partial^2}{\partial z^2} & \frac{c_{33}h^4}{80a}\frac{\partial^2}{\partial z^2} - \frac{c_{44}h^2}{6a} & \frac{c_{33}h^4}{80}\frac{\partial^2}{\partial z^2} - \frac{c_{44}h^2}{3} \end{bmatrix}, \tag{8b}$$

$$\mu = \ln\frac{a+h/2}{a-h/2}, \tag{8c}$$

and

$$\beta = \frac{h^2}{12}. \tag{8d}$$

The mechanical boundary conditions for a free–free shell are obtained by equating the transforms of the coefficients of δu, δw, and $\partial\delta w/\partial z$ in the single integrals to zero. They can also be obtained by using Eqs. (1) and (2) and

$$2\pi\int_{-h/2}^{h/2}(a+x)T_z(r,z,\omega)dx = 0, \quad \text{at } z=0,\ L, \tag{9a}$$

$$2\pi\frac{\partial}{\partial z}\int_{-h/2}^{h/2}(a+x)xT_z(r,z,\omega)dx = 0, \quad \text{at } z=0,\ L, \tag{9b}$$

and

$$2\pi\int_{-h/2}^{h/2}(a+x)xT_z(r,z,\omega)dx = 0, \quad \text{at } z=0,\ L. \tag{9c}$$

Equation (9) indicates that the axial force, shear force, and bending moment are zero at the ends of the shell. After using Eqs. (1)–(4) in Eq. (9) and integrating, the resulting equations are expressed as

$$[B]\begin{Bmatrix} U(a,z,\omega) \\ W(a,z,\omega) \\ \Phi_0(z,\omega) \\ \Phi_1(z,\omega) \\ \Phi_2(z,\omega) \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix}, \tag{10a}$$

where

TABLE I. Properties of the piezoelectric material used in the computations.

Material property	Value
ρ (kg/m^3)	7500
s_{11}^E (m^2/N)$\times 10^{-11}$	1.56−j0.031
s_{12}^E	−0.42+j0.012
s_{13}^E	−0.825+j0.017
s_{33}^E	1.89−j0.034
s_{44}^E	392−j0.13
d_{31} (C/N)$\times 10^{-12}$	−295+j9.9
d_{33}	564−j15
d_{15}	560−j30
ϵ_{11}^T (F/m)$\times 10^{-9}$	21.4−j1.3
ϵ_{33}^T	30.6−j1.1

TABLE II. Computed resonance and antiresonance frequencies of shells with no internal losses.

Dimensions (L, a, and h in mm)	Branch	Model	Resonance frequency (kHz)	Antiresonance frequency (kHz)
10,10,1	Lower	Flugge	45.96, 59.96, 178.5	51.36
		ATILA	45.98, 58.49, 150.9	51.40
	Upper	Flugge	160.3, 604.2	206.2, 617
		ATILA	160.0, 596.2	205.9, 609
10,10,2	Lower	Flugge	46.03, 84.1	51.41
		ATILA	46.09, 72.7	51.57
	Upper	Flugge	160.5, 605.5	206.4, 618
		ATILA	159.2, 563.7	205.1, 577

$$[B]=\begin{bmatrix} c_{22}\dfrac{\partial}{\partial z} & \dfrac{c_{12}}{a}-\dfrac{c_{22}\beta}{a}\dfrac{\partial^2}{\partial z^2} & c_{23}\dfrac{\partial}{\partial z} & \beta+\dfrac{c_{23}}{a}\dfrac{\partial}{\partial z} & c_{23}\beta\dfrac{\partial}{\partial z} \\ \dfrac{c_{22}\beta}{a}\dfrac{\partial^2}{\partial z^2} & -c_{22}\beta\dfrac{\partial^3}{\partial z^3} & \dfrac{c_{23}\beta}{a}\dfrac{\partial^2}{\partial z^2} & c_{23}\beta\dfrac{\partial^2}{\partial z^2} & \dfrac{6c_{23}\beta^2}{a}\dfrac{\partial^2}{\partial z^2} \\ \dfrac{c_{22}\beta}{a}\dfrac{h^2}{12}\dfrac{\partial}{\partial z} & -c_{22}\beta\dfrac{\partial^2}{\partial z^2} & \dfrac{c_{23}\beta}{a}\dfrac{\partial}{\partial z} & c_{23}\beta\dfrac{\partial}{\partial z} & \dfrac{6c_{23}\beta^2}{a}\dfrac{\partial}{\partial z} \end{bmatrix}. \tag{10b}$$

It can be shown that the third, fourth, and fifth equations in the displacement–potential relations in Eq. (8) are together equivalent to the electrostatic condition:

$$\frac{\partial D_z(r,z,\omega)}{\partial z}+\frac{\partial D_r(r,z,\omega)}{r\ \partial r}=0. \tag{11}$$

The general solution to the displacement–potential relations in Eq. (8) is expressed in the form

$$\begin{Bmatrix} U(a,z,\omega) \\ W(a,z,\omega) \\ \Phi_0(z,\omega) \\ \Phi_1(z,\omega) \\ \Phi_2(z,\omega) \end{Bmatrix} = V(\omega)\begin{Bmatrix} \sum_{s=1}^{10} A_s(\omega)e^{\lambda_s(\omega)z} \\ \sum_{s=1}^{10} A_s(\omega)\psi_s e^{\lambda_s(\omega)z}+K_1(\omega) \\ \sum_{s=1}^{10} A_s(\omega)\xi_s e^{\lambda_s(\omega)z}+\alpha K_1(\omega)z+K_2(\omega) \\ \sum_{s=1}^{10} A_s(\omega)\gamma_s e^{\lambda_s(\omega)z} \\ \sum_{s=1}^{10} A_s(\omega)\chi_s e^{\lambda_s(\omega)z} \end{Bmatrix}, \tag{12a}$$

where it is seen by substituting Eq. (12a) in the second Eq. in (8) and equating the coefficient of the z-independent term to zero that

$$\alpha=\left(\rho\omega^2-\frac{c_{11}\mu}{ah}\right)\frac{a}{c_{31}}, \tag{12b}$$

$\lambda_s, s=1,\ldots 10$ are the solutions to the characteristic equation obtained by dropping the summation in Eq. (12), substituting it in Eq. (8), and equating the determinant of the resulting set of homogeneous equations to zero and numerically solving the tenth order characteristic equation. The values of ψ_x, ξ_s, γ_s, and χ_s, $s=1,\ldots 10$ are then determined by using the known values of $\lambda_S(\omega)$ and the homogeneous set of equations obtained above.

The coefficients A_s and K_i, $i=1,2$, are now determined by using the boundary conditions. After substituting the expressions for the displacements and potential in Eq. (12) in the six mechanical boundary conditions in Eq. (8) and the six electrical boundary conditions in (3b) and (5), the real and imaginary parts of the resulting equations are separately equated to zero. Twenty-four equations are obtained. These are then used to find the real and imaginary parts of A_s and K_i.

All quantities of interest are now obtained easily. For example, the input electrical admittance is expressed, by using Eqs. (1)–(6) and (12a), as

$$\begin{aligned} Y(\omega)=\frac{j\omega 2\pi}{V(\omega)}\int_{-h/2}^{h/2}(a+x)D_z(r,z,\omega)dx=\frac{j\omega 2\pi ah}{V(\omega)} \\ \times\left[c_{32}\frac{\partial U(a,z,\omega)}{\partial z}+c_{31}\frac{W(a,z,\omega)}{a}\right. \\ -c_{32}\frac{\beta}{a}\frac{\partial^2 W(a,z,\omega)}{\partial z^2}+c_{33}\left(\frac{\partial\Phi_0(z,\omega)}{\partial z}\right. \\ \left.\left.+\frac{\beta}{a}\frac{\partial\Phi_1(z,\omega)}{\partial z}+\beta\frac{\partial\Phi_2(z,\omega)}{\partial z}\right)\right]. \end{aligned} \tag{13a}$$

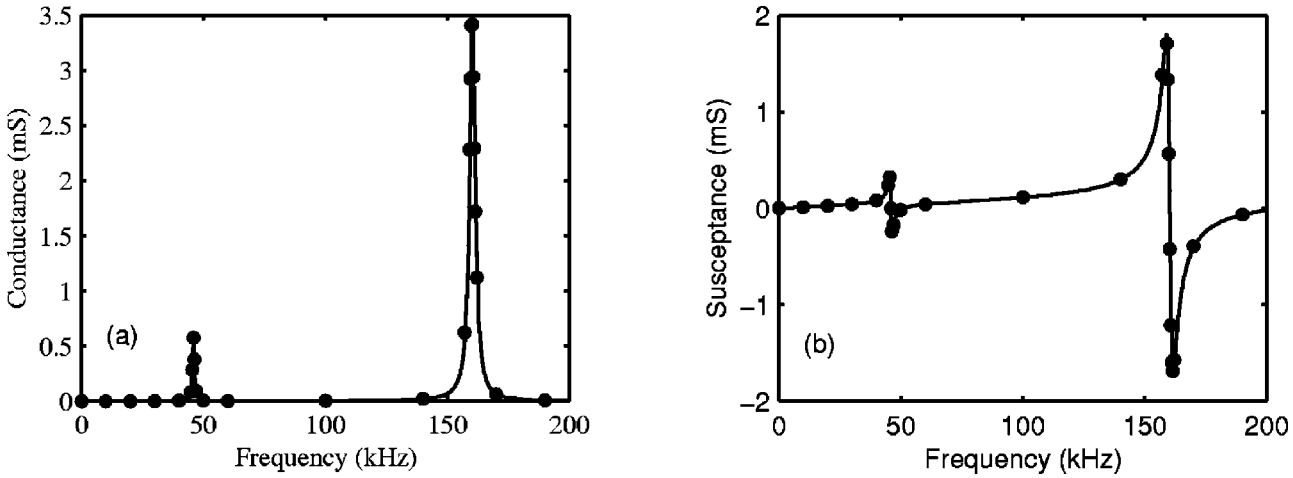


FIG. 2. The input electrical admittance of a shell. $L=10$ mm, $a=10$ mm, and $h=1$ mm. Solid line: present model. Dots: finite element model (ATILA). (a) Conductance and (b) susceptance.

It appears from Eq. (13a) that $Y(\omega)$ is dependent on z. However, it is seen by using the third equation in (8b) and Eq. (12a) in Eq. (13a) that

$$Y(\omega)=j\omega 2\pi a h\left(\frac{c_{31}}{a}+c_{33}\alpha\right)K_1(\omega), \tag{13b}$$

is independent of z. The expressions for displacements and potential are sums of a finite number of terms but the expression for input electrical admittance contains only one of the 12 coefficients.

III. NUMERICAL RESULTS AND DISCUSSION

Numerical results obtained using the Flugge model are presented for shells with and without internal losses. They are also compared with those obtained using the general-purpose finite-element program ATILA.[10] Second-order, isoparametric, eight-noded, quadrilateral, axisymmetric, piezoelectric elements are used in ATILA with internal losses represented by complex coefficients. The electroelastic coefficients and density of the piezoceramic are shown in Table I and can be found in Ref. 15.

First, consider a shell without internal losses whose co-

TABLE III. Computed behavior near resonance frequencies of shells with internal losses.

Case	Dimensions (L, a, and h in nm)	Branch	Model	Conductance Frequency in kHz and maximum G in mS	Susceptance		
					Frequency in kHz and maximum B in mS	Frequency in kHz and minimum B in mS	Frequency in kHz at which B is zero
A	10,10,1	Lower	Flugge	45.95(0.58)	45.49(0.32)	46.40(−0.25)	51.40
			ATILA	45.97(0.58)	45.51(0.32)	46.42(−0.25)	51.37
		Upper	Flugge	160.31(3.50)	158.92(1.80)	161.71(−1.70)	206.20
			ATILA	160.0(3.48)	158.62(1.79)	161.38(−1.69)	205.95
B	10,10,2	Lower	Flugge	46.01(1.2)	45.55(0.64)	46.44(−0.50)	51.38
			ATILA	46.08(1.17)	46.62(0.65)	46.5(−0.51)	51.54
		Upper	Flugge	160.48(7.07)	159.30(3.6)	161.83(−3.4)	206.35
			ATILA	159.14(6.92)	157.75(3.57)	160.52(−3.35)	205.94
C	5,5,1	Lower	Flugge	92.03(1.20)	91.09(0.64)	92.91(−0.50)	102.80
			ATILA	92.16(1.17)	91.23(0.65)	93.06(−0.51)	103.09
		Upper	Flugge	320.99(7.0)	318.24(3.60)	323.71(−3.40)	412.70
			ATILA	318.28(6.92)	315.50(3.56)	320.70(−3.32)	410.24
D	5,10,1	Lower	Flugge	46.38(1.10)	45.95(0.60)	46.83(−0.45)	51.41
			ATILA	46.43(1.07)	45.95(0.60)	46.87(−0.45)	51.46
		Upper	Flugge	317.45(14.70)	314.74(7.60)	320.32(−7.10)	411.53
			ATILA	314.74(14.5)	312.0(7.46)	317.5(−7.05)	409.30
E	5,10,2	Lower	Flugge	46.46(2.14)	45.99(1.20)	46.91(−0.91)	51.47
			ATILA	46.55(2.16)	46.10(1.09)	47.00(−1.06)	51.54
		Upper	Flugge	319.20(28.70)	316.40(14.80)	321.92(−13.90)	412.71
			ATILA	306.40(28.3)	303.60(14.0)	309.20(13.80)	401.87
F	5,10,5	Lower	Flugge	46.87(5.30)	46.40(3.0)	47.33(−2.30)	51.85
			ATILA	47.35(5.90)	46.88(3.30)	47.81(−2.59)	52.90
		Upper	Flugge	312.49(69.10)	310(35.50)	315.17(−33.40)	398.43
			ATILA	243.24(40.21)	240.8(22.2)	245.60(−16.6)	369.67

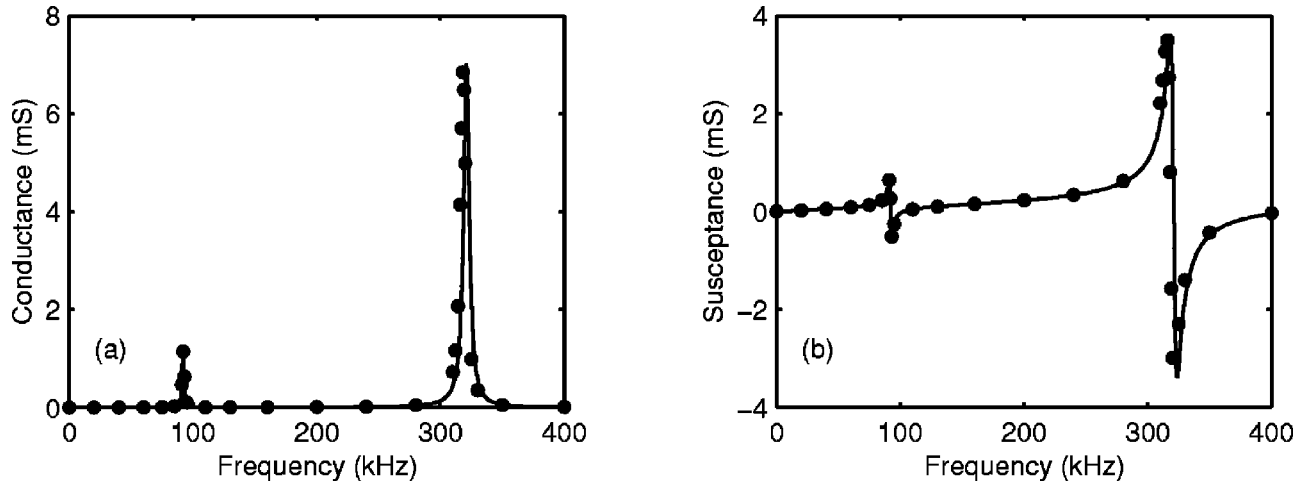


FIG. 3. The input electrical admittance of a shell. $L=5$ mm, $a=5$ mm, and $h=1$ mm. Solid line: present model. Dots: finite element model (ATILA). (a) Conductance and (b) susceptance.

efficients and density are the real parts of the values shown in Table I. The coefficients in Eq. (2b) are computed using the coefficients shown in Table I. The resonance and antiresonance frequencies obtained using the present Flugge and ATILA models are shown in Table II for two shells with $L=10$ mm and $a=10$ mm. The thickness is $h=1$ and 2 mm, respectively. The shells are modeled in ATILA by using 40 elements along the axis and 8 elements in the radial direction, i.e., 40×8 elements. Identical results are obtained by using 30×6 elements.

The frequency at which there is a discontinuity in the susceptance and the frequency at which the susceptance is zero are the resonance and antiresonance frequencies, respectively. It is seen from Table II that the first resonance and antiresonance frequencies in the lower and upper branches computed using the present model and ATILA are nearly the same. When the thickness increases the present model and ATILA indicate that the first resonance frequency in the lower branch increases. However, the present model indicates that the first resonance frequency in the upper branch increases slightly, whereas ATILA shows that there will be small decrease. The higher-order resonance frequencies in the lower and upper branches computed using the present model are greater than those obtained using ATILA. The difference increases as the thickness of the shell increases. The first antiresonance frequencies in the lower and upper branches are also in good agreement. When only the first resonance and antiresonance frequencies in the lower and upper branches are considered, the difference is less than 0.2% for $h=1$ mm and less than 0.9% for $h=2$ mm. Therefore, the present model can be used to find the electroelastic coefficients of thin shells.

The complex input electrical admittance (IEA) of a shell with $L=10$ mm, $a=10$ mm, and $h=1$ mm computed using the complex coefficients in Table I is shown in Fig. 2. A solid line and dots are used to show the values obtained using the present Flugge model and ATILA, respectively. Only the first resonance frequencies in the lower and upper branches can be seen. The maximum values of conductance occur at 45.95 kHz (0.58 mS) and 160.31 kHz (3.50 mS), respectively, when using the Flugge model and at 45.97 kHz (0.58 mS) and 160.0 kHz (3.48 mS) when using ATILA and are also shown in case A of Table III. The resonance and antiresonance frequencies, the maximum value of conductance, the

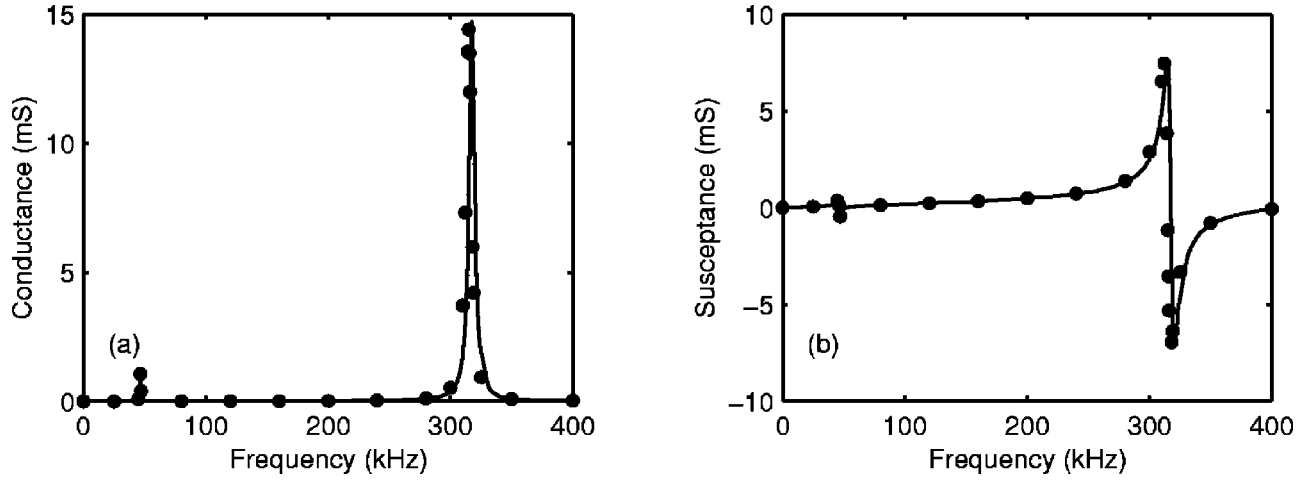


FIG. 4. The input electrical admittance of a shell. $L=5$ mm, $a=10$ mm, and $h=1$ mm. Solid line: present model. Dots: finite element model (ATILA). (a) Conductance and (b) susceptance.

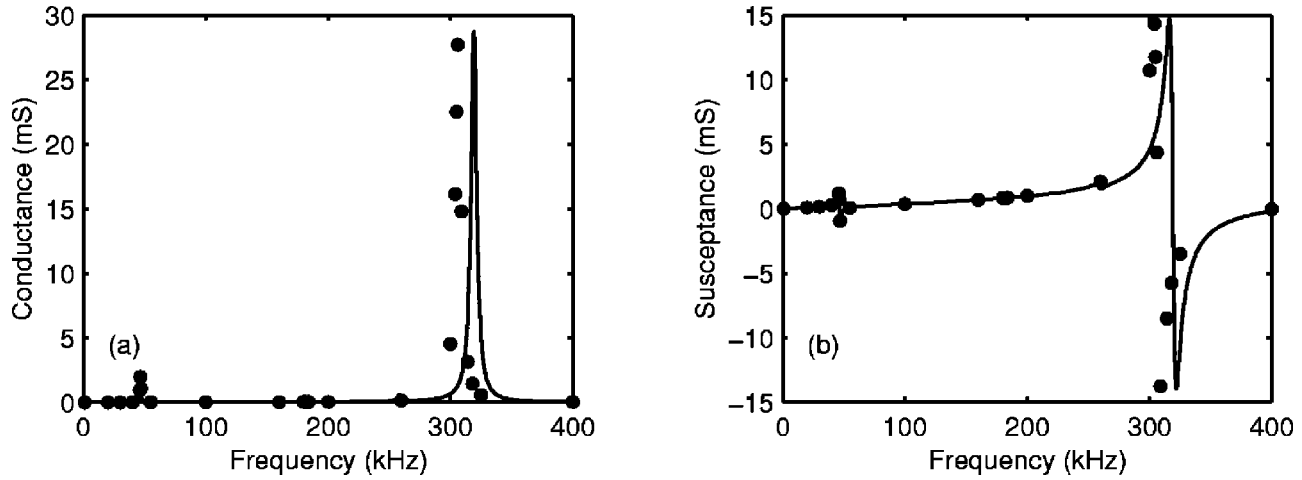


FIG. 5. The input electrical admittance of a shell. $L=5$ mm, $a=10$ mm, and $h=2$ mm. Solid line: present model. Dots: finite element model (ATILA). (a) Conductance and (b) susceptance.

maximum and minimum value of susceptance, and the frequencies at which they occur are all in good agreement with the values obtained using ATILA. It is seen by comparing these resonance frequencies with those shown in Table II that the resonance frequencies for shells with and without losses are not exactly the same. The IEA computed using ATILA is shown only at selected frequencies to make the comparison clear. A 10 Hz resolution is used in the actual ATILA calculations. The same results are obtained by using 40×4 and 30×3 elements in ATILA. The 40×4 elements are used in ATILA and the same results were obtained using 30×3 elements. The good agreement again indicates that the Flugge model can be used to find the piezoelectric coefficients of thin shells by using the measured values of IEA. The effect of losses is also seen by comparing the results for shells with $L=10$ mm, $a=10$ mm, and $h=2$ mm in Tables II and III.

In Table III, results are also shown for other shells with either h/L or h/a, or both, greater than or equal to 0.2. However, L/a is less than or equal to 1 for all the cases. There is fairly good agreement between the present model and ATILA for cases A–E. In cases B–E, 40×8 elements are used. In case F, 20×20 elements are used and the same results are obtained by using 15×15 elements. The error in the first resonance frequency in the lower branch is less than 0.2% for these cases. However, for case F, $h/L=1$ and $h/a=0.5$ and the error is 1%. Similarly, the error in the first resonance frequency in the upper branch is less than 1% for cases A–D and increases to 4% for case E and 28% for case F. For thick shells, it is necessary to include not only the effects of shear stress and rotary inertia, but also radial stress[16]—all of which have been neglected here. The error in the upper branch resonance frequency is greater than the error in the lower branch resonance frequency for all the cases. In case F, at the lower branch resonance frequency, the maximum radial stress, computed using ATILA, is about 3 times the maximum axial stress and the maximum tangential stress is about 20 times the maximum radial stress. However, the maximum radial stress at the upper branch resonance frequency is of the same order as the maximum axial and tangential stresses. Therefore, radial stress cannot be neglected when computing the upper branch resonance frequency and this explains the 28% difference in resonance frequency between the present model and ATILA.

The first resonance frequency in the lower and upper branches, of the shells that are not very thick (cases A–E), is nearly inversely proportional to a and L, respectively. The resonance frequencies are also weakly dependent on the thickness, as can be seen from the results for cases A–B and cases D–E.

The error in the maximum value of the conductance, G, at the first resonance frequency in the lower branch, is approximately 0% in case A, 1% in case E, and 10% in case F. The error in the value of G at the first resonance frequency in the upper branch is greater than that in the lower branch in all the cases. It is seen, from Table III, that the values of G at the first resonance frequencies in the lower and upper branches are approximately proportional to h/L and ah/L^2, respectively.

Similar conclusions are drawn about the values of the susceptance, B. The maximum and minimum values of B are of interest because these values, along with other measured values, can be used to find the electroelastic coefficients of cylindrical shells. The maximum and minimum values of B in the lower and upper branches, shown in Table III, are also approximately proportional to h/L and ah/L^2, respectively.

The IEA is shown in Figs. 3–5 for cases C–E of Table III. Only the first resonance frequency of the lower branch can be seen in Figs. 2–4, and, as noted earlier, the values computed using the present model are in good agreement with ATILA. The higher-order resonances cannot be seen. In Fig. 5, the second resonance frequency in the lower branch is seen. When $L=5$ mm, $a=10$ mm, and $h=2$ mm, the Flugge model has a resonance at 258 kHz, but the ATILA model has a very weak resonance at 181 kHz. The agreement is very poor in this case but, as noted earlier, better agreement is obtained for the thin shells in Table II.

IV. CONCLUSIONS

A solution for the displacement–potential equations of thin axially polarized shells is presented. The 12 coefficients in the solution are complex when the shell has internal losses

and are determined by using the boundary conditions at the ends of the cylinder and the equipotential conditions on the flat surfaces. The displacements and potential are expressed as the sum of a finite number of terms but the input electrical admittance is a single term. Similar solutions can be obtained, even if other thin shell theories are used.

Numerical values of the input electrical admittances of cylinders with complex piezoelectric coefficients are presented. They are compared with values obtained using finite element analysis. It is shown that the frequencies at which the input electrical conductance, G, reaches a local maximum and the input electrical susceptance, B, reaches a local maximum or minimum as well as the values of G and B at these frequencies, computed using the model presented here and the finite element program are in good agreement for shells that are not very thick. However, when the wall thickness of the shell is comparable to the length, the radial stress that is assumed to be zero everywhere in the cylinder cannot be neglected at the upper branch resonance frequency, and the error is large.

It is inferred from the numerical results that the maximum value of the input electrical conductance, and the maximum and minimum values of the input electrical susceptance, in the neighborhood of the first resonance frequencies of the lower and upper branches are approximately proportional to h/L and ah/L^2, respectively.

The model used here can be used to accurately find the response of electrically excited thin shells with internal losses. It can also be used to find the electroelastic coefficients of shells using measured values of input electrical admittance.

ACKNOWLEDGMENT

Permission from Director, NPOL, to publish this paper is gratefully acknowledged.

[1] D. Stansfield, *Underwater Electroacoustic Transducers* (Bath University Press, Bath, UK, 1990).
[2] O. B. Wilson, *Introduction to the Theory and Design of Sonar Transducers* (Peninsula, Los Altos, CA, 1988).
[3] M. C. Dokmeci, ''Shell theory for vibrations of piezoceramics under a bias,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **17**, 369–385 (1990).
[4] M. C. Dokmeci, ''On the higher order theories of piezoelectric crystal surfaces,'' J. Math. Phys. **15**, 2248–2252 (1974); ''Recent advances: Vibrations of piezoelectric crystals,'' Int. J. Eng. Sci. **18**, 431–448 (1980); ''Dynamic applications of piezoelectric crystals. Part II. Theoretical Studies,'' Shock Vib. Dig. **15**, 15–26 (1983); ''Recent progress in the dynamic applications of piezoelectric crystals,'' *ibid.* **20**, 3–20 (1988).
[5] N. N. Rogacheva, *The Theory of Piezoelectric Shells and Plates* (CRC Press, London, 1993).
[6] H. S. Tzou and J. P. Zhong, ''A linear theory of piezoelastic shell vibrations,'' J. Sound Vib. **175**, 77–88 (1994); H. S. Tzou, *Piezoelectric Shells—Distributed Sensing and Control of Continua* (Kluwer, Boston, 1993).
[7] D. S. Drumheller and A. Kalnins, ''Dynamic shell theory for ferroelectric ceramics,'' J. Acoust. Soc. Am. **47**, 1343–1353 (1970).
[8] D. D. Ebenezer and P. Abraham, ''Piezoelectric thin shell theoretical model and eigenfunction analysis of radially polarized ceramic cylinders,'' J. Acoust. Soc. Am. **105**, 154–163 (1999).
[9] D. D. Ebenezer and A. J. Sujatha, ''New methods to characterize radially polarized piezoelectric ceramic cylindrical shells of finite length,'' J. Acoust. Soc. Am. **102**, 1540–1548 (1997).
[10] J.-N. Decarpigny, J.-C. Debus, B. Tocquet, and D. Boucher, ''In-air analysis of piezoelectric Tonpilz transducers in a wide frequency band using a mixed finite element-plane wave method,'' J. Acoust. Soc. Am. **78**, 1499–1507 (1985).
[11] W. Flugge, *Stresses in Shells*, 2nd ed. (Springer-Verlag, Berlin, 1973).
[12] R. Holland, ''Representation of dielectric, elastic and piezoelectric losses by complex coefficients,'' IEEE Trans. Sonics Ultrason. **14**, 18–20 (1967).
[13] H. F. Tiersten, ''Hamilton's principle for linear piezoelectric media,'' Proc. IEEE **55**, 1523–1524 (1967).
[14] W. Soedel, *Vibrations of Shells and Plates*, 2nd ed. (Marcel Dekker, New York, 1993).
[15] B. K. Mukherjee and S. Sherrit, ''Characterisation of piezoelectric and electrostrictive materials for acoustic transducers: I. Resonance methods,'' *5th International Congress on Sound and Vibration*, Adelaide, Australia, 15–18 December 1997, pp. 385–393.
[16] J. E. Greenspon, ''Vibrations of a thick-walled cylindrical shell—Comparison of the exact theory with approximate theories,'' J. Acoust. Soc. Am. **32**, 571–578 (1960).

Compressional and torsional wave amplitudes in rods with periodic structures

A. Morales,[a] J. Flores,[b] L. Gutiérrez, and R. A. Méndez-Sánchez
Centro de Ciencias Físicas UNAM, P.O. Box 48-3, 62251 Cuernavaca, Morelos, México

(Received 11 December 2001; revised 29 July 2002; accepted 3 August 2002)

To measure and detect elastic waves in metallic rods a low-frequency electromagnetic-acoustic transducer has been developed. Frequencies range from a few hertz up to hundreds of kilohertz. With appropriate configuration of the transducer, compressional or torsional waves can be selectively excited or detected. Although the transducer can be used in many different situations, it has been tested and applied to a locally periodic rod, which consists of a finite number of unit cells. The measured wave amplitudes are compared with theoretical ones, obtained with the one-dimensional transfer matrix method, and excellent agreement is obtained. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1509431]

PACS numbers: 43.38.Dv, 43.40.At, 43.20.Ks, 43.40.Cw [SLE]

I. INTRODUCTION

Waves in rods and bars have been the subject of extensive research for many years.[1] Besides the treatment of the uniform rod in textbooks, many variants of the problem have been tackled, due to its scientific and technological significance. Two of the most important applications are vibration isolation[2–6] and suppression of unwanted resonant modes.[7] Of growing interest is the scattering on spatially extended systems, for which an important question is to understand how the transport properties are related to the resonance spectrum.[8,9] Torsional waves in a rod with a diameter step have also been analyzed recently,[10,11] as well as normal modes and frequencies in periodic and nonperiodic rods.[12–14] In a closely related problem, the spatial analysis of torsional waves in a cylindrical waveguide has been presented.[15] However, most of these works are theoretically oriented and a comparison with experimental data is seldom performed. In particular, normal mode wave amplitudes for torsional and compressional waves are not usually measured.

It is the aim of the present paper to apply the electromagnetic-acoustic transducer (EMAT) recently developed[16] to excite and detect compressional and torsional waves in cylindrical metallic rods with notches arranged in different configurations. Although other techniques, such as magnetostrictive[17,18] and interferometric methods[19,20] among others, can be used for this purpose, this EMAT provides a simpler approach.

In Sec. II the EMAT and the experimental setup are described. In Sec. III the one-dimensional transfer matrix method used to obtain numerical results is discussed. Boundary conditions are imposed as is normally done in textbooks. Being one dimensional, in this calculation the coupling between different kinds of waves is neglected. In Sec. IV experimental results are compared with numerical values for both frequency spectrum and wave amplitudes. Since the usual boundary conditions to join rods with different cross sections are approximate, it is necessary to adjust a parameter to fit theory with experiment.

II. EXPERIMENTAL SETUP

An EMAT is a noncontacting device to excite and detect elastic waves. The EMAT recently developed consists of a simple coil and a permanent magnet (see Fig. 1). When used as an exciter, the time-varying magnetic field of the coil induces a time-varying eddy current in the metal. The field of the permanent magnet then generates a Lorentz force on the eddy current accelerating the rod. When used as a detector, the motion of the rod induces a variable magnetic flux in the rod due to the field of the permanent magnet. This, according to Faraday's law, generates an emf, which in turn produces eddy currents. The variable magnetic field arising from these eddy currents is then detected by the coil of the EMAT. The emf produced by the EMAT detector is proportional to the acceleration of the rod surface. This is because the eddy current is proportional to the time derivative of the magnetic flux, and this is proportional, in a first approximation, to the speed of the surface of the metal. Since the emf induced in the coil is proportional to the time derivative of the magnetic field produced by the eddy current, this emf is, again in a first approximation, proportional to the acceleration of the rod surface.

Different configurations of the coil and permanent magnet can produce distinct types of waves.[16] Compressional waves are excited or detected using setup (a) or (b) of Fig. 1. Setup (b) can measure the amplitudes along the rod. On the other hand, setups (c) and (d) are appropriate for torsional waves. Furthermore, it can also be used to excite bending oscillations. It is, therefore, very versatile notwithstanding the fact that it is composed of readily accessible commercial components. In contradistinction with a typical EMAT,[21–23] which is designed to operate at megahertz frequencies, ours has been tested from a few hertz to hundreds of kilohertz. To our knowledge, there is no other EMAT working in this frequency range.

[a]Author to whom correspondence should be addressed; electronic mail: mori@fis.unam.mx
[b]Permanent address: Instituto de Física, UNAM, P.O. Box 20-364, 01000 México, D. F. México.

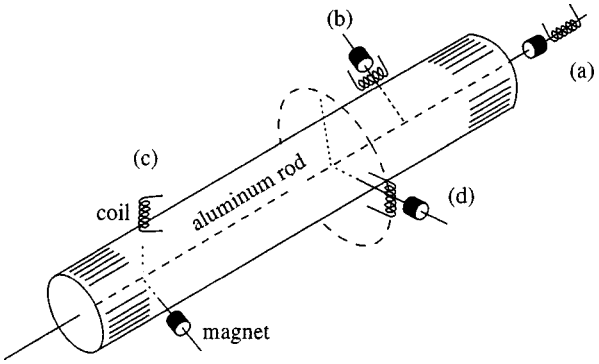


FIG. 1. Schematic drawing showing different EMAT configuratioins. By changing the configuration of the coil and magnet as well as its position different kinds of waves can be excited or detected: (a) and (b) compressional, (c) and (d) torsional.

In Fig. 2 the block diagram of the experimental setup is given. The signal of an oscillator (Stanford Research Systems, Function Generator DS345) is sent to a power amplifier (Krhon-Hite 7500) and then to the EMAT exciter. The frequency is measured with an HP 5382A Frequency Counter. The signal from the EMAT detector is amplified by a high-input impedance amplifier (EG&G PARC, Selective Amplifier 189). To monitor the resonances the signal is sent to the oscilloscope and to the EG&G PARC, Lock-In Amplifier 128A. The latter converts the ac signal to a dc voltage, which is applied to the *Y* input of the recorder. This voltage is proportional to the wave amplitude when exciting a normal mode of the rod. The reference signal for the lock-in was taken from the oscillator. The unwanted signal of the exciter picked up by the detector was dephased 90° with respect to the reference by the lock-in amplifier. This technique improves the measurement of the wave amplitudes.

The position *z* of the scanning detector along the rod is measured by mechanically coupling it to a cursor in contact with a nichrome wire, depicted by the bold line of Fig. 2, and then sent to the *X* input of the recorder. The normal mode amplitudes are then scanned along the rod. We should remark, however, that the resonant frequencies depend on temperature. Instead of controlling the latter, we changed the frequency of the exciter to follow the resonance. To do this, a third EMAT (denoted MONITOR in Fig. 2) is used to maintain the phase of the wave constant. The change in frequency is typically of the order of 0.1%, which is therefore the experimental error in the frequency measurements.

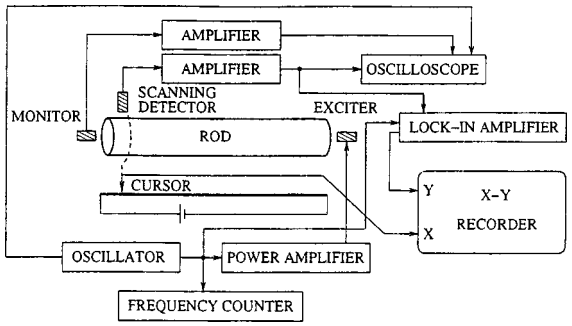


FIG. 2. Block diagram for the experimental setup. The EMATs (exciter, scanning detector, and monitor) are shown shaded and in a configuration to excite or detect compressional modes. The dashed line indicates mechanical coupling.

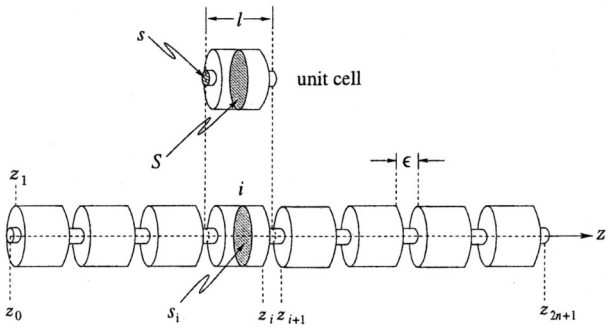


FIG. 3. Geometry of the rod with *n* cells and therefore $(n-1)$ notches. Here z_i denotes the position where the radius changes and s_i is the cross-section area of the *i*th cylinder of radius r_i. We take $z_0=0$ and $z_{2n+1}=L$.

III. THE ONE-DIMENSIONAL TRANSFER MATRIX METHOD

The rods analyzed here consist of *n* copies of a unit cell periodically arranged. Each cell is formed by three cylinders as shown in Fig. 3. The radius of the rod is *R* and that of the $(n-1)$ notches is $r<R$. The cross-section areas corresponding to *r* and *R* are *s* and *S*, respectively. Here $R \ll L$, the length of the rod, and $R \ll \lambda = 2\pi/k$, the normal mode wavelength, so the system indeed behaves as one dimensional.

The wave amplitude in cylinder *i* is

$$\psi_i(z) = A_i e^{ik(z-z_{i-1})} + B_i e^{-ik(z-z_{i-1})} \tag{1}$$

for $z_{i-1} \leqslant z \leqslant z_i$, $i=1,2,...,2n+1$. For compressional modes the boundary conditions at $z=z_i$ are

$$\psi_i|_{z_i} = \psi_{i+1}|_{z_i} \tag{2}$$

and

$$s_i \frac{\partial \psi_i}{\partial z}\bigg|_{z_i} = s_{i+1} \frac{\partial \psi_{i+1}}{\partial z}\bigg|_{z_i}, \tag{3}$$

where $s_i = \pi r_i^2$ is *s* or *S*, according to the value of *i*.

These boundary conditions are, however, approximate. In fact the wave amplitude ψ is a function of both *z* and the radial coordinate, since the boundary conditions at z_i are mixed. The semiquantitative argument given in the Appendix gives a simple result: if instead of the actual value $\eta = r/R$ an effective η_{eff}^C is used, a better approximation holds. It is obtained

$$\eta_{\text{eff}}^C = \frac{\eta}{(1+\alpha_C/\epsilon)^{1/2}}, \tag{4}$$

where ϵ is the width of the notch and α_C is a constant for fixed η that depends on the Young's modulus.

For torsional waves, the boundary conditions read

$$\psi_i|_{z_i} = \psi_{i+1}|_{z_i} \tag{5}$$

and

$$s_i^2 \frac{\partial \psi_i}{\partial z}\bigg|_{z_i} = s_{i+1}^2 \frac{\partial \psi_{i+1}}{\partial z}\bigg|_{z_i}. \tag{6}$$

As in the compressional case, these boundary conditions are approximate and an effective value $\eta_{\rm eff}^T$ should also be used. A similar semiquantitative argument gives

$$\eta_{\rm eff}^T = \frac{\eta}{(1+\alpha_T/\epsilon)^{1/4}}, \tag{7}$$

where α_T is a constant which now depends on the shear modulus. Taking into account Eqs. (2) and (3), the transfer matrix linking the wave amplitude coefficients A_{i+1} and B_{i+1} to A_i and B_i is given for compressional waves by

$$\begin{pmatrix} A_{i+1} \\ B_{i+1} \end{pmatrix} = \mathbf{M}^C_{i\to i+1} \begin{pmatrix} A_i \\ B_i \end{pmatrix}, \tag{8}$$

where

$$\mathbf{M}^C_{i\to i+1} = \frac{1}{2} \begin{pmatrix} \left[1+\left(\frac{r_i}{r_{i+1}}\right)^2\right] e^{ik(z_i-z_{i-1})} & \left[1-\left(\frac{r_i}{r_{i+1}}\right)^2\right] e^{-ik(z_i-z_{i-1})} \\ \left[1-\left(\frac{r_i}{r_{i+1}}\right)^2\right] e^{ik(z_i-z_{i-1})} & \left[1+\left(\frac{r_i}{r_{i+1}}\right)^2\right] e^{-ik(z_i-z_{i-1})} \end{pmatrix}. \tag{9}$$

Equations (5) and (6) yield for torsional modes:

$$\begin{pmatrix} A_{i+1} \\ B_{i+1} \end{pmatrix} = \mathbf{M}^T_{i\to i+1} \begin{pmatrix} A_i \\ B_i \end{pmatrix}, \tag{10}$$

with

$$\mathbf{M}^T_{i\to i+1} = \frac{1}{2} \begin{pmatrix} \left[1+\left(\frac{r_i}{r_{i+1}}\right)^4\right] e^{ik(z_i-z_{i-1})} & \left[1-\left(\frac{r_i}{r_{i+1}}\right)^4\right] e^{-ik(z_i-z_{i-1})} \\ \left[1-\left(\frac{r_i}{r_{i+1}}\right)^4\right] e^{ik(z_i-z_{i-1})} & \left[1+\left(\frac{r_i}{r_{i+1}}\right)^4\right] e^{-ik(z_i-z_{i-1})} \end{pmatrix}. \tag{11}$$

In this way the coefficients A_{2n+1} and B_{2n+1} are expressed in terms of A_1 and B_1 as

$$\begin{pmatrix} A_{2n+1} \\ B_{2n+1} \end{pmatrix} = \mathbf{M}^D \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} \tag{12}$$

where

$$\mathbf{M}^D = \mathbf{M}^D_{2n\to 2n+1} \mathbf{M}^D_{2n-1\to 2n} \cdots \mathbf{M}^D_{2\to 3} \mathbf{M}^D_{1\to 2} \tag{13}$$

where D stands for C or T. We shall restrict ourselves to a rod with free ends. In this situation, Neumann boundary conditions hold for both the compressional and torsional cases. At $z_0 = 0$, they imply

$$A_1 = B_1, \tag{14}$$

and at $z_{2n+1} = L$,

$$A_{2n+1} e^{ik(L-z_{2n})} - B_{2n+1} e^{-ik(L-z_{2n})} = 0. \tag{15}$$

Inserting Eq. (12) in Eq. (15), together with Eq. (14), then leads to two homogeneous equations for A_1 and B_1:

$$\mathbf{P}^D \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{16}$$

where

$$\mathbf{P}^D = \begin{pmatrix} 1 & -1 \\ M^D_{11} e^{ik(L-z_{2n})} + M^D_{21} e^{-ik(L-z_{2n})} & M^D_{12} e^{ik(L-z_{2n})} + M^D_{22} e^{-ik(L-z_{2n})} \end{pmatrix}, \tag{17}$$

where M^D_{ij} are the matrix elements of $\mathbf{M}^D$.

The eigenvalues of the normal modes correspond to the roots of $\det(\mathbf{P}^D) = 0$. The allowed wave numbers are obtained numerically by scanning the determinant as a function of k. When a change of sign is found, the FORTRAN routine DZERO for root finding is called.[24] Once the roots are obtained, the wave amplitudes can be computed using Eq. (1).

IV. COMPARISON BETWEEN THEORY AND EXPERIMENT: WAVE AMPLITUDES AND THE EMERGENCE OF BAND SPECTRA

We have both measured and calculated the normal-mode frequencies as a function of the number n of unit cells. We show the spectrum for torsional waves in Fig. 4, where we compare theory with experiment. The emergence of bands is

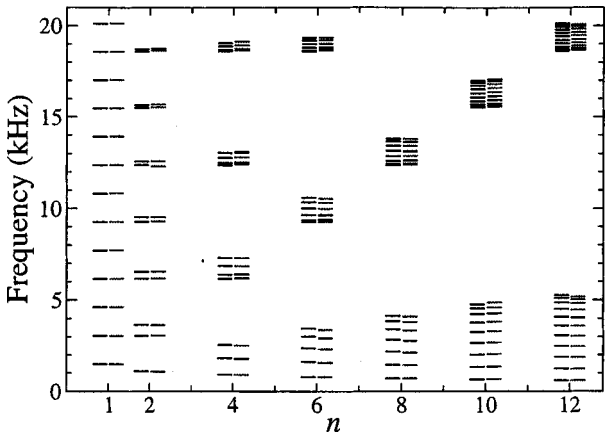


FIG. 4. Spectra for torsional waves showing the emergence of bands of allowed frequencies as a function of n. The theoretical and experimental values are given for each n by the left-and right-hand side lines, respectively. In this and in the following plots the frequency equal to zero is excluded. In the experiment $\eta=0.236$, $\epsilon=0.275$ mm, $L=1.0005$ m, and $R=6.35$ mm. The value of $\alpha_T=0.89$ mm for the theoretical values is obtained by a least-squares fit of the theory with the experimental data for $n=2$, 4, 6, 8, 10, 12. A least-squares fit for each n gives $\alpha_T=0.91\pm0.09$ mm where the uncertainty is the standard deviation on the data. This uncertainty arises from mechanical imperfections in the construction of the rods.

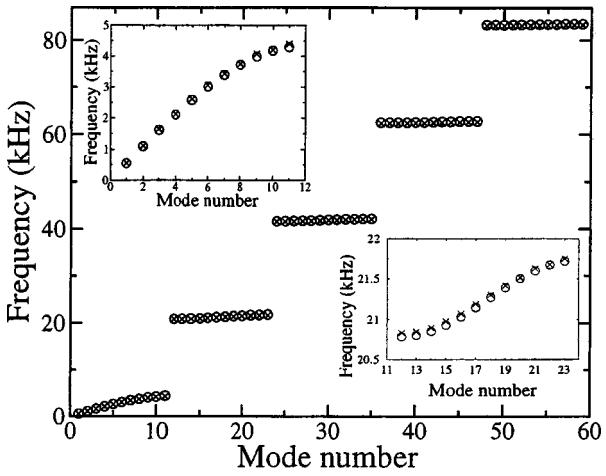


FIG. 6. Torsional wave spectrum showing the agreement between the experimental values (crosses) and transfer matrix results (circles) for $n=12$ using the η^T_{eff} obtained from Eq. (4) with $\alpha_T=0.88$ mm. This value is compatible with that obtained in Fig. 4. The experimental parameters are $\eta=0.209$, $\epsilon=0.33$ mm, $L=0.9106$ m, and $R=6.45$ mm. The upper-left and lower-right insets show a zoom of the first and second bands, respectively, to make the small disagreement apparent.

evident as n grows. This effect has been previously shown in theoretical calculations[8,9,25] but not for the elastic case and even less experimentally. Theoretical values fit the experimental ones very well for all frequencies and all values of n if one uses a single value of α_T and Eq. (7). The value of α_T was obtained using a least-squares fit.

Bands also appear for compressional waves and, even more, when the shape of the obstacles is altered. This is shown in Fig. 5, where experimental values for a rod with drilled circular holes instead of notches are given. In what follows, all other results are given for rods with notches.

The band structure appears more clearly in Fig. 6 where the frequencies for torsional waves forming the first five bands for $n=12$ are compared with the theoretical results. When the frequency increases, the gap grows until the bands look like an equidistant, n-fold degenerate spectrum, corresponding to that of the unit cell. This behavior contrasts with that of the typical quantum mechanical case.[26] The first level of each band is independent of the value of η because in this case the notches correspond to those points where the derivative of the wave amplitude vanishes. The gap is produced when the Bragg condition $kl=m\pi$ is fulfilled, where l is the length of the unit cell and m is an integer.

A similar band structure occurs for compressional waves as shown in Fig. 7 but with two differences: (i) a larger value η^2 instead of η^4 rules the wave motion, as seen from Eqs. (9) and (11), and (ii) there is a small deviation between theory and experiment in the second and third bands. This can be understood if the Rayleigh correction, not present for torsional waves, is taken into account.[16]

The wave amplitudes were also measured and compared with those predicted by the transfer matrix method. In Fig. 8

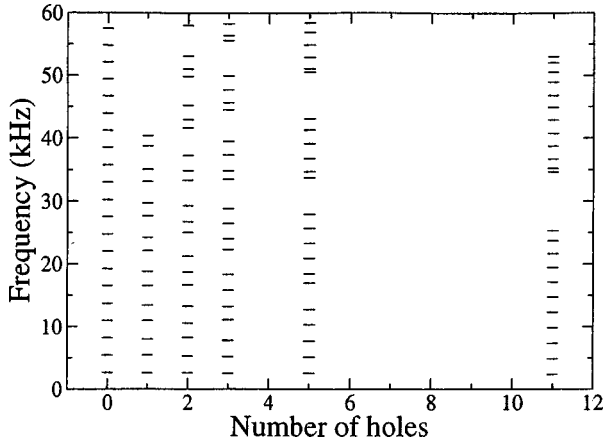


FIG. 5. Emergence of the bands for compressional waves in an aluminum rod ($R=6.45$ mm, $L=0.910$ m) with holes instead of notches. The radius of the holes is 3.5 mm. Notice that the number of holes corresponds to $(n-1)$.

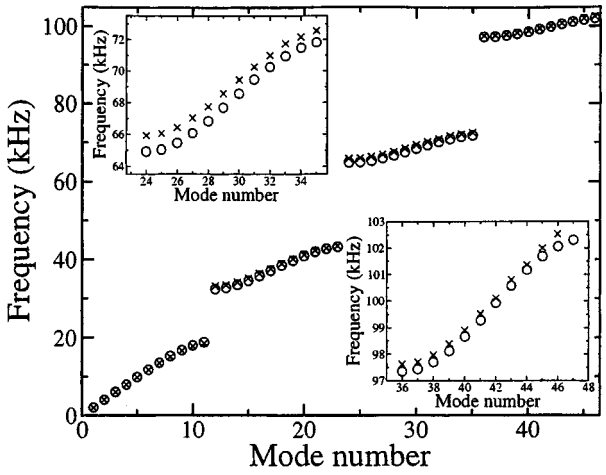


FIG. 7. Compressional wave spectrum showing the agreement between the experimental (crosses) and the transfer matrix theory results (circles) using $\alpha_C=2.17$ mm. The experimental parameters of the rod are the same as in Fig. 6. The upper-left and lower-right insets show a zoom of the third and fourth bands, respectively.

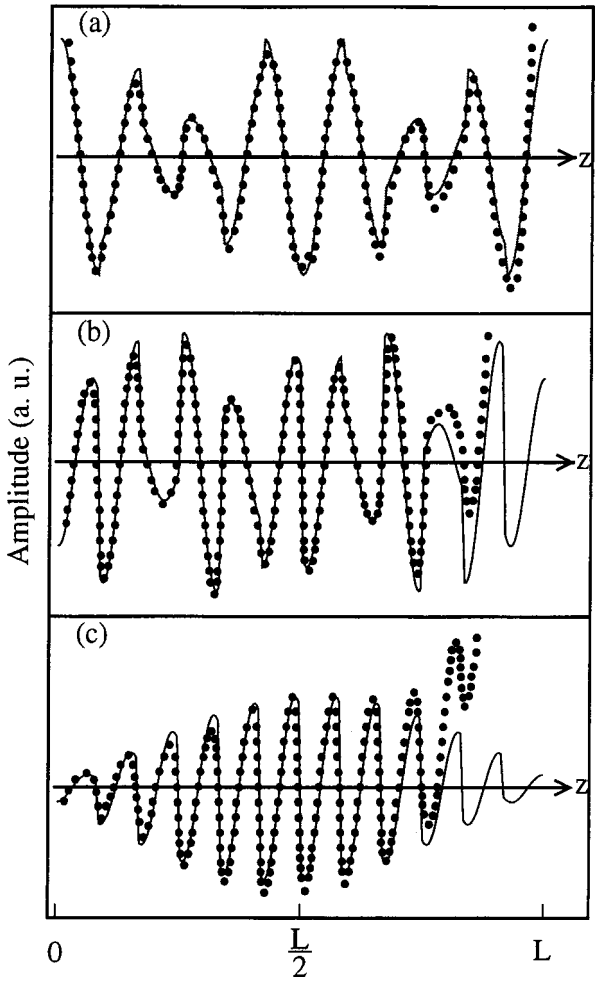


FIG. 8. Compressional wave amplitudes showing the agreement between experimental results (dots) and the transfer matrix theory (continuous line): (a) mode 14, (b) mode 19, and (c) mode 23. The same parameters as in Fig. 6 are used. The amplitudes are measured continuously along the rod but only some points are plotted for a better presentation.

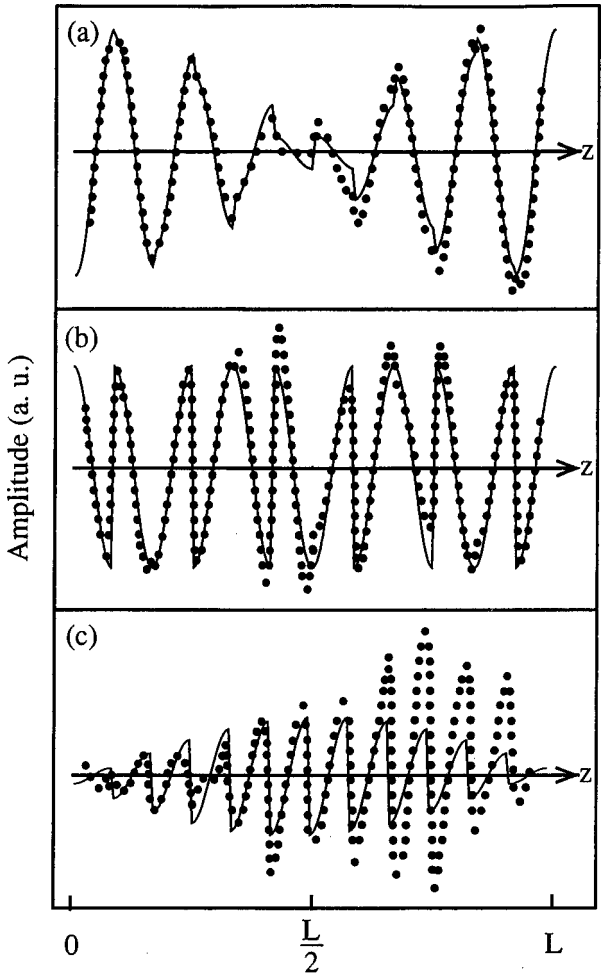


FIG. 9. Wave amplitudes for torsional waves (a) mode 13, (b) mode 18, and (c) mode 23. The experimental results are denoted by dots and the transfer matrix theory results correspond to the continuous line. We have used the same parameters as in Fig. 6.

a comparison is done for three different compressional modes of the second band. Excellent agreement, except for $z \approx L$, is observed. These deviations come from the magnetic field of the exciter that is picked up by the coil. In spite of the dephasing procedure mentioned in Sec. II, this effect cannot be eliminated near the exciter.

The same comparison can be performed for torsional waves. In this case further asymmetries in the wave amplitudes are observed. This is not an artifact coming from the magnetic field picked up by the detector. In order to see this, we measured the wave amplitudes only along the part of the rod that is far away from the exciter and later on we turned the rod around and repeated the measurement. The experimental results presented in Fig. 9 are obtained by matching these two amplitudes. The asymmetry persists. The reason for this is the dependence on η^4 of the transmission amplitudes which makes them very sensitive to errors in the radii. For example, a change of the order of 10% in the radii of the fourth and fifth notches can produce such nonsymmetric wave amplitudes. In fact, when fitting the theoretical wave amplitude for mode 13 to the experimental one, a value $\eta_{\mathrm{eff}}^{T}=0.130$ is obtained for these notches instead of that obtained from Eq. (7) used for the rest of the obstacles. Using these values other normal mode wave amplitudes agree very well with the experimental ones, as shown by examples (b) and (c) of Fig. 10. Note that uncertainties of this order of magnitude on the radii of the notches are compatible with the machining tolerance.

V. CONCLUSION

We have developed an EMAT which can be used to measure the normal-mode wave amplitudes for compressional and torsional waves in metallic rods. The EMAT is highly selective, works for low frequencies, and is constructed with commercial components. In this work, we have applied this EMAT to study resonances in rods with notches arranged in a periodic structure. A band spectrum emerges as the number of notches increases. Theoretical normal-mode frequencies and wave amplitudes, calculated with a transfer matrix approach, agree very well with the experimental ones.

The EMATs of Fig. 1 can also be used in a variety of systems. For example, bending waves in rods[16] and plates may be analyzed very conveniently with it. And not only this. It is also useful in studying three-dimensional systems. Since the wave amplitudes can be measured, besides their

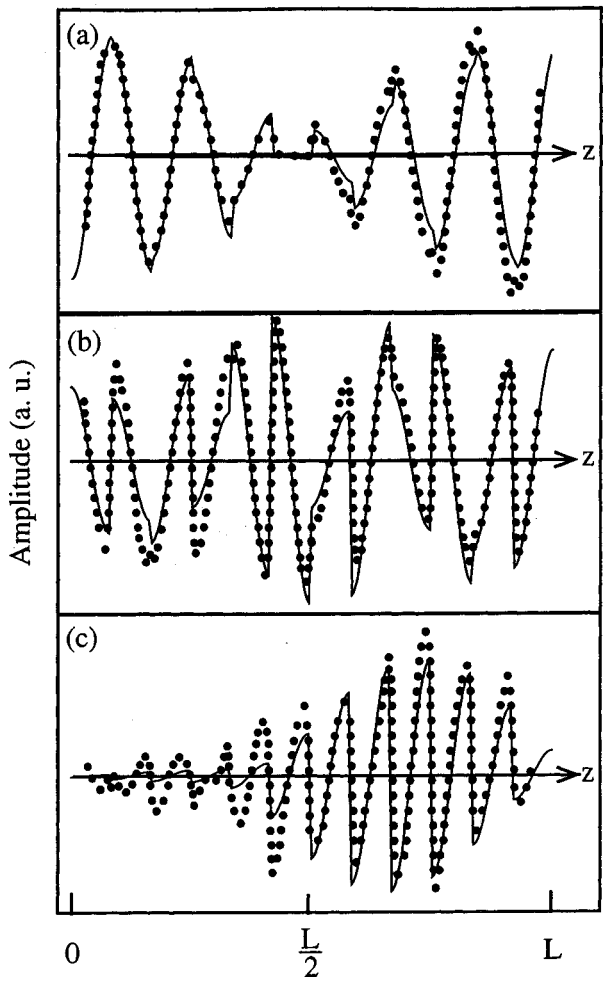


FIG. 10. Wave amplitudes for torsional waves showing good agreement between the experimental results (dots) and the transfer matrix theory (continuous line) with two defects in the fourth and fifth notches; $\eta_{\text{eff}}^{T}=0.130$ while for the rest of the notches the value obtained from Eq. (4) was used.

intrinsic interest these elastic systems may be used to further understand quantum mechanical ones, for which the wave function is not normally available. Results obtained for elastic vibrations are useful, for example, to test many results in the physics of waves. These studies are nowadays of great interest and are giving rise to what can be called acoustic (or elastic) chaos.[27]

ACKNOWLEDGMENTS

We thank A. Díaz-de-Anda for his help in the experimental measurements. R.A.M.S. wishes to thank G. Báez, H. Larralde, and J. Maytorena for useful discussions. This work was supported by DGAPA-UNAM Project No. IN104400 and by CONACyT, México.

APPENDIX

A semiquantitative argument to justify Eq. (4) is presented here. A similar argument is valid for the torsional case. To evaluate the transfer matrix of Eq. (9) we have used the boundary condition

$$E_1 s_1 \frac{d\psi_1}{dz} = E_2 s_2 \frac{d\psi_2}{dz}, \tag{A1}$$

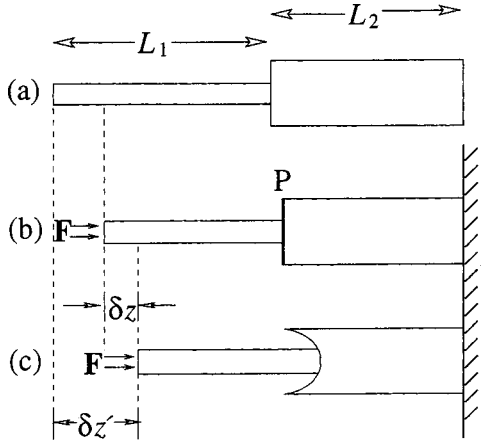


FIG. 11. Diagram showing the deformation of the stepped rod. In (b) a nondeformable thin plate P is present and in (c) this plate is eliminated.

with E_i the Young modulus, as is normally done in textbooks. In Eq. (3), E_1 is taken equal to E_2. Equation (A1) is approximate because the boundary condition at the radius step is not correctly taken into account.

As shown in Sec. IV, the experimental data are fitted very well using only one parameter α_C for all the frequencies considered. In particular, this is true for very low-frequency values. Therefore in what follows only the static case will be discussed.

Consider two circular rods of different cross-section areas $s_1 < s_2$, arranged as in Fig. 11(a). Imagine the two rods separated by a thin nondeformable plate P as in Fig. 11(b). In such a case, a force F applied to the free end of rod 1 is transmitted uniformly to rod 2. By Hooke's law, rod 1 is deformed by

$$\delta L_1 = \frac{F L_1}{E_1 s_1} \tag{A2}$$

and rod 2 by

$$\delta L_2 = \frac{F L_2}{E_2 s_2}. \tag{A3}$$

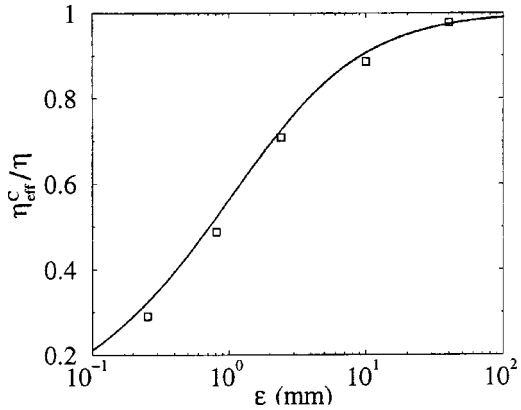


FIG. 12. The ratio $\eta_{\text{eff}}^{C}/\eta$ as a function of ϵ for $n=6$. The curve corresponds to Eq. (4) valid for the compressional case and the squares to the experimental values. The value of $\alpha_C = 2.17$ mm in the theoretical curve was taken from Fig. 7.

Notice that the same F appears in both Eqs. (A2) and (A3). If we equate F from these two expressions, we arrive at Eq. (A1) if $d\psi/dz$ is constant such as it is in the static case. We might say that the presence of the rigid plate P makes the boundary condition (A1) exact.

If plate P is eliminated, however, an extra deformation δz appears, as in Fig. 11(c). If one makes the reasonable assumption that δz is proportional to F, then $\delta z = cF$. Here c depends in a complicated way on many factors, but is independent of L_1. The total deformation $\delta z' = \delta L_1 + \delta L_2 + \delta z$ is given by

$$\delta z' = \left(\frac{FL_1}{E_1 s_1} + cF \right) + \frac{FL_2}{E_2 s_2}. \tag{A4}$$

To obtain an expression such as Eq. (A1) for the boundary condition, we introduce an effective area s_1^{eff} for rod 1, defined as

$$\frac{L_1}{E_1 s_1^{\text{eff}}} = \frac{L_1}{E_1 s_1} + c. \tag{A5}$$

This yields

$$s_1^{\text{eff}} = \frac{s_1}{1 + c s_1 E_1 / L_1}. \tag{A6}$$

If one realizes that $(\eta_{\text{eff}}^C)^2 = s_1^{\text{eff}}/s_2$, Eq. (A6) is equivalent to Eq. (4) in the text after replacing L_1 by ϵ and α_C by $c s_1 E_1$.

We have verified the dependence of η_{eff}^C on ϵ, by measuring the spectrum of five rods with different values of ϵ. For each rod an effective value of η_{eff}^C was obtained from a least-squares fit as was done in connection with Fig. 7. These values are shown by the squares in Fig. 12. They agree with Eq. (4) as shown in this figure by the continuous line obtained with $\alpha_C = 2.17$ mm (see Fig. 7).

[1] K. F. Graff, *Wave Motion in Elastic Solids* (Dover, New York, 1991).
[2] J. C. Snowdon, "Vibration isolation: Use and characterization," J. Acoust. Soc. Am. **66**, 1245–1274 (1979).
[3] A. S. Bansal, "Free-wave propagation through combinations of periodic and disordered systems," J. Acoust. Soc. Am. **67**, 377–389 (1980).
[4] B. Ravindra, "Love-theoretical analysis of periodic system of rods," J. Acoust. Soc. Am. **106**, 1183–1186 (1999).
[5] B. Ravindra and A. K. Mallik, "Harmonic vibration isolation characteristics of periodic systems," J. Sound Vib. **154**, 249–259 (1992).
[6] C. H. Hodges and J. Woodhouse, "Vibration isolation from irregularity in a nearly periodic structure: Theory and measurements," J. Acoust. Soc. Am. **74**, 894–905 (1983).
[7] W. Johnson, B. A. Auld, E. Segal, and F. Pasarelli, "Trapped torsional modes in solid cylinders," J. Acoust. Soc. Am. **100**, 285–293 (1996).
[8] D. J. Griffiths and C. A. Steinke, "Waves in locally periodic media," Am. J. Phys. **69**, 137–154 (2001).
[9] F. Barra and P. Gaspard, "Scattering in periodic systems: From resonances to band structure," J. Phys. A **32**, 3357–3375 (1999).
[10] H. E. Engan, "Torsional wave scattering from a diameter step in a rod," J. Acoust. Soc. Am. **104**, 2015–2024 (1998).
[11] K. Liu, S. Xie, and H. Sun, "Reflection characteristic of torsional waves in transversely isotropic stepped bars," Comput. Methods Appl. Mech. Eng. **153**, 107–115 (1998).
[12] A. S. Bansal, "Flexural waves and deflection mode shapes of periodic and disordered beams," J. Acoust. Soc. Am. **72**, 476–481 (1982).
[13] D. Li, H. Benaroya, "Waves, normal modes, and frequencies in periodic and near-periodic rods. I," Wave Motion **20**, 315–338 (1994); "Waves, normal modes and frequencies in periodic and near-periodic rods. II," **20**, 339–358 (1994).
[14] D. J. Mead, "A new method of analyzing wave propagation in periodic structures: Applications to periodic Timoshenko beams and stiffened plates," J. Sound Vib. **104**, 9–27 (1986).
[15] J. C. Aime, M. Brissaud, and L. Laguerre, "Spatial analysis of torsional wave propagation in a cylindrical waveguide. Application to magnetostrictive generation," J. Acoust. Soc. Am. **109**, 51–58 (2001).
[16] A. Morales, L. Gutiérrez, and J. Flores, "Improved eddy current driver-detector for elastic vibrations," Am. J. Phys. **69**, 517–522 (2001).
[17] H. Kwun and C. M. Teller, "Magnetostrictive generation and detection of longitudinal, torsional and flexural waves in a steel rod," J. Acoust. Soc. Am. **96**, 1202–1204 (1994).
[18] H. Kwun and K. A. Bartels, "Experimental observation of elastic-wave dispersion in bounded solids of various configurations," J. Acoust. Soc. Am. **99**, 962–968 (1996).
[19] M. Gindre, W. Urbach, R. H. Coursant, and M. Fink, "Mechanical displacement induced in a piezoelectric structure: Experimental measurement by laser interferometry and simulation by a finite element method," J. Acoust. Soc. Am. **84**, 11–19 (1988).
[20] C. Campos-Pozuelo and J. A. Gallego-Juárez, "Finite amplitude flexural vibrations at ultrasonic frequencies in metallic bars," J. Acoust. Soc. Am. **98**, 1742–1750 (1988).
[21] R. B. Thompson, "Physical principles of measurements with EMATs," in *Physical Acoustics*, edited by R. N. Thurston and D. Pierce (Academic, Boston, 1990), Vol. XIX.
[22] G. A. Alers, "Noncontacting electromagnetic and magnetostrictive transducers-their unique feature and applications," Acta Phys. Sin. **8**, S9–S15 Suppl. S (1999).
[23] H. Ogi, "Field dependence on coupling efficiency between electromagnetic field and ultrasonic bulk waves," J. Appl. Phys. **82**, 3940–3949 (1997).
[24] See http://gams.nist.gov/ and http://www.netlib.org/
[25] E. Cota, J. Flores, and G. Monsivais, "A simple way to understand the origin of the electron band structure," Am. J. Phys. **56**, 366–372 (1988).
[26] J. D. Patterson, *Introduction to the Theory of Solid State Physics* (Addison–Wesley, Reading, MA, 1971), p. 109.
[27] C. Ellegaard, K. Schaadt, and P. Bertelsen, "Acoustic chaos," Phys. Scr., T **T90**, 223–230 (2001).

Topology optimization of the kerf fillings in linear phased arrays for therapy

Yuan Lin[a] and Karl Grosh[b]
Department of Mechanical Engineering, University of Michigan CoE, 2250 G. G. Brown Laboratory, 2350 Hayward Street, Ann Arbor, Michigan 48109-2125

(Received 20 July 2001; revised 3 July 2002; accepted 25 July 2002)

Cross talk in ultrasonic arrays is an important problem that affects the array performance. For therapy arrays, cross talk decreases the acoustic power efficiency and array steering capabilities. Since arrays used for therapy are generally air-backed, structure-borne cross talk from the kerf fillings that connect the elements is predominant. In this paper, design strategies are developed to reduce the cross talk in linear phased arrays by optimizing the topology of kerf fillings. Two topology optimization schemes, an element-by-element design and layer-by-layer design, are developed. Sequential linear programming is used to solve the nonlinear optimization problem. Using an array subset technique as the basis for the design, an objective function based on maximizing the normal velocity response of the wet surface of an electrically driven element is shown to improve both the power output and steering of the array. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1510531]

PACS numbers: 43.38.Hz [SLE]

I. INTRODUCTION

Ultrasonic arrays have been widely used in imaging applications, such as medical diagnosis, nondestructive evaluation and underwater acoustics. Recently, the use of phased arrays for noninvasive therapy treatment has received increased attention. Hyperthermia treatment and high intensity focused ultrasound tissue ablation surgery represent two important therapeutic applications, where ultrasonic waves are used to induce changes in tissues through the thermal and/or mechanical interactions (Driller and Lizzi, 1987; Bihrle *et al.*, 1994; Sanghvi *et al.*, 1996). Phased arrays consist of an assembly of elements and provide unique advantages over geometrically focused transducers. Their ultrasonic beams are electronically focused and steered. Thence they allow for rapid on-line control of the ultrasound distribution during the surgery without mechanical movement of the applicator head. Furthermore, phased arrays can be used to synthesize multiple-foci patterns and the focal patterns can be moved at electronic speed to compensate for the body/organ movement. Algorithms correcting for tissue aberrations can be implemented to enhance the quality of the focal patterns (Ebbini and Cain, 1989, 1991).

As phased arrays become more prevalent in ultrasound therapy, efficient computer-aided modeling and design of these phased arrays for high power output and maximal steering capabilities also gains importance. For ideal beam forming, each element would act independently allowing for complete electronic control of the element output velocity and be acoustically omnidirectional. Cross talk, either via structural or acoustical paths, destroys this perfect control through the interelement induced motion. The transfer of power to the other elements can reduce the radiation efficiency of the element and the array's angular steering capabilities by increasing the effective aperture of each element (Kino and DeSilets, 1979; Turnbull and Foster, 1991).

In imaging arrays, the cross talk between array elements principally arises from the backing and the matching layer (Lerch, 1990; Greenstein *et al.*, 1997). It leads to the loss in the range resolution, limits the steering angles and increases the ringdown time through delayed signals from neighbor elements (Dias, 1982). Some strategies have been proposed to reduce the cross talk in imaging arrays by experimentally changing the slot depth or slot angle into the backing (Dias, 1982; Kino and Baer, 1983). While in therapy arrays, which are generally air-backed, cross talk from kerf fillings is predominant. Little work has been published for the reduction of cross talk in therapy arrays. The application of systematic topology optimization techniques in a finite element framework to the design of kerf fillings for therapy or imaging arrays has not been investigated.

In this paper, linear phased arrays for therapy are modeled and designed using the finite element method. Array subset synthesis (Assaad and Bruneel, 1996), and a far-field pressure projection algorithm (Lin *et al.*, 1997), are used in conjunction with density-based topology optimization schemes to design for enhanced acoustic performance. The optimization of the kerf fillings is linked so that the topology of each segment is identical. The effectiveness of different design schemes using this framework is determined in a fluid-loaded setting.

II. MODELING OF LINEAR PHASED ARRAYS

Because of the computational requirements engendered by a topology optimization scheme, efficiency in modeling the transducer system is of paramount importance. In this section, the numerical methodology by which the transducer system is modeled is presented. In ultrasound therapy, since the acoustic power is a main figure of merit, a continuous

[a]Current address: Structural Research and Analysis Corporation, 12121 Wilshire Blvd., Suite 700, Los Angeles, CA 90025.
[b]Electronic mail: grosh@engin.umich.edu

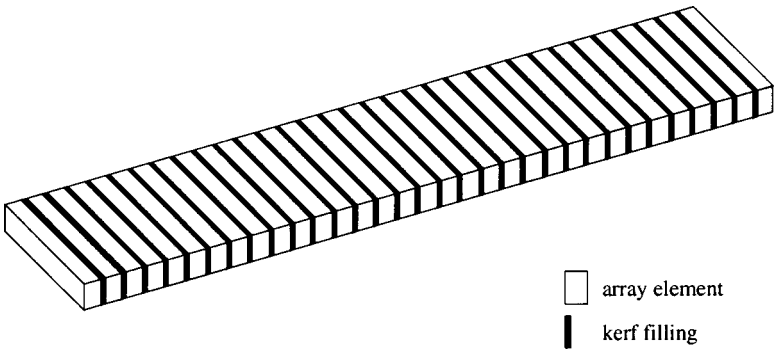


FIG. 1. A representative linear phased array for therapy.

wave, time harmonic signal is typically used to heat and ablate tissue in the focal region of an ultrasonic transducer (Driller and Lizzi, 1987). The underlying model for the transducer problem is that of time harmonic acoustics and elastodynamics ($e^{-i\omega t}$ time dependence is assumed).

A. Direct solution

A representative linear phased array for therapy is shown in Fig. 1, consisting of piezoelectric elements and kerf fillings. The typical kerf fillings are epoxy (Turnbull and Foster, 1992; Seyed-Bolorforosh, 1995), silicon rubber and silicon adhesive (Buchanan and Hynynen, 1994), or mixture of hollow glass microspheres in an epoxy matrix (Greenstein *et al.*, 1997).

Since the elements in linear phased arrays are generally parallelepipeds, with a high aspect ratio in one dimension, the longitudinal modes along the length direction are usually far below the frequency of interest. A two-dimensional plane strain model is appropriate. The time harmonic, fluid–structure interaction problem for the design of these piezoelectric elements is solved using a finite element code, as described in Lin *et al.* (1997); Lin and Grosh (2000b). If sufficient storage is available, the radiation of a fluid-loaded array driven by phased voltage signal can be directly modeled and solved. Since the system is linear, the frequency response of the entire array can also be obtained via superposition of the response from subproblems. The subproblems are formulated by sequentially driving only one array element with unity voltage while all other elements are short-circuited (voltage is zero). Finally, the response from subproblems is then superposed with weighted voltage amplitude and phase shift.

To show the equivalence of modeling in general, consider a fluid-loaded linear array of N electroded elements, each grounded on the wet side and excited on the other side. The drive voltage vector on each array element is $V_i e^{-j\varphi_i}\mathbf{I}$, where V_i, φ_i are, respectively, the voltage magnitude and phase, $\mathbf{I}=\langle 1,1,...,1\rangle^T$ is a vector and $\mathbf{I}\in\mathcal{R}^{i\text{node}\times 1}$ ($i$node is the number of electroded nodes on each array element), $i=1,2,...,N$. Then the discrete finite element matrix equations can be written as (Lin *et al.*, 1997; Lin and Grosh, 2000b)

$$\mathbf{A}\langle \mathbf{u},\mathbf{0},V_1e^{-j\varphi_1}\mathbf{I},V_2e^{-j\varphi_2}\mathbf{I},\ldots,V_ie^{-j\varphi_i}\mathbf{I},\ldots,V_Ne^{-j\varphi_N}\mathbf{I},\ \boldsymbol{\Phi}_{\text{un}},\mathbf{p}\rangle^T=\mathbf{0}, \tag{1}$$

where $\mathbf{A}$ is the global dynamic stiffness coefficient matrix, $\mathbf{u}$, $\mathbf{p}$ are the vectors of nodal mechanical displacement and fluid pressure, respectively, vector zero after the $\mathbf{u}$ denotes the voltage of the grounded wet side of array elements, and $\boldsymbol{\Phi}_{\text{un}}$ is the vector of unknown nodal electric voltage inside the piezoelectric element. Note that there are no mechanical or acoustical force terms in this study, indicated by the vector zero in the right-hand side of Eq. (1). Further details of the finite element formulation are omitted here in the interest of focusing on the superposition technique.

Although written as a set of homogeneous equations, these equations are forced (nonhomogeneous) through the applied voltage, which results in a right-hand side term. The frequency response of the whole array can be computed by directly solving the above-given matrix equations. Because of the linearity of the problem, the matrix equation in Eq. (1) is equivalent to

$$\begin{aligned}\mathbf{A}\langle \mathbf{u},\mathbf{0},V_1e^{-j\varphi_1}\mathbf{I},V_2e^{-j\varphi_2}\mathbf{I},\ldots,V_ie^{-j\varphi_i}\mathbf{I},\ldots,V_Ne^{-j\varphi_N}\mathbf{I},\\ \boldsymbol{\Phi}_{\text{un}},\mathbf{p}\rangle^T=\sum_{i=1}^{N}\mathbf{A}\langle V_ie^{-j\varphi_i}\mathbf{u}_i,\mathbf{0},\delta_{i1}V_1e^{-j\varphi_1}\mathbf{I},\\ \delta_{i2}V_2e^{-j\varphi_2}\mathbf{I},\ldots,\delta_{ii}V_ie^{-j\varphi_i}\mathbf{I},\ldots,\\ \delta_{iN}V_Ne^{-j\varphi_N}\mathbf{I},V_ie^{-j\varphi_i}\boldsymbol{\Phi}_{\text{un}_i}V_ie^{-j\varphi_i}\mathbf{p}_i\rangle^T,\end{aligned} \tag{2}$$

where δ_{il} is the Kronecker delta function and $\mathbf{u}_i$, $\boldsymbol{\Phi}_{\text{un}_i}$, and $\mathbf{p}_i$ are the solution of subproblem

$$\mathbf{A}\langle \mathbf{u}_i,\mathbf{0},\delta_{i1}\mathbf{I},\delta_{i2}\mathbf{I},\ldots,\delta_{ii}\mathbf{I},\ldots,\delta_{iN}\mathbf{I},\boldsymbol{\Phi}_{\text{un}_i},\mathbf{p}_i\rangle^T=0 \tag{3}$$

such that, for example,

$$\mathbf{u}=\sum_{i=1}^{N}V_ie^{-j\varphi_i}\mathbf{u}_i. \tag{4}$$

From Eq. (3), the zero voltage (or short-circuited) boundary conditions on electronically passive elements other than the ith driven element are justified.

B. Pressure projection

Although the ultimate design of a one-dimensional array will likely consist of many elements (greater than 50), for the topology optimization design iterations, it is desirable to use a smaller model on which to perform the optimization. The smaller model is denoted in this paper as an array subset. Once the design is completed, numerical tests are required to determine if the design enhancement is conveyed from the

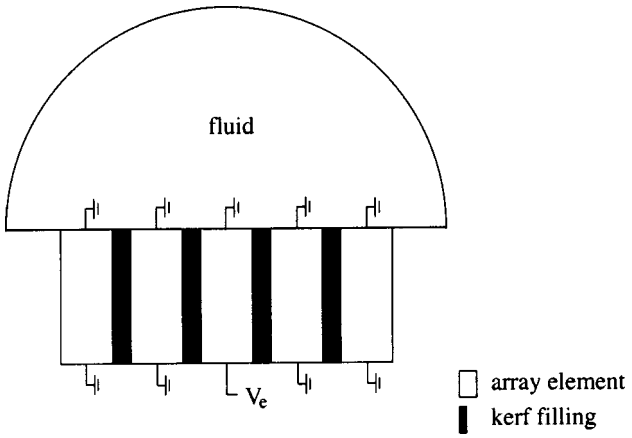


FIG. 2. Fluid-loaded five-element array subset for the acoustic field of a linear phased array, where V_e is the excitation voltage on the middle element.

array subset to the original large array. In Fig. 2, a five-element array is shown, typical of the size used in this study for the topology optimization studies.

The radiated pressure within the truncated acoustic domain of the array subset is directly obtained by solving the finite element equations. However, in most situations, the point of interest is outside the small acoustic domain of the array subset. The pressure radiated by the array is calculated by propagating the computed pressure on the semicircular truncation boundary using the following (Lin *et al.*, 1997):

$$p(r,\theta) = \frac{2}{\pi a}\sum_{n=1}^{\infty}\frac{H_n(kr)}{H_n(ka)}\int_{\Gamma_t}\sin n\theta \sin n\theta' p(a,\theta')d\Gamma_t(\theta'), \tag{5}$$

where a is the radius of the truncation boundary Γ_t and H_n is the Hankel function of the first kind.

III. TOPOLOGY OPTIMIZATION

A. Density-based topology optimization

Topology optimization seeks to determine an optimum material layout within a specified region (design domain) of a structure so as to minimize/maximize the objective function (Bendsoe, 1995). The density-based design approach is utilized in this work. It is simple and suits the multidiscipline optimization design. In this approach, the volume density of material in each finite element within the design domain is the design variable. An empirical relationship is assumed between material properties and volume density.

Let the design variable be represented by h_i, where $i=1,2,...,$ndv, and ndv is the total number of design variables. Each design variable ranges from h^l to 1, where h^l is a positive lower bound. The lower bound avoids singularities in the finite element stiffness matrix. The element is "void" if $h_i=h^l$ and is "full" if $h_i=1$. In this work, h^l is set as 10^{-3}. Our experiments show that small changes in this number do not affect the final design. The design problem is to find the optimal amount of material h_i in each finite element so that the objective function is minimized or maximized. The finite element mesh remains unchanged during the optimization procedure, making it much easier to calculate the sensitivities of the objective function and constraints with respect to the design variables.

Since the kerf fillings in transducer arrays are generally isotropic and elastic materials, the material properties in an element with an intermediate density can be approximated by a simple artificial material model [e.g., Yang and Chuang (1994); Silva (1998)],

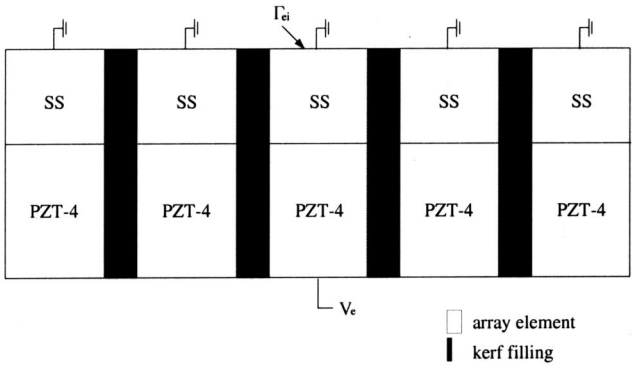
$$\mathbf{E}_i = h_i\mathbf{E}_0, \quad \rho_i = h_i\rho_0, \quad \nu_i = h_i\nu_0, \tag{6}$$

where $\mathbf{E}$ is the elastic Young's modulus, ρ is the mass density, and ν is the Poisson's ratio. The subscripts 0 and i represent the original material and intermediate material, respectively. Allowing for intermediate densities during optimization iterations and the approximate material model greatly simplify the design process. Note the final optimized design only has "full" material or "void" (i.e., only has a real material property).

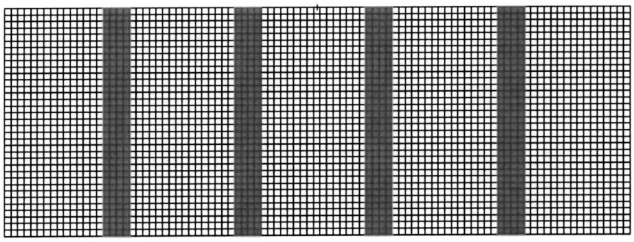
In most cases, the converged optimization solution has some intermediate densities. Since it is hard to interpret the complex, intermediate densities as a real structure, an interpretation procedure based on the image processing (Kikuchi *et al.*, 1997) is introduced. First, the optimized density configuration is transformed into sufficiently fine bitmap scales called pixels with gray-scale values between 0 and 255. The gray-scale values in each pixel are obtained via the interpolation of gray-scale functions using the moving least-squares method. Then an appropriate threshold is selected to obtain the final digital image of the real structure. That is, if the gray-scale value is larger than the threshold, the pixel lies within the real structure. Otherwise, it does not. Since the interpretation process, the final optimized design must be validated. The digital image of the structure is then transformed back to the finite element model for verification of the optimized design.

B. Array subset model and the objective function

An array subset consisting of five elements in which only the middle element is excited is used for the design of kerf fillings. Since the design in a fluid-loaded setting mixes the structure and fluid-borne cross talk, the latter of which is difficult to eliminate, an array subset model without any fluid loading provides a more sensible framework for design. In addition, our previous study of cross talk indicates that the cross talk through the kerf fillings is dominant compared to the acoustical path (Lin and Grosh, 2000b). Therefore, an *in vacuo* array subset model is used, targeting the kerf filling design and removing the acoustical path from the optimization process. The effectiveness of this strategy will be evaluated later in a fluid-loaded setting. The design wireframe and finite element mesh for the array subset are shown in Fig. 3, where each array element consists of a PZT-4 block and a stainless steel layer. The array subset is mechanically free and grounded on the wet side (when the array is in contact with the fluid). On the other side, the middle element is driven by a unit voltage V_e while the others are electronically passive.



(a) Design wireframe of the *in vacuo* model



(b) Finite element mesh

FIG. 3. Design wireframe and finite element mesh of the *in vacuo* five-element array, where SS is the stainless steel layer.

Cross talk in imaging arrays has been extensively studied and is usually represented by the ratio of the induced electric voltage on the adjacent nondriven element to the excitation voltage on the driven element (Kino and Baer, 1983; Lerch, 1990; Turnbull and Foster, 1992). However, using an objective that seeks to minimize the induced electric voltage with a fixed V_e may not give the best design for therapy arrays. In view of the analysis in Eq. (2), one can see that the induced mechanical displacements on elements short-circuited (zero voltage) clearly give the radiated sound associated with the interelement cross talk. For therapy, while the induced voltage on the other elements does not directly imply radiation from these elements, any design effort placed on reducing the induced voltage on the electronically passive elements simply results in lower overall response. Initial optimization experiments (not shown in this paper) seeking to minimize the induced electric voltage on adjacent elements have demonstrated this. Hence, while such an optimization may reduce the induced voltage, the final design results in a very poor sound radiator. Therefore, a more effective objective function to reduce cross talk in therapy arrays was indicated.

In general, an element for a therapy array is designed individually to vibrate at a thickness mode at the operating frequency, outside of the array setting. Therefore, the normal volume velocity on the wet surface, which is proportional to the acoustic power in light acoustic medium, such as air, is the maximum at that frequency. However, when elements are connected together by filling passive materials into the kerf, the normal velocity response decreases due to interelement cross talk and from the connection stiffness. An objective function to maximize the normal velocity response of the driven element was proposed. Specifically, the integral of normal velocity squared along the surface of the driven element face was used as the measure of the response:

$$\mathcal{F}=\int_{\Gamma_{ei}}|v_{ey}|^2 d\Gamma, \tag{7}$$

where Γ_{ei} is the wet surface of the excited element, and v_{ey} is the normal velocity on Γ_{ei}.

While maximizing the volume of normal velocity on the driven element is the objective function used in this study, other objective functions were also considered, such as to combine the minimization of the induced volume velocity on the adjacent nondriven element

$$\mathcal{F}=\frac{\int_{\Gamma_{ei}}|v_{ey}|^2 d\Gamma}{\int_{\Gamma_{nei}}|v_{ney}|^2 d\Gamma}, \tag{8}$$

where Γ_{nei} is the wet surface of the first adjacent nondriven element and v_{ney} is the normal velocity on that surface. Our experience with the combined objective function in Eq. (8) showed that it gave little improvement compared with the single objective function in Eq. (7) while increasing the computing complexity. Therefore, the single objective function to maximize the volume of normal velocity on the driven element was chosen for the design.

C. Topology design schemes

The design wireframe and typical finite element mesh for the *in vacuo* array subset are shown in Fig. 3. The entire kerf region is the design domain. To maintain array periodicity and for manufacturing convenience, identical kerf filling topology in each segment is required. The imposition of this constraint considerably reduces the design variables, thereby the computational cost. Two different design schemes are developed.

In the first design scheme (see Fig. 4), consider each kerf region identically discretized into N finite elements, then there are N design variables. The density of the corresponding element in each kerf segment changes simultaneously during the optimization iterations. We call this design the element-by-element design, a modification of the conventional approach by the linkage of kerf fillings in each segment.

The second design scheme is shown in Fig. 5. It is called the layer-by-layer design. In this design scheme, each kerf region is discretized into N finite elements with N_h layers along the thickness and N_w elements per layer ($N=N_hN_w$), just as in the element-by-element approach. Nevertheless, more constraints are imposed which require that the elements in each layer have the same density. The total number of design variables reduces to N_h.

The general formulated optimization problem for both design schemes can be stated as:

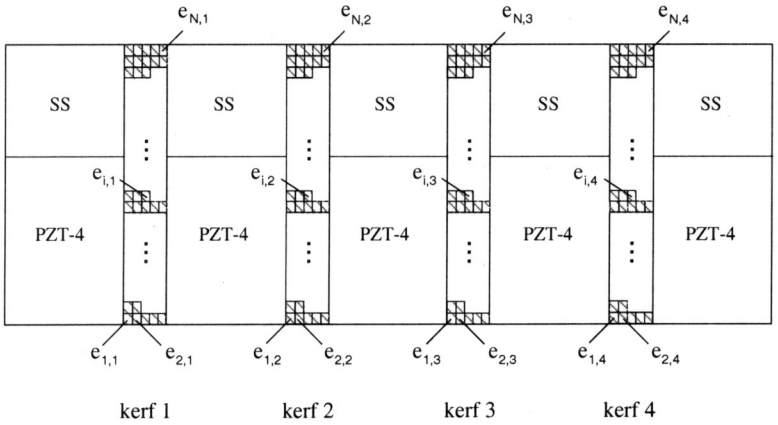


FIG. 4. Topology design scheme one: element-by-element design, each kerf segment has an identical topology and each element density may vary.

Maximize: $\mathcal{F}(\mathbf{h})$ $(\mathbf{h}=[h_1,h_2,\ldots,h_{\mathrm{ndv}}])$

Subject to: $\hat{\mathbf{A}}\mathbf{U}=\mathbf{F}$

$$\left(0<h^l\leqslant h_i\leqslant 1;\ \ \Omega^l\leqslant\int_{\Omega_D} h\,d\Omega\leqslant\Omega^u\right), \tag{9}$$

where ndv equals the number of finite elements in each kerf segment in the element-by-element design and the number of layers along the thickness in each kerf segment in the layer-by-layer design, respectively. Ω_D is the design domain, Ω^l and Ω^u are, respectively, the lower bound and upper bound of the volume constraint, which are determined by the filling material and manufacturing process. The lower bound ensures that the kerf fillings have the required stiffness to support and stabilize the array element while the upper bound controls the amount of material used for filling the kerf. The finite element matrix, $\hat{\mathbf{A}}$, is the condensed and invertable version of the global finite element matrix $\mathbf{A}$ [see Eq. (1)], where the equations associated with the degrees of freedom on which electrical essential boundary conditions are imposed (i.e., the prescribed voltage on the electrodes) are eliminated. The forcing associated with these boundary conditions is held in the vector $\mathbf{F}$ ($\mathbf{F}$ is formed by multiplying the nodal values of the prescribed voltage with the global finite element matrix).

D. Sensitivities and sequential linear programming

In general, topology optimization problems are highly nonlinear with respect to the design variables. Numerous local minima and maxima of the objective function may exist within the design domain. One has to resort to approximate linearization techniques to solve the problem. One widely used category of these techniques is the mathematical programming method. It is flexible and theoretically can deal with any objective function and constraints (Yang and Chuang, 1994; Silva, 1998). The sequential linear programming method is used to solve this complex nonlinear problem. It consists of solving sequential approximate linearized subproblems that can be obtained by the Taylor series expansion of the objective and constraint functions around the current design variables at each iteration step (Vanderplaats, 1984). One great advantage of this method is that it can handle a large number of design variables. The commercial package DSPLP from the SLATEC library (Hanson and Hiebert, 1981) is integrated with the finite element code to do the optimization. The linearization at each iteration step requires the sensitivities of the objective function and constraints with respect to the design variables. They form the key basis of the optimal design.

The sensitivities of the objective function to the design variables are

$$\frac{\partial\mathcal{F}}{\partial\mathbf{h}}=\int_{\Gamma_{ei}}\left(\frac{\partial v_{ey}}{\partial\mathbf{h}}v_{ey}^{*}+\frac{\partial v_{ey}^{*}}{\partial\mathbf{h}}v_{ey}\right)d\Gamma. \tag{10}$$

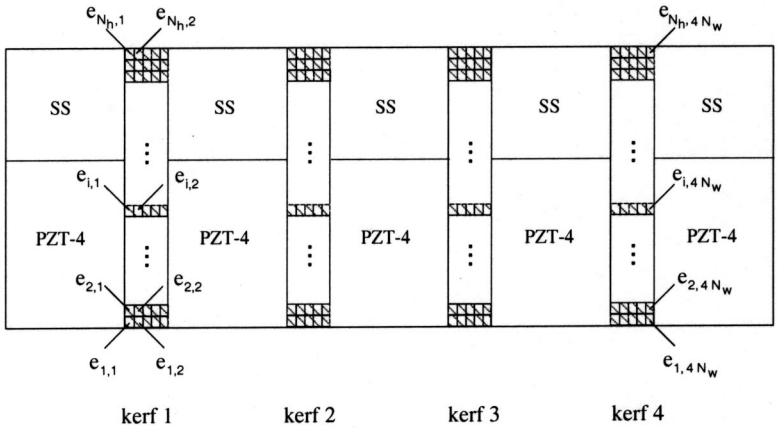


FIG. 5. Topology design scheme two: layer-by-layer design, each kerf segment has an identical topology and a horizontally stratified design is achieved through the layerwise density variations.

Since $v_{ey}=(-i\omega)u_{ey}$, where u_{ey} is the normal displacement, the derivative of the normal velocity on the excited element can be obtained by computing the sensitivities of the normal displacement. Differentiation of the discrete matrix equations in Eq. (9) with respect to the design variables **h** and inverting yield

$$\frac{\partial \mathbf{U}}{\partial \mathbf{h}} = -\hat{\mathbf{A}}^{-1}\frac{\partial \hat{\mathbf{A}}}{\partial \mathbf{h}}\mathbf{U}. \tag{11}$$

Note $\partial\hat{\mathbf{A}}/\partial\mathbf{h}$ has nonzero entries only for the element matrices in which the design variable is defined. In the element-by-element design scheme, there are N design variables since each kerf segment consists of N finite elements. The $\partial\hat{\mathbf{A}}/\partial h_i$, $i=1,2,...,N$, is achieved through the assemblage of sensitivities of corresponding local element stiffness matrices $\partial\hat{\mathbf{A}}_j^e/\partial h_i$, $j=1,2,...,4$, in each kerf segment of the array subset.

On the other hand, in the layerwise design, there are N_h design variables since each kerf segment has N_h layers through the thickness. The sensitivities of the stiffness matrix to each design variable $\partial\hat{\mathbf{A}}/\partial h_i$ are acquired by the assemblage of $4N_w$ local element stiffness sensitivities $\partial\hat{\mathbf{A}}_j^e/\partial h_i$, $j=1,2,...,4N_w$, where N_w is the number of finite elements per layer in each kerf segment.

In both design schemes, solving Eq. (11) requires the solution of a complex linear system with multiple right-hand sides. Efficient solution of this problem involves factorization of the coefficient matrix $\hat{\mathbf{A}}$ once and the sensitivities $\partial\mathbf{U}/\partial\mathbf{h}$ are then obtained by forward and backward substitutions (Golub and Van Loan, 1989). Algorithms for the sensitivities are computed in closed form and included in the finite element code.

IV. NUMERICAL RESULTS

Numerical experiments using the above-given topology design strategies are presented in this section. The array subset used for the kerf filling design is shown in Fig. 3, consisting of five elements. Each element is composed of a PZT-4 block which is 1.5 mm wide and 2.1 mm thick, and a 1.4-mm-thick stainless steel layer attached on the wet side. All the elements are grounded on the wet surface and mechanically free. Note that each single element is designed to have a distinctive thickness mode with maximum acoustic power at 500 kHz by using the design approach developed in a separate paper (Lin and Grosh, 2000a). The center-to-center spacing between elements is 2.0 mm, which is about 0.67λ, where λ is the acoustic wavelength in water at 500 kHz. Material properties for araldite epoxy which has a Young's modulus of 4.24 GPa, Poisson's ratio of 0.38, and density of 1160 kg/m^3 (Mills and Smith, 1996) are used for the kerf filling material. To compensate the internal transducer loss, a simplified Rayleigh damping, which is proportional to the structural stiffness, is applied (Lerch, 1990). The structural damping factors which correspond, respectively, to the reciprocal of mechanical Q of PZT-4 (400), stainless steel (250), and araldite epoxy (100) are used in the numerical model. In the finite element mesh for the array subset, each kerf segment is discretized into 175 finite elements, 5 elements along the width and 35 elements along the thickness.

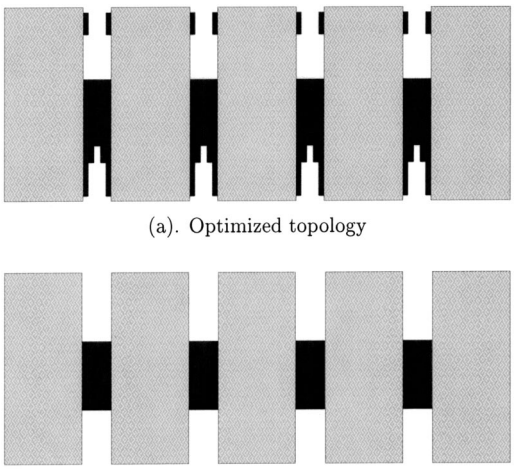


FIG. 6. Optimized kerf filling topology by the elementwise design scheme.

A. Optimized topology by the element-by-element design scheme

First, topology optimization is undertaken using the element-by-element design scheme. In this scheme, the volume density in each finite element in the design domain is the design variable. The total number of design variables is 175. At each iteration, the move limit, which is the allowable change of design variables, is set as 15% of original values of the design variables.

Since the number of design variables is very large, the objective function is highly nonlinear and not well-behaved. Many local optimal solutions exist in this enormous design space. As initial topology and volume constraint vary, optimization results may also change. Though it is desired to have more filling materials in the kerf to ensure the strong connections between array elements, optimization results were not successful when the lower bound of volume constraint is higher than 45%. The optimized topology on the whole was very obscure, dominanted by regions of intermediate densities which were difficult to interpret using image based methods. Also the optimal solutions were overall stuck in the very vicinity of the initial design. Improved optimization results were obtained when the lower bound was less than 45%, such that less than half of the kerf was filled with material.

Figure 6(a) shows the optimized topology of kerf fillings when the initial density is uniformly set to 45% and the lower bound and upper bound of the volume constraint are 30% and 65%, respectively. One can see that a clear trend of the optimization is to break off the connections between elements (interelement coupling) by reducing the material densities either around the centerline of the kerf or close to the array element. Similar trends are observed in other elementwise designs with low volume constraint. Finally, to simplify the optimized topology in Fig. 6(a), the upper and lower epoxy materials are taken away. The final optimized kerf fillings are shown in Fig. 6(b), where the epoxy filled about 35% of the whole kerf.

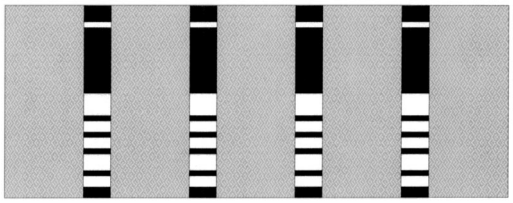


FIG. 7. Optimized kerf filling topology by the layerwise design scheme.

B. Optimized topology by the layer-by-layer design scheme

The layer-by-layer design is motivated by the desire for continuous connections (kerf fillings) between array elements as well as manufacturing convenience. In this design scheme, the volume density along each layer is the design variable. Since the kerf is discretized into 35 elements throughout the thickness, the number of design variables reduces to 35, considerably reducing the optimization complexity.

With significant reduction of the number of design variables using the layerwise approach, the optimization problem was found to be much better behaved. It was found that the optimization solution becomes less sensitive to the selection of design parameters and distinct topology of kerf fillings which has much fewer intermediate densities is achieved when more materials are filled in the kerf.

A representative of the optimized kerf filling topology is displayed in Fig. 7, where the initial density is uniformly 0.7 and the lower and upper bound of the volume constraint are, respectively, 60% and 100%. Usually, the topology optimization gives some intermediate densities, with this set of design parameters, a perfect black–white structure consisting of exclusively solid material and ''void'' is achieved directly from the optimization. Since there are no intermediate densities the image based processing was not needed for this case. The total volume percentage of the kerf fillings is 60% in the final optimized design. A similar optimized topology with distinctive black-white structure is also obtained when the initial density is 0.6 and the lower bound and upper bound of the volume constraint are 50% and 100%, respectively. Only one middle layer has an intermediate density of 0.53 and the position of middle discrete layers is a bit different from the topology shown in Fig. 7. The effect of these middle epoxy layers on the acoustic radiation will be discussed in Sec. IV C 2, when evaluating the performance of the array. More gray-scale densities emerge when other design parameters are used. Since the enhancement of other designs is not as great, they are not presented here.

In both design schemes (elementwise and layerwise), we note that sound optimization results are usually attained near the lower bound of volume constraint, exhibiting the tendency to weaken the interelement connections. The upper bound has much less control over the optimization search, especially when a wide volume bound is imposed. In array fabrications, it is desired that the kerf be filled with more materials to ensure the required support and stabilization of array elements. In order to reduce cross talk, the optimization process tends to eliminate the fillings. The topology optimization process helps in directing the designer where to put the material, something which is not known *a priori*. Thence, we recommend the lower bound of volume constraint set the lowest limit of material needed to ensure structural stability.

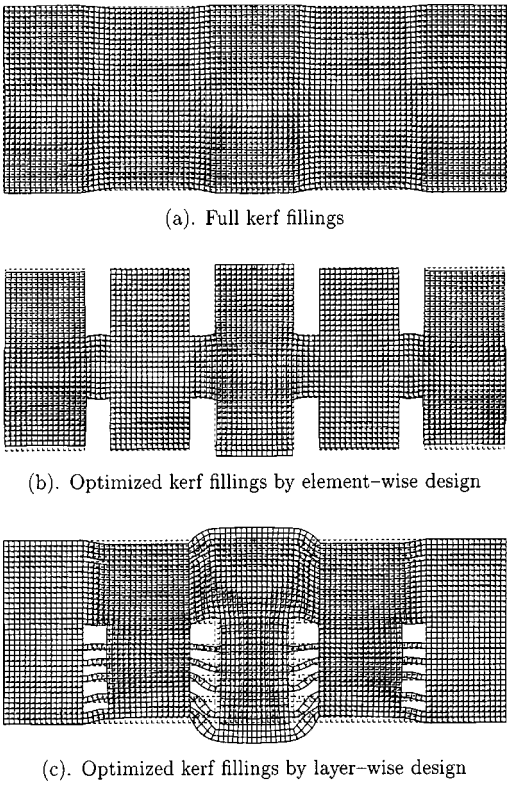


FIG. 8. Displacement field of the five-element array at 500 kHz when only the middle element is driven. The solid line is the deformed shape and the dotted line is the undeformed shape.

C. Evaluation of optimized kerf fillings in the array subset

1. Single element performance in the array subset

To evaluate the performance of optimized kerf filling topology obtained by the above-mentioned design schemes, the real displacement field and acoustic beam pattern of the five-element array subset when only the middle element is driven are first examined. The deformed array subset with different kerf filling topology is displayed in Fig. 8. One can see a dramatic improvement of the displacement field with the layerwise optimization approach [Fig. 8(c)]. The driven element vibrates the most compared with the neighboring elements. The elementwise design also results in reduced cross talk compared to full kerf filling, though not as remarkable as that with the layer-by-layer design.

The corresponding acoustic field is presented next. The far-field acoustic pressure at a radius of 60 mm from the center of the five-element array is shown in Fig. 9. Relative to an ''idealized'' array with no kerf fillings, the array with full kerf fillings performs poorly, with a 10 dB reduction of the pressure level and greater angular variations, a result arising from the severe structural cross talk between elements. However, with the optimized kerf filling topology, the acous-

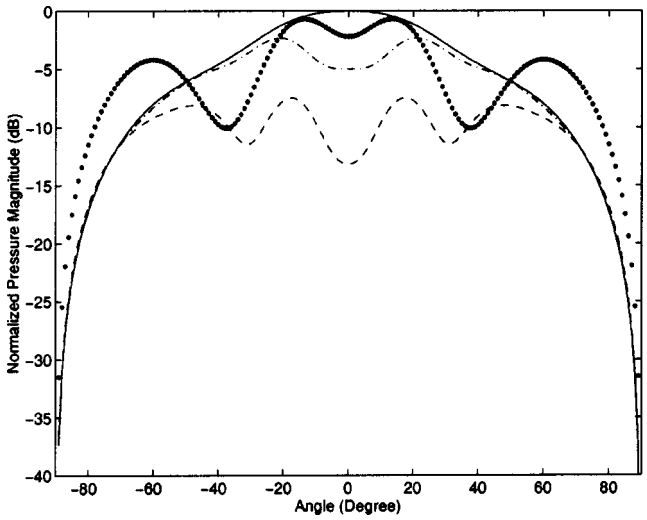


FIG. 9. Far-field acoustic pressure of the array subset when only the middle element is excited: (—) no kerf fillings, (— — —) full kerf fillings, (-·-) optimized kerf fillings by elementwise design, (···) optimized kerf fillings by layerwise design.

tic field is improved. The pressure magnitude with the elementwise optimized kerf fillings is increased by about 8 dB over the full kerf design on the central axis and the main lobe is much flatter. Even stronger acoustic response is seen in the layerwise design with about 6–11 dB improvement of the pressure magnitude within the steering angle of 20°.

2. Steering of the array subset

The steering of the array focus, which is very important for array applications, is also evaluated. The acoustic field when five elements are excited by phased voltage is demonstrated in Fig. 10. Correlated with the response of single element, significant improvement is attained by the array subset with layerwise optimized kerf fillings when it is steered within an angle of 20° as compared to the full kerf filling design. An increase of more than 8 dB is seen in the layerwise optimized design and about 3–5 dB improvement with the elementwise optimized design is obtained, relative to the fully filled design. At greater steering angles, around 30°, the performance of the two optimized designs is comparable [see Fig. 10(c)].

In addition, the effect of the middle discrete epoxy layers (Fig. 7) in the layerwise optimization design on the array performance was studied by combining the middle discrete layers or taking some layers out. It was observed that the radiated acoustic field is fairly insensitive to these changes. The variation of the acoustic pressure around the focus is usually less than 9%.

In summary, the maximization of the normal velocity response achieved through an *in vacuo* optimization correlates to improved acoustic radiation in the fluid-loaded case in both design schemes. The layerwise optimized kerf filling design (Fig. 7) achieves the greatest improvement of acoustic radiation for the trails preformed in this study for steering angles less than 30°.

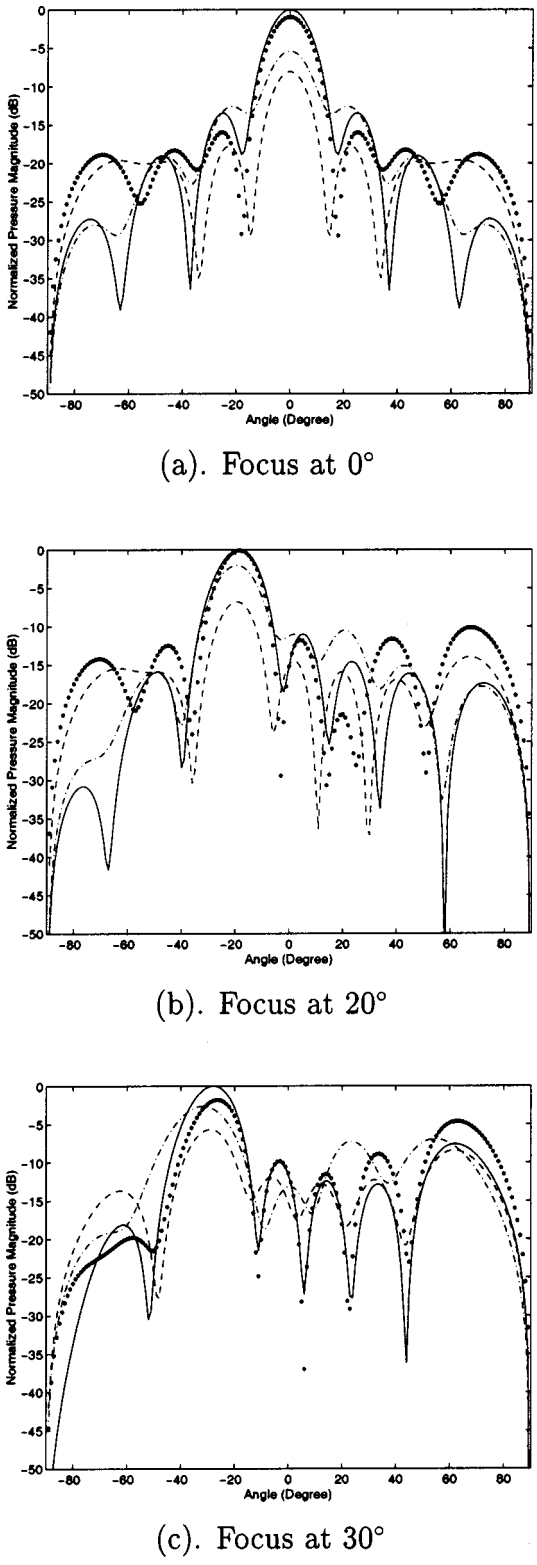


FIG. 10. Far-field acoustic pressure of the five-element array: (—) no kerf fillings, (— — —) full kerf fillings, (–·–) optimized kerf fillings by elementwise design, (···) optimized kerf fillings by layerwise design.

D. Acoustic field of large linear arrays

To validate the effectiveness of optimized kerf fillings on arrays with more elements, the modeling technique described in Sec. II is approximated to calculate the radiation by large linear arrays. First, a 9-element array provides the

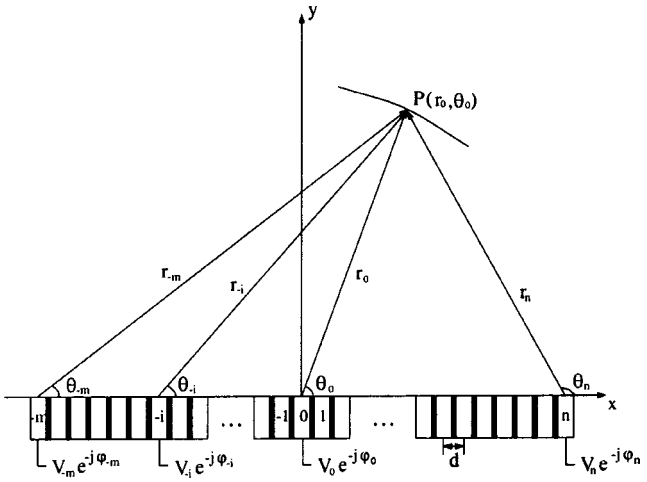
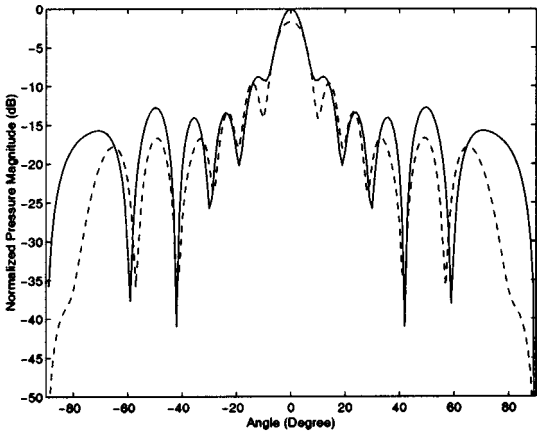
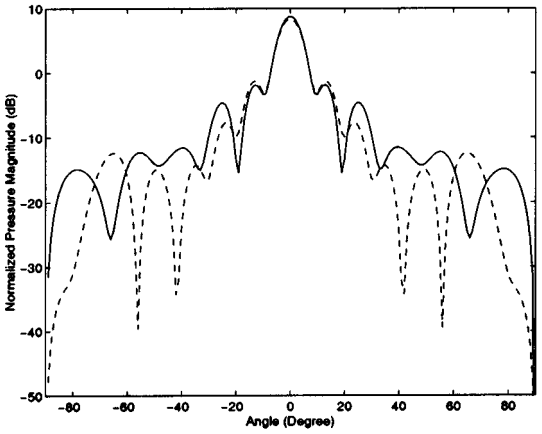


FIG. 11. Calculation schematic of the acoustic field of a linear phased array with $m+n+1$ elements.

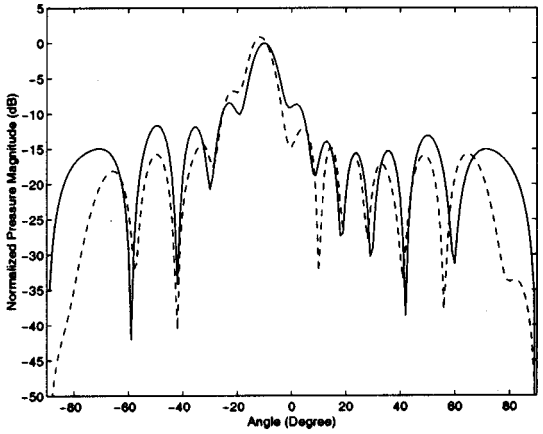


(a). Full kerf fillings

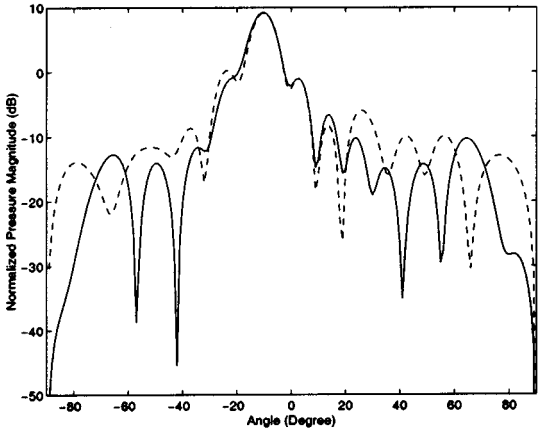


(b). Layer-wise optimized kerf fillings

FIG. 12. Far-field acoustic pressure of the nine-element array focusing at 0°: (—) direct modeling, (— — —) synthesis of the five-element array subset. All the pressure magnitude is normalized with respect to the pressure on the focus of direct modeling with full kerf fillings.



(a). Full kerf fillings



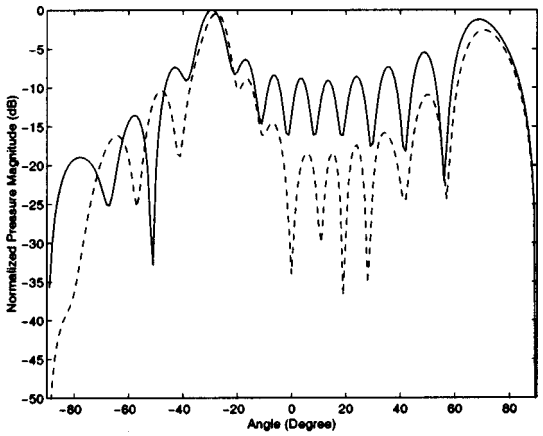
(b). Layer-wise optimized kerf fillings

FIG. 13. Far-field acoustic pressure of the nine-element array focusing at 10°: (—) direct modeling, (— — —) synthesis of the five-element array subset. All the pressure magnitude is normalized with respect to the pressure on the focus of direct modeling with full kerf fillings.

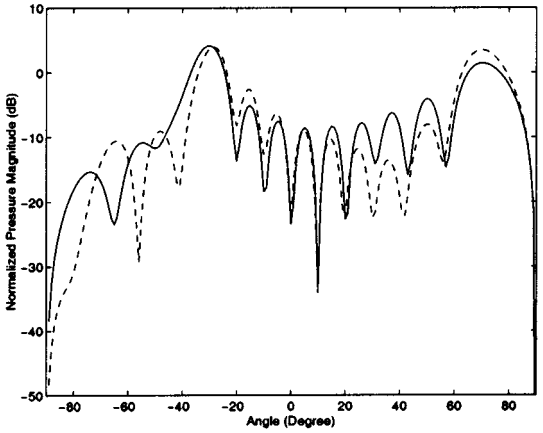
setting to describe and test the approximate technique followed by modeling of a fifty-one-element array.

1. Subset array modeling

The direct modeling of an entire array can be prohibitively expensive, from a computational and storage perspective. In this study, an array subset synthesis technique is utilized (Assaad and Bruneel, 1996). Instead of modeling an entire array, the array response is approximated by synthesizing the response from a subset of a complete array in which only the middle element is driven and other elements are short-circuited. In addition to using this technique for modeling a fifty-one-element array, determination of the accuracy of this approximate approach is presented. In Fig. 11, an array with $N=n+m+1$ elements is shown. Assad's synthesis technique involves the simulation of a fluid-loaded subset array of $M<N$ elements. The same boundary conditions as described in Sec. II are used for the subset. The full array is simulated by moving the center element of the subset array to each driven element. The benefit of this approach is that a large array can be modeled with only a few elements.

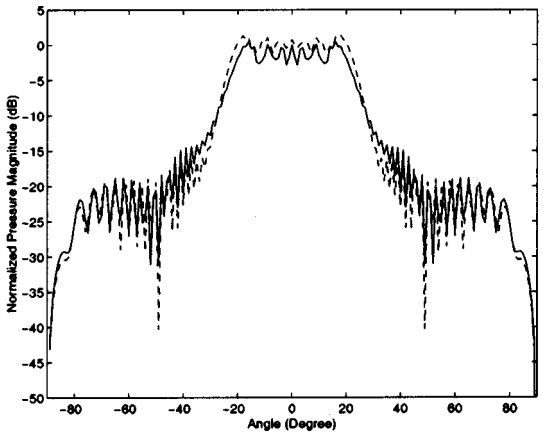


(a). Full kerf fillings

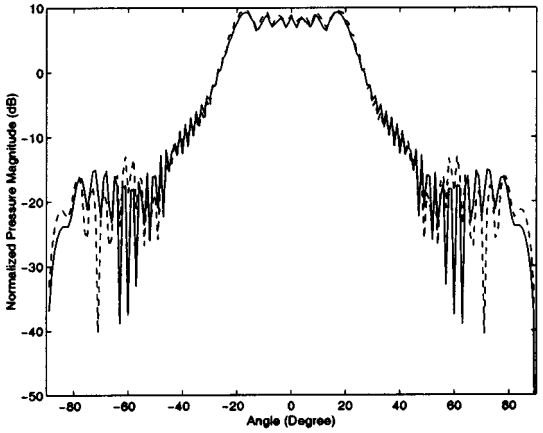


(b). Layer-wise optimized kerf fillings

FIG. 14. Far-field acoustic pressure of the nine-element array focusing at 30°: (—) direct modeling, (— — —) synthesis of the five-element array subset. All the pressure magnitude is normalized with respect to the pressure on the focus of direct modeling with full kerf fillings.



(a). Full kerf fillings



(b). Layer-wise optimized kerf fillings

FIG. 15. Far-field acoustic pressure of the fifty-one-element array focusing at 0°: (—) synthesis of the five-element array, (— — —) synthesis of the nine-element array. All the pressure magnitude is normalized with respect to the pressure on the focus of five-element synthesis with full kerf fillings.

The accuracy of the approach is determined by performing the simulations with increasing subset size and examining the attendant change in predicted response with an increase in the subset size.

Therefore, based on the subset synthesis, the acoustic pressure at point $P(r_0,\theta_0)$ (Fig. 11) can be approximated by

$$p=\sum_{i=1}^{N} V_i e^{-j\varphi_i} p_i(r_i,\theta_i), \tag{12}$$

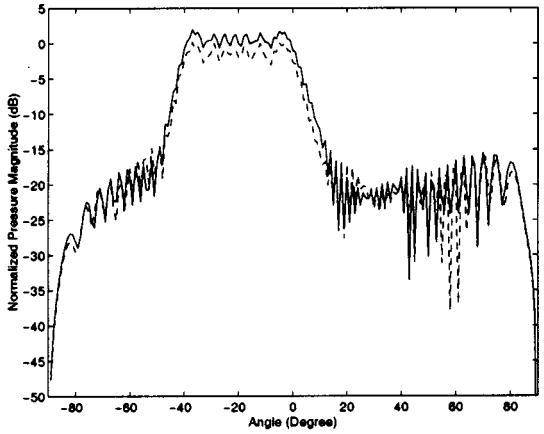
where

$$r_i=[(x_0-id)^2+y_0^2]^{1/2}, \quad \theta_i=\arctan\frac{y_0}{(x_0-id)}.$$

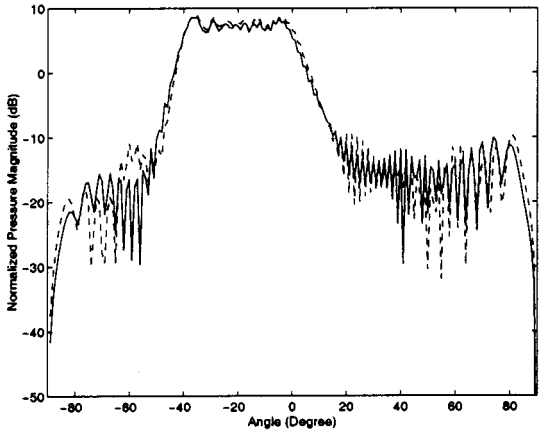
Here $p_i(r_i,\theta_i)$ is the acoustic pressure generated by the ith array subset in which only the middle element (the ith element) is driven by a unit voltage, d is the center-to-center spacing between elements and x_0, y_0 are the Cartesian coordinates of point P. When $P(r_0,\theta_0)$ is within the truncated acoustic domain of the array subset, the pressure at P radiated by the subset can be directly obtained by solving the finite element equations. However, in most situations, the point of interest is outside the small acoustic domain of the array subset. Hence, the far-field projection algorithm is used to determine the contribution of each array subset to the total far-field pressure.

2. Accuracy of five-element array subset synthesis for the acoustic field of nine-element array

A finite element model for a water-loaded nine-element array is built, where the dimensions for the array element and kerf as well as mesh resolutions are the same as those of the five-element array subset. The nine-element array enables the determination of both the effectiveness of the kerf filling design and the accuracy of the subset synthesis technique, because the nine-element array can be directly simulated. The acoustic field obtained by direct modeling of the nine-element array and that by the five-element array subset synthesis, focusing, respectively, at 0°, 20°, and 30°, are compared in Figs. 12–14. In general, good correlation between these two modeling approaches is obtained around the focus region, with full kerf fillings or layerwise optimized kerf

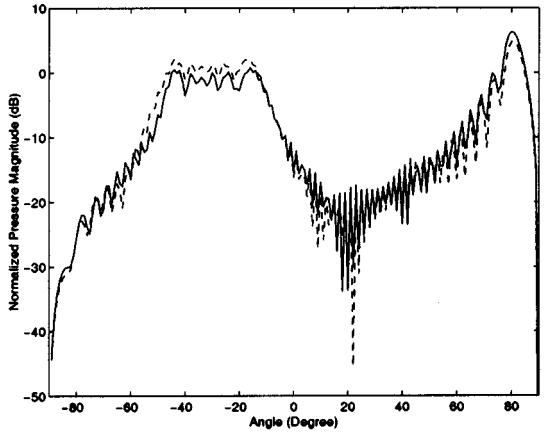


(a). Full kerf fillings

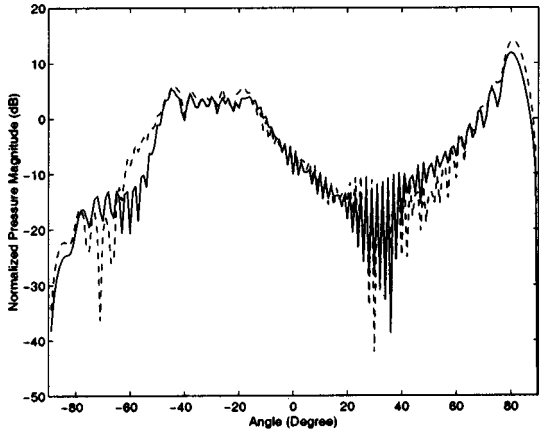


(b). Layer-wise optimized kerf fillings

FIG. 16. Far-field acoustic pressure of the fifty-one-element array focusing at 10°. (—) synthesis of the five-element array, (— — —) synthesis of the nine-element array. All the pressure magnitude is normalized with respect to the pressure on the focus of five-element synthesis with full kerf fillings.



(a). Full kerf fillings



(b). Layer-wise optimized kerf fillings

FIG. 17. Far-field acoustic pressure of the fifty-one-element array focusing at 30°: (—) synthesis of the five-element array, (— — —) synthesis of the nine-element array. All the pressure magnitude is normalized with respect to the pressure on the focus of five-element synthesis with full kerf fillings.

fillings. The main discrepancy is seen in the low pressure region where the angle is far away from the focus.

Two sources contribute to the discrepancies in the array subset synthesis. The first source, which is the most important one, is the accuracy of the five-element subset as the basis for synthesis. The second source of the error is the edge effect that is not accounted for in the synthesis. The edge effect occurs in the end few elements of the array and will decrease with the increase of number of elements in a large array. In general, the five-element array subset synthesis is sufficient when the acoustic field around the focus is of main concern and the steering angle is not large. Higher accuracy away from the focus can be achieved by increasing the number of elements in the array subset.

The main goal of this work is the development of effective design strategies for the kerf fillings in arrays. Most importantly, the comparison of the radiation by a nine-element (Figs. 12–14) to that of the five-element array (Fig. 10) clearly demonstrates the effectiveness of the topology design approach on the array subset. That is, the enhancement of acoustic radiation from the five-element array with optimized layerwise design is effectively carried to that of the nine-element array. An increase of about 3–10 dB of the acoustic pressure around the focus is found in the nine-element array with optimized design, very similar to the acoustic field enhancement in the five-element array. A further characterization of the effectiveness of the optimized design in a large linear array setting is conducted next.

3. Acoustic field of a fifty-one-element array

Using the above five-element and nine-element subarray as the basis of synthesis, the acoustic field of a fifty-one-element array with fully filled kerf and layerwise optimized kerf fillings, steering at angles from 0° to 30°, is shown in Figs. 15–17. The radius of the acoustic field is 120 mm away from the center of the array. Overall, very good agreement with these two subsets is attained, corresponding to the above-given comparisons for the nine-element array. With more elements in the array, discrepancies between the five-element synthesis and nine-element synthesis are decreased. Note larger discrepancies are seen on the main beam with full kerf fillings than those with layerwise kerf fillings. Once again, a remarkable increase of the acoustic pressure around

the focus of about 10 dB is observed when the array with optimized kerf fillings steers within the angle of 20°. When the steering angle is at 30°, there is still about a 4 dB increase of the pressure magnitude around the focus. Note due to the large center-to-center spacing between elements ($>\lambda/2$, where λ is the acoustic wavelength at 500 kHz), a grating lobe appeared when the array steers at 30°.

V. CONCLUSIONS

The density-based topology design approach is utilized to improve the design of kerf fillings in linear phased arrays to enhance the acoustic radiation for therapy applications. An *in vacuo* array subset is used as the design model. Of several possible approaches (and attendant optimization cost functions) for reducing the interelement cross talk, a cost function that maximizes the mechanical response of a driven element (as opposed to minimizing the electrical response of nondriven elements) was found most effective. Two different topology design schemes are proposed and formulated (element-by-element and layer-by-layer design schemes). Simulations in a fluid-loaded setting determined that the *in vacuo* optimized kerf filling topology indeed conveyed enhancements to the steering capabilities of the fluid-loaded array and maximum pressure level around the array focus. For the designs presented in this paper, the layerwise design approach was more robust and efficient, delivering higher levels of acoustic radiation while reducing the optimization complexity. The accuracy of the array subset design is verified. For the more complicated topologies arising from the optimization procedure, the manufacturing process is more complicated. However, with the advance of multilayer fabrication methods, the layerwise optimization design is feasible. The topology optimization techniques presented here provide the designer with a unique tool to steer the process in a direction that otherwise might not have been envisioned from *a priori* considerations. While a powerful tool for guiding the designer, the converse is also true, that designer's experience and insight, which constrain the design space, are also essential for the effectiveness of optimization process.

ACKNOWLEDGMENTS

This work was supported by the Whitaker Foundation and the Office of Naval Research. We would like to thank Professor Noboru Kikuchi for his assistance with the topology design methodology.

Assaad, J., and Bruneel, C. (**1996**). ''Influence of interaction between elementary transducer on directivity pattern of an array,'' Ultrasonics **34**, 107–110.

Bendsoe, M. P. (**1995**). *Optimization of Structural Topology, Shape, and Material* (Springer, Berlin).

Bihrle, R., Foster, R. S., Sanghvi, N. T., Donohue, J. P., and Hood, P. J. (**1994**). ''High intensity focused ultrasound for the treatment of benign prostatic hyperplasia: Early United States experience,'' J. Urol. (Baltimore) **151**, 1271–1275.

Buchanan, M. T., and Hynynen, K. (**1994**). ''Design and experimental evaluation of an intracavitary ultrasound phased array system for hyperthermia,'' IEEE Trans. Biomed. Eng. **41**, 1178–1187.

Dias, J. F. (**1982**). ''An experimental investigation of the cross-coupling between elements of an acoustic imaging array transducer,'' Ultrason. Imaging **4**, 44–55.

Driller, J., and Lizzi, F. L. (**1987**). ''Therapeutic applications of ultrasound: A review,'' IEEE Eng. Med. Biol. Mag. **38**, 33–39.

Ebbini, E. S., and Cain, C. A. (**1989**). ''Multiple-focus ultrasound phased-array pattern synthesis: Optimal driving-signal distributions for hyperthermia,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **36**, 540–548.

Ebbini, E. S., and Cain, C. A. (**1991**). ''Experimental evaluation of a prototype cylindrical section ultrasound hyperthermia phased-array applicator,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**, 510–520.

Golub, G., and Van Loan, C. F. (**1989**). *Matrix Computations*, 2nd ed. (Johns Hopkins, Baltimore, MD).

Greenstein, M., Lum, P., Yoshida, H., and Seyed-Bolorforosh, M. S. (**1997**). ''A 2.5 MHz 2D array with *Z*-axis electrically conductive backing,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**, 970–977.

Hanson, R., and Hiebert, K. (**1981**). ''A sparse linear programming subprogram,'' Technical report, Sandia National Laboratories.

Kikuchi, N., Hollister, S., and Yoo, J. (**1997**). ''A concept of image-based integrated CAE for production engineering,'' in Proceedings of the International Symposium on Optimization and Innovative Design in JSME, Tokyo, Japan, pp. 75–90.

Kino, G. S., and Baer, R. (**1983**). ''Theory for cross-coupling,'' in Proceedings of the IEEE Ultrasonics Symposium, Vol. 2, pp. 1013–1019.

Kino, G. S., and DeSilets, C. S. (**1979**). ''Design of slotted transducer arrays with matched backings,'' Ultrason. Imaging **1**, 189–209.

Lerch, R. (**1990**). ''Simulation of piezoelectric devices by two- and three-dimensional finite element methods,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**, 233–247.

Lin, Y., Dodson, J. M., Hamilton, J., Kluiwstra, J.-U., Cain, C., and Grosh, K. (**1997**). ''Theory and experiment for the design of piezoelectric elements for phased arrays,'' in Proceedings of the IEEE Ultrasonics Symposium, Toronto, Canada, Vol. 2, pp. 1697–1700.

Lin, Y., and Grosh, K. (**2000a**). ''Design of ultrasonic array elements for acoustic power considerations,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control. **49**, 20–28.

Lin, Y., and Grosh, K. (**2000b**). ''Topology optimization of the element connections in transducer arrays,'' in Proceedings of SPIE Smart Structures and Materials. Mathematics and Control in Smart Structures, Newport Beach, CA.

Mills, D. M., and Smith, S. W. (**1996**). ''Combing multi-layers and composites to increase SNR for medical ultrasound transducers,'' in Proceedings of the IEEE Ultrasonics Symposium, Vol. 2, pp. 1509–1512.

Sanghvi, N. T., Hynynen, K., and Lizzi, F. L. (**1996**). ''New developments in therapeutic ultrasound,'' IEEE Eng. Med. Biol. Mag. **15**, 83–92.

Seyed-Bolorforosh, M. S. (**1995**). ''Design requirements for an integrated acoustic impedance matching layer,'' Electron. Lett. **31**, 1702–1703.

Silva, E. C. N. (**1998**). ''Design of piezocomposite materials and piezoelectric transducers using topology optimization,'' Ph.D. thesis, The University of Michigan.

Turnbull, D. H., and Foster, F. S. (**1991**). ''Beam steering with pulsed two-dimensional transducer arrays,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**, 320–333.

Turnbull, D. H., and Foster, F. S. (**1992**). ''Fabrication and characterization of transducer elements in two-dimensional arrays for medical ultrasound imaging,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 464–475.

Vanderplaats, G. N. (**1984**). *Numerical Optimization Techniques for Engineering Design: with Applications* (McGraw–Hill, New York).

Yang, R. J., and Chuang, C. H. (**1994**). ''Optimal topology design using linear programming,'' Comput. Struct. **52**, 265–275.

Microphone array measurement system for analysis of directional and spatial variations of sound fields

Bradford N. Gover[a)]
Department of Physics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada and Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada

James G. Ryan[b)] and Michael R. Stinson
Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada

(Received 15 May 2001; accepted for publication 30 July 2002)

A measurement system has been developed that is capable of analyzing the directional and spatial variations in a reverberant sound field. A spherical, 32-element array of microphones is used to generate a narrow beam that is steered in 60 directions. Using an omnidirectional loudspeaker as excitation, the sound pressure arriving from each steering direction is measured as a function of time, in the form of pressure impulse responses. By subsequent analysis of these responses, the variation of arriving energy with direction is studied. The directional diffusion and directivity index of the arriving sound can be computed, as can the energy decay rate in each direction. An analysis of the 32 microphone responses themselves allows computation of the point-to-point variation of reverberation time and of sound pressure level, as well as the spatial cross-correlation coefficient, over the extent of the array. The system has been validated in simple sound fields in an anechoic chamber and in a reverberation chamber. The system characterizes these sound fields as expected, both quantitatively from the measures and qualitatively from plots of the arriving energy versus direction. It is anticipated that the system will be of value in evaluating the directional distribution of arriving energy and the degree and diffuseness of sound fields in rooms. [DOI: 10.1121/1.1508782]

PACS numbers: 43.38.Hz, 43.55.Mc, 43.55.Br, 43.60.Gk [SLE]

I. INTRODUCTION

The idea of a diffuse sound field is one that is central to many facets of acoustical research, especially those that deal with measurements or tests carried out in enclosures. Oftentimes, the procedure in question hinges on an assumption of diffuseness, and it is of interest to be able to gauge to what degree that assumption is valid. Additionally, there are instances where it is conceded that the sound field is not diffuse, but there is interest in learning details about the spatial variations. Attempts to address these concerns have led to much discussion and interest in spatial measurement systems and techniques.

To renew the discussion of spatial variations of sound fields, it is worth reviewing the definition of ''diffuseness.'' The ANSI definition[1] of a diffuse sound field is a ''*Sound field in which the time average of the mean-square sound pressure is everywhere the same and the flow of acoustic energy in all directions is equally probable*.'' One implication of this statement is that at any single point in the sound field, the incidence of sound energy is equally likely in all directions. That is, viewed from such a point, the sound field is isotropic. Moreover, a diffuse sound field must also be homogeneous, that is, isotropic at *every* point. These facts can be of use in assessing the quality of diffuseness.

In fact, in attempting to ascertain whether or not a sound field is diffuse, measures of the homogeneity are of great importance. These point-to-point measures can reveal the suitability of diffuse-field assumptions, but do not necessarily shed light on the nature of any deviations from such. In contrast, the ability to measure the degree of isotropy in a sound field—namely, how the incident sound energy varies with direction at a single point—cannot necessarily indicate whether the sound field is diffuse, but can reveal if the sound field is *not* diffuse.

In those instances where one does not necessarily *desire* to have a diffuse field in a room, one would typically like to know from which directions the sound arrives at a point. This, practically by definition, implies that it is the characterization of the isotropy that is of importance.

The system described in this paper was developed out of a desire to shed light on this latter problem—to measure arriving energy with direction at a point. The apparatus is also capable of some analysis of the point-to-point variations in sound fields, with the caveat that the separation of those points is limited by the spatial extent of the device.

II. EXISTING METHODS OF STUDYING SPATIAL CHARACTERISTICS OF SOUND FIELDS

Much work has been done in the area of measurement of spatial characteristics of sound fields in rooms and auditoria. Overviews of early approaches are presented by Schultz[2] and by Cremer and Müller.[3] More recently, cumulative summaries have been given by Abdou and Guy[4] and by Nélisse and Nicolas.[5]

[a)]Electronic mail: bradg@audiolab.uwaterloo.ca
[b)]Present address: Gennum Corporation, 232 Herzberg Rd., Kanata, Ontario K2K 2A1, Canada.

There is a large body of work, both theoretical and experimental, concerned with the point-to-point variations of sound fields in rooms. The most commonly investigated parameters are the point-to-point cross-correlation of pressure,[6–8] and the spatial uniformity of pressure.[5,9]

Approaches that attempt to more directly measure isotropy, in the sense that some quantity that varies with the direction of energy flow at a point is measured, have also been extensively studied. Some of the earliest work was by Thiele.[10] He employed an omnidirectional microphone coupled with a parabolic reflector to study arriving energy for steady-state warble tones in concert halls. It is in this paper that the so-called *directional diffusion of* a sound field is proposed. Meyer and Thiele[11] used a similar apparatus, but chose to present the data in the form of directional ''hedgehogs,'' showing graphically the arrival directions and strengths. By using impulses as sources, they were also able to investigate the early and late sounds separately.

The idea of using an array of microphones was proposed by Schroeder.[12] In his discussion, the microphone positions are distributed on one wall of a rectangular reverberation chamber. By replicating the array in room images, a very narrow beam can be generated with minimal perturbation of the sound field. The practical difficulties associated with this have prevented its successful use. Rather than a planar array dependent on the room symmetry, Broadhurst[13] used a cubic grid of 125 microphone positions to steer a narrow beam in many directions. Limited by hardware considerations to a total measurement time of only 167 ms, early reflections were located and some spatial results were presented for a concert hall and a lecture theater.

Two similar four-microphone techniques were developed by Sekiguchi, Kimura, and Hanyuu[14] and by Yamasaki and Itow.[15] In these approaches, locations are assigned to virtual image sources that correspond to arriving reflections. There are problems with ambiguities as reflections arrive simultaneously, which is certain to occur after a modest amount of time. Using much the same four-microphone geometry, Guy and Abdou[16] developed an instantaneous intensity technique, computing the vector intensity at each time instant. From the point of view of detecting reflection arrivals, problems arise when sound waves simultaneously arrive from more than one direction—the intensity vector, corresponding to net flow, does not necessarily indicate the direction of any particular arrival.[17,18]

The key information that all the existing approaches appear unable to yield is the direction of arriving wave fronts, for every time instant, in a general sound field. It is for this reason that the present system has been developed.

III. MEASURING SYSTEM

Conceptually, and as evidenced by much of the literature, the simplest and most direct way to study the sound arriving at a point is with a directional receiver. For a number of steering orientations, the output response to a given sound input in the room is measured. An analysis of the measured responses allows a characterization of the isotropy at the receiving position. This is the approach taken in this study.

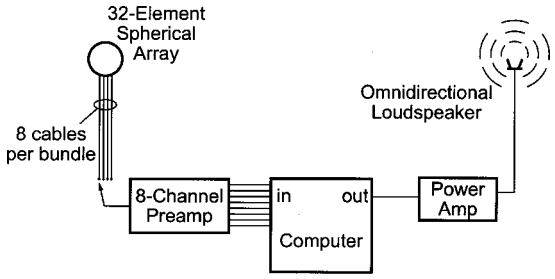


FIG. 1. Block diagram of measurement system. The microphone array picks up the direct acoustic signal from the loudspeaker and all subsequent room reflections.

An overview of the system developed in this work is shown in Fig. 1. An array of 32 omnidirectional microphones is used as the receiver, the microphone outputs being sampled by the control computer eight at a time. While sampling each group of eight microphones, the loudspeaker is simultaneously being driven with a maximum-length sequence (MLS) signal by the same computer. The (omnidirectional) room impulse response to each of the array microphones is determined by post-processing of the recorded signals, and subsequent beamforming generates a directional impulse response for each steering direction. From this set of directional impulse responses, quantitative measures and graphical representations describing the isotropy at the array center are generated. From the set of omnidirectional microphone impulse responses themselves, quantitative measures describing the point-to-point variations of the sound field are computed.

A. The microphone arrays

The microphone array geometry used is shown in Fig. 2. There are 32 microphones, located at the positions of the vertices of a geodesic sphere. The microphones in this arrangement are semiregularly spaced, in that there are two unique nearest-neighbor distances. The high degree of symmetry can be exploited if one chooses the primary steering axis as the direction through the center of the triangle defined by any three neighboring microphones. Then, by simple rotation, the exact same beam can be steered through all 60

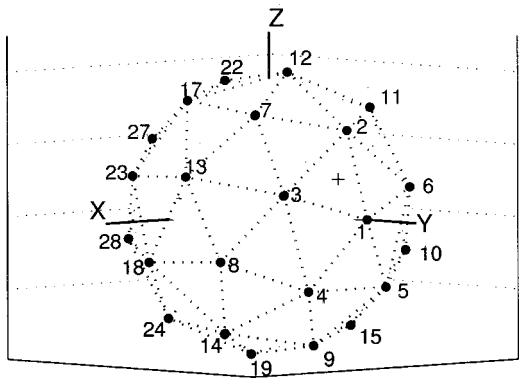


FIG. 2. Array geometry depicting microphone locations (black dots) and the primary steering direction (+). The 32 microphone positions (not all shown) lie at the vertices of a geodesic sphere (shown with dotted lines).

such triangles, and so achieve the identical array response in a total of 60 (albeit not quite evenly spaced) steering directions.

The beamformer design was conducted in two steps. First, the expected array gain was maximized while taking into account a measured microphone mismatch.[19–21] This resulted in an array design that could be implemented with reasonable microphone tolerances, but that had very low white noise gain at particular frequencies. Second, this white noise gain curve was manually edited at these frequencies to ensure a smooth function over the frequency range of interest. This modified white noise gain was then used as a design target in a separate array gain maximization step.[22] The microphone weights resulting from the second optimization became the design target for the individual microphone filters.

Each of the 32 beamformer filters was implemented as a 30-tap FIR filter. The FIR coefficients were calculated using a least-squared error design procedure.[23] These coefficients are too numerous to include in printed tabular form but have been made available through the EPAPS service.[24] The filters are enumerated in such a way that the array output is the response corresponding to the direction between the microphones whose signals are fed through filters #1, #2, and #3. By cyclically permuting the 32 microphone responses, the directional impulse responses in all 60 steering directions are generated from one set of measurements, with no need for actual physical movement of the array.

The beam has a half-power width of about 28°, which is slightly larger than the 22° or 24° angles between adjacent steering directions. All angles are therefore certain to be adequately covered. Note that this beamwidth is approximately the same as for a fourth-order differential microphone.[25] As a comparison, the well-known first-order hypercardioid pattern has a beamwidth of 105°.

The usable frequency range, over which the peak-to-average directional response (the so-called *directional gain*) remains above 14 dB is about 1.7 octaves, so two separate arrays were constructed to allow the system to cover the speech band of about 300–3300 Hz. The high-frequency array has a diameter of 16 cm, intermicrophone distances of 5.1 and 5.7 cm, and an operating frequency range of 1000–3300 Hz. The low-frequency array is three times larger, with a diameter of 48 cm, intermicrophone spacings of 15.4 and 17.1 cm, and an operating frequency range of 300–1000 Hz. The beam properties are shown in Fig. 3 for the high-frequency array and in Fig. 4 for the low-frequency array. In both figures, panel (a) shows the directivity pattern in the plane passing through the primary steering direction, microphone #1, and the center of the array (refer to Fig. 2). This is the wideband pattern for the entire operating range of the array, calculated numerically from the filter coefficients. The asymmetry of the pattern is due to the asymmetry of the array with respect to this plane. Panel (b) shows the directional gain versus frequency, also calculated from the beamforming coefficients, with the operational range indicated by the vertical lines.

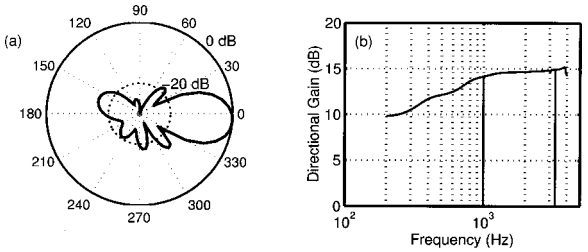


FIG. 3. High-frequency array beam properties, computed from the beamformer coefficients: (a) Beam pattern in the plane that passes through the primary steering direction, microphone 1, and the center of the array (see Fig. 2). This is the wideband beam pattern over the entire operating range of the array (1000–3300 Hz). The steering direction is at 0°, microphone 1 is at 21°. The asymmetry of the pattern is due to the asymmetry of the array with respect to this plane. (b) Directional gain versus frequency. The operational range is indicated by the vertical lines. The breakdown of the performance of the array outside this range is evident by the drop in the curve at lower frequencies.

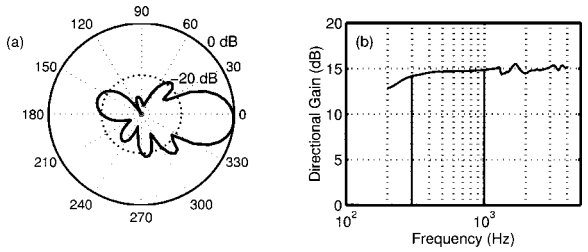


FIG. 4. Low-frequency array beam properties computed from the beamformer coefficients, the same as Fig. 3. (a) Wideband beam pattern over the entire operating range of the array (300–1000 Hz). (b) The performance breaks down at frequencies above the operational range due to spatial aliasing.

A photograph of the arrays is shown in Fig. 5. Notice that the construction of the stands takes advantage of the fact that groups of five microphones have the same *Y* coordinate (refer to Fig. 2). In this photo, microphone 1 is at the left-hand end of the central shaft, the closest set of ''spokes'' carries microphones 2–6, the next set carries microphones 7–12, and so on. Microphone 32 is at the opposite end of the central shaft. This construction simplifies matters, in that there is no need to actually construct the geodesic sphere, and has the added advantage of greater acoustic transparency, since there is much less scattering material in close proximity to each microphone capsule.

The microphones themselves are 6 mm omnidirectional electrets (Panasonic WM-61A102B), wired with 1.8 mm-diam shielded, twisted-pair cable (Jaguar Industries JFX-

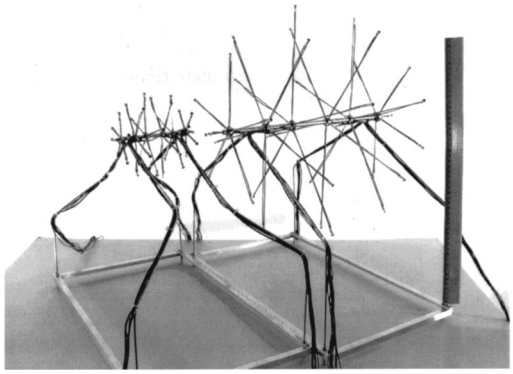


FIG. 5. Photograph of arrays, shown with a 50 cm ruler. The high-frequency array is the smaller one on the left; the low-frequency array is on the right.

SJ232). The microphones were calibrated in an anechoic chamber and matched sets were selected for each array. The mean sensitivity of each microphone was measured and is used for calibration in all measurements. The standard deviation of sensitivity of each set is less than 0.1 dB over each operating range. When mounted, the microphones are not actually touching the stainless steel stands, but are sitting 1 mm away from the ends of the 1.6 mm-diam spokes, held in place by the cables, which are tightly taped to the frame. Both arrays sit at the same height atop rectangular leg frames with the same footprint so as to allow precise positioning and repositioning.

B. The data acquisition system

The control computer is a portable ''lunch-box'' type, equipped with a multichannel sound card (Echo Audio Layla24). This card has eight simultaneous analog inputs, allowing only eight microphones to be measured at a time. Hence, the microphone cables are wired in groups of eight into four connectors, to enable quick patching into the eight-channel preamplifier. The preamplifier provides 29.4 dB of gain, each channel being individually corrected for variations.

The playback side consists of a power amplifier (Bryston 2B) driving an omnidirectional loudspeaker (B&K Omnisource Type 4295). The playback and recording is effected using multichannel recording software (Syntrillium Software's CoolEdit SE).

To prepare for the actual measurement process, several sound files each containing nine replicas of a length-N MLS were generated, where $N=2^L-1$, for MLS orders L ranging from 14 to 18. At a sampling rate of 44.1 kHz, this means the test MLS is between 0.37 and 5.94 s; hence, a wide range of enclosures can be measured.

For the data acquisition, the appropriate file is played and eight microphones are simultaneously recorded. The microphone connector is switched, and the process repeated using the same MLS until all four sets of microphones have been measured. Each of the 32 microphone recordings is processed in the same manner:[26] the first N samples are discarded, the next eight groups of N samples are averaged to reduce noise, resulting in a length-N microphone signal. This microphone signal is cross-correlated with the original MLS and an N-point impulse response results. This set of 32 impulse responses is taken as the input to the beamformer.

It should be noted that, as constructed, the system is capable of measuring non-time-varying aspects of the sound field only, namely the reverberant field. With an additional 3 sound cards, a full real-time 32 channel capability could be realized and time-varying sound fields—such as those due to talkers or machinery noise—could be analyzed.

IV. ANALYSIS

A. Quantifiers relating to variations at a point

Once the 60 directional pressure impulse responses are on hand, several analysis options exist. The energy arriving from each direction during a given time range and the rate of this energy decay can both be computed and investigated.

1. Directional diffusion

The most prevalent quantifier describing how the incident sound varies with direction at a receiving point is the *directional diffusion*, d.[10,27] This quantity is a function of the measured intensity,

$$I_m(t,\theta,\phi)=\int\int I(t,\theta',\phi')|\Gamma(\theta-\theta',\phi-\phi')|^2 \times\cos\theta'\, d\theta' d\phi', \tag{1}$$

which is the convolution in elevation θ and azimuth ϕ of the receiver directivity pattern Γ with the true incident acoustic intensity, $I(t,\theta,\phi)$. The narrower the beam pattern, the more closely Γ approaches a delta function, and the more closely I_m approximates I at each time instant, t. The directional diffusion is defined as

$$d=\left(1-\frac{m}{m_0}\right)\times 100\%, \tag{2}$$

where m is the average of the absolute differences of I_m from the spatial average, normalized by the average

$$m=\frac{1}{4\pi\langle I_m\rangle}\int\int |I_m-\langle I_m\rangle|\cos\theta\, d\theta\, d\phi, \tag{3}$$

where the average is given by

$$\langle I_m\rangle=\frac{1}{4\pi}\int\int I_m\cos\theta\, d\theta\, d\phi, \tag{4}$$

and m_0 is the value of m for a single plane wave incidence. This normalization results in d taking on a value of 100% in a perfectly diffuse sound field, and a value of 0% under anechoic conditions. Since I_m depends on the receiver pattern, so does d, and hence the same sound field measured with different receivers can yield differing values. In this work the same receiver is used in all cases, however, so comparisons between subsequent measurements can be safely made.

It is possible to compute m at each time instant and obtain d as a function of time, but a simple integration of I_m over time will allow d to be determined for a given time range.

In this work, d is computed from the 60 directional pressure responses $p_{d,i}(t)$ $(i=1,...,60)$ as follows. First, the squared pressure responses are integrated from $t_1=n_1T_s$ to $t_2=n_2T_s$ to get the energy E_i arriving from each direction i over that time period,

$$E_i=\frac{AT_s}{\rho c}\sum_{n=n_1}^{n_2} p_{d,i}^2(nT_s), \tag{5}$$

where the effective cross-sectional area of incidence A, the sampling period T_s, the density of air ρ, and the speed of sound c are constants. Then

$$\langle E\rangle=\frac{1}{60}\sum_{i=1}^{60} E_i, \tag{6}$$

and the value of m is computed from

$$m = \frac{1}{\langle E \rangle} \sum_{i=1}^{60} |E_i - \langle E \rangle|, \tag{7}$$

and d follows from Eq. (2).

2. Directivity index

A coarser, but possibly more intuitive, measure is the *directivity index*, DI. The DI, which is commonly used to quantify microphone and antenna patterns, is equally applicable to the measured directional energies, and is given by

$$\mathrm{DI} = 10 \log \left[\frac{\max(E_i)}{\langle E \rangle} \right], \tag{8}$$

which is merely the peak-to-average ratio, expressed in dB. Admittedly, it could yield the same value for differing sound fields in which the directional diffusion would reveal a difference. Nonetheless, it is a quick measure of the variation of incident energy over all angles. Notice, however, that this quantifier too depends on the directivity pattern of the receiver.

3. Directional decay rates

To investigate the energy decay rate in each direction, the reverberation time can be computed from each directional impulse response.[27] The times are compared, rather than the rates themselves, to maintain consistency with convention. For a directional response $p_{d,i}(t)$, the ''directional reverberation time'' is given by

$$T_{60d,i} = 60 \left(\frac{\Delta L_i}{\Delta t} \right)^{-1}, \tag{9}$$

where $\Delta L_i / \Delta t$ is the slope in dB/s of the decay curve given by

$$L_i(nT_s) = 10 \log \sum_{l=N}^{n} p_{d,i}^2(lT_s), \tag{10}$$

where N is the last measurement sample. To determine the variation of the decay rate with direction, the standard deviation of the set of 60 directional reverberation times is computed.

B. Quantifiers relating to point-to-point variations

Knowing the 32 omnidirectional responses at the microphone positions allows the determination of several acoustic indicators at different points in the space, as well as the correlation of the sound field between pairs of those points. These could well be considered local variations, as the spatial separation of these points is on the order of a wavelength at most.

1. Reverberation time

As was done for the directional impulse responses, the reverberation time is computed for each of the 32 omnidirectional microphone responses. This is obtained as above, with $p_{m,j}(t)$, the omnidirectional impulse response for position $j = 1,...,32$ substituted for $p_{d,i}(t)$ in Eq. (10). Then Eq. (9) yields $T_{60,j}$, the traditional reverberation time for microphone position j. The reverberation times can be computed for various frequency bands and compared by computing the standard deviation.

2. Spatial variation of pressure

The time-averaged sound pressure level for microphone position j, $L_{p,j}$, is computed from

$$L_{p,j}(\omega_C) = 10 \log \int_{\omega_L}^{\omega_U} |P_{m,j}(\omega)|^2 \, d\omega, \tag{11}$$

where ω_C is the center frequency of the band defined by upper frequency ω_U and lower frequency ω_L, and $P_{m,j}(\omega)$ is the Fourier transform of $p_{m,j}(t)$. The 32 levels can be computed in various frequency bands, and compared to each other by computing the standard deviation.[9]

3. Point-to-point cross-correlation

Knowing the pressure in the sound field at the 32 microphone positions, the cross-correlation between pairs of microphone signals can be computed. Given the geometry of the array, no more than two microphones are collinear, so variation of the cross-correlation coefficient at a fixed distance, but in different directions, is investigated.

From the Fourier transforms of two microphone signals $p_{m,j_1}(t)$ and $p_{m,j_2}(t)$, the coefficient for the frequency band centered at ω_C is computed from[5]

$$C_{j_1 j_2} = \frac{\int_{\omega_L}^{\omega_U} \mathrm{Re}\{P_{m,j_1}(\omega) P_{m,j_2}^*(\omega)\} d\omega}{[\int_{\omega_L}^{\omega_U} |P_{m,j_1}(\omega)|^2 \, d\omega \int_{\omega_L}^{\omega_U} |P_{m,j_2}(\omega)|^2 \, d\omega]^{1/2}}. \tag{12}$$

In general, the cross-correlation coefficient depends on kr, where $k = \omega/c$ is the wave number and r is the distance between the microphones. In particular, for a truly diffuse field, the expected coefficient is[5]

$$C_{3D} = \frac{\sin(kr)}{kr} \frac{\sin(\Delta k \, r/2)}{\Delta k \, r/2}, \tag{13}$$

where $\Delta k = k_U - k_L$ is the difference in wave number over the frequency band of interest.

In the special case of a single plane-wave arrival (i.e., anechoic), the expected expression is

$$C_{1D} = \frac{\sin(k_U r \cos\alpha) - \sin(k_L r \cos\alpha)}{\Delta k \, r \cos\alpha}, \tag{14}$$

where α is the angle the plane-wave arrival makes with the line joining the microphones.

The coefficients can be computed over various frequency bands and for various microphone pairs. For a fixed distance, different correlations in different directions imply anisotropy of the field.

C. Graphical output

The computation of numerical quantifiers is certainly of importance but, to get a feel for from *where* the sound arrives, it is important to be able to see the distribution. Plots of three-dimensional surfaces have been generated to aid in

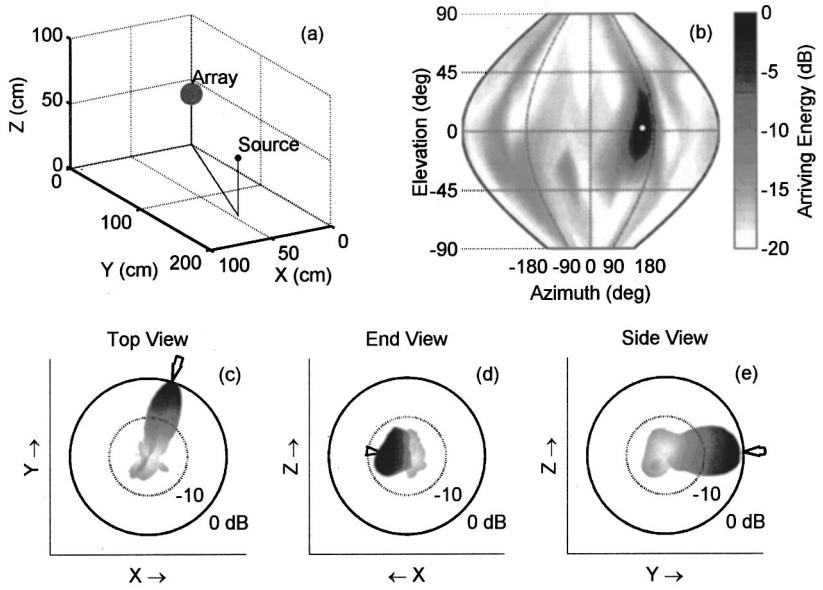


FIG. 6. Anechoic chamber measurement with high-frequency array (1000–3300 Hz): (a) Geometry of the measurement setup. (b) The arriving sound energy as a gray scale map, plotted against direction. The angles are measured with respect to the array position shown in (a): azimuth in a plane parallel to the *XY* plane, increasing from the *X* axis, elevation out of this plane, increasing toward the $+Z$ axis. The expected directions of sound arrival are indicated by small white circles. (c)–(e) Three different views of a surface whose radius and gray level in a particular direction are proportional to the arriving energy in dB. The gray of the surface is consistent with the scale in (b). Directions of expected sound arrival are indicated by arrows pointing inward. The directional diffusion $d=21\%$, the directivity index DI$=12.8$ dB, and the directional reverberation time $T_{60,d}=28\pm8$ ms.

the visualization. The radius and color of the surface in a given direction are proportional to the arriving energy. The measured energies are interpolated among the 60 steering directions to yield a complete surface. In a spherically isotropic field, this surface would be a uniformly colored sphere. In a field with only one sound wave arrival, it would be a single lobe, corresponding to the beam pattern.

To show all directions of arrival at the same time, the three-dimensional surface can also be shown ''flattened'' onto a plane, plotted in much the same style as a cartographic map. While some perspective is lost in this projection, it simplifies the identification of key directions of sound arrival.

V. VERIFICATION OF SYSTEM PERFORMANCE: ANECHOIC MEASUREMENT

The measurement system was first tested in an anechoic chamber, where only a single sound wave arrival (i.e., the direct wave) is expected. The plots of the total arriving energy, summed over the entire decay time of the chamber, are shown in Fig. 6 for the high-frequency array measurement, and in Fig. 7 for the low-frequency array measurement.

In these and all similar subsequent figures, the geometry of the measurement setup is shown in panel (a). The dimensions are to scale, and the array is shown as a sphere of the appropriate diameter (16 cm for the high-frequency array, 48

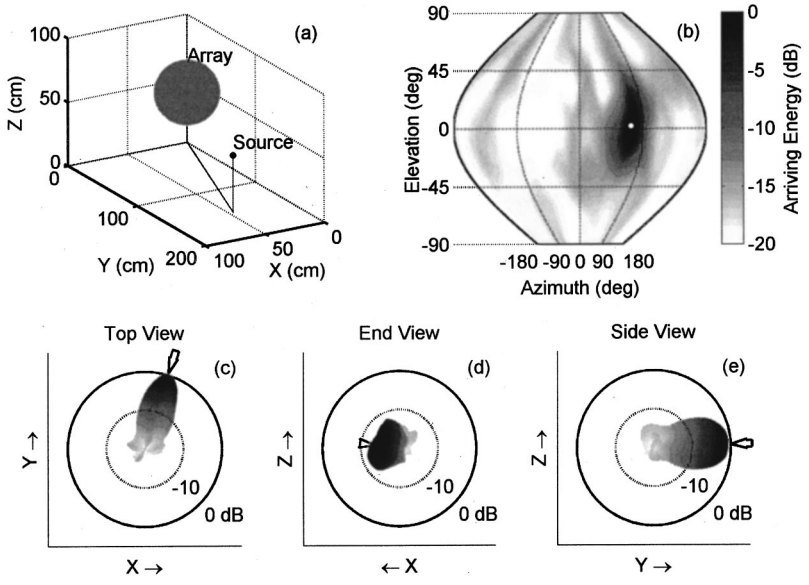


FIG. 7. Anechoic chamber measurement with low-frequency array (300–1000 Hz), the same as Fig. 6. $d=17\%$, DI$=12.8$ dB, $T_{60,d}=29\pm8$ ms.

cm for the low-frequency array). Panel (b) shows the measured energy as a gray scale map, plotted against direction. The elevation and azimuth angles are measured from the array position, in the usual manner with regard to the axes in panel (a): azimuth in a plane parallel to the *XY* plane, increasing from the *X* axis, elevation out of this plane, increasing toward the $+Z$ axis. Lines of constant elevation are straight and horizontal; lines of constant azimuth are curved to yield a somewhat more ''equal-area'' projection. The gray scale on the right is in dB, 100% gray (i.e., black) corresponds to the peak value for this measurement (0 dB), 0% gray (i.e., white) corresponds to -20 dB relative to this. Expected directions of sound arrival, corresponding to the direct sound and specular reflections (as appropriate), are indicated by small white circles. Panels (c)–(e) show three different views of a surface whose radius and gray scale in a given direction are proportional to the arriving energy. The gray of the surface is consistent with the scale in panel (b). Directions of expected sound arrival are indicated by arrows pointing inward. The values for the directional diffusion, the directivity index, and the mean and standard deviation of directional reverberation time are given in the captions.

For both arrays, the direction of peak measured energy lines up with the expected direction of sound arrival (72° azimuth, 2° elevation). That is, the single sound wave is correctly located. The measured values for *d* (21% for the high-frequency measurement, 17% for the low-frequency measurement) and for DI (12.8 dB both measurements) define the anechoic baseline performance of the system. Deviations from the theoretical limits of 0% and 14 dB, respectively, are ascribed to noise in the measurement system, scattering from decking and other materials, and a faint reflection from the door frame in the chamber (approximately 12 dB down, $-120°$ azimuth, 6° elevation).

For a single-wave incidence in the anechoic chamber, the decay rate in each direction is expected to be the same, however, it should be a stepwise decay. For such a field, finding the slope of the decay curve described by Eq. (10) is difficult, which is reflected in the large standard deviation of decay times (28% both measurements). For this reason, reverberation time measurements for this sound field are not very meaningful.

Figures 8 and 9 show the results for the analysis of the point-to-point variations of reverberation time and of sound pressure level for the high-frequency and low-frequency array measurements, respectively. For the set of third-octave reverberation times, panel (a) of each figure shows the mean values plotted versus the frequency band centers, the standard deviations indicated with the error bars. Panel (b) in each figure depicts analogous plots for the third-octave sound pressure level.

As with the directional decays, the measured standard deviations of the omnidirectional reverberation times are larger than is expected, due to the difficulty in obtaining the slope of the decay curves computed from Eq. (10) for such a nonreverberant field. As measured with the microphones in both arrays, the mean values of the sound-pressure level are about the same, but the standard deviations are somewhat larger for the low-frequency array measurement, where the

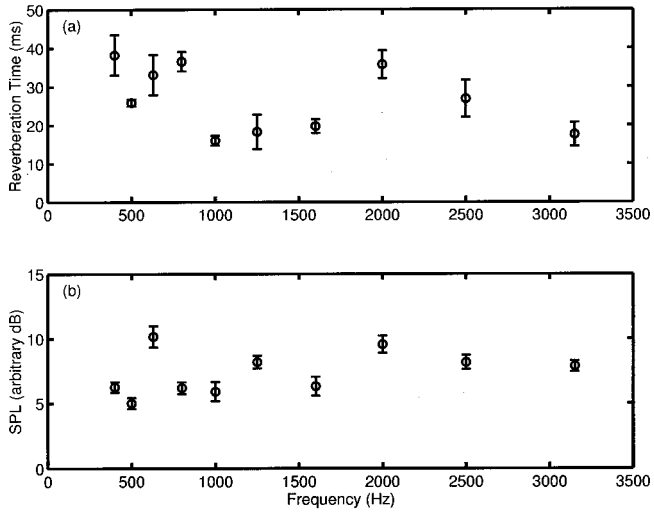


FIG. 8. An anechoic chamber measurement with the high-frequency (16 cm) array: (a) Mean and standard deviation of $\frac{1}{3}$-octave reverberation time, computed from the 32 omnidirectional microphone signals; (b) mean and standard deviation of $\frac{1}{3}$-octave sound pressure level, computed from the 32 omnidirectional microphone signals. The microphone sensitivities are matched to within 0.1 dB over the operating range of the array (1000–3300 Hz), and within 0.2 dB from 300–1000 Hz. All the error bars are a total of two standard deviations long.

microphones are spaced farther apart. Even so, these deviations are only on the order of 1 dB, which is small enough to imply homogeneity of the sound field.[9]

Figures 10 and 11 show the third-octave cross-correlation coefficients of microphone #1 with its nearest neighbors, microphones #2–6. The distance of each from microphone #1 is the same; only the directions vary. The data points in each figure are measured values, the solid curves are obtained from Eq. (14) for the appropriate microphone spacing and sound wave angle of incidence. The data points follow different trends for different directions, which means the sound field is not isotropic. This is as must be

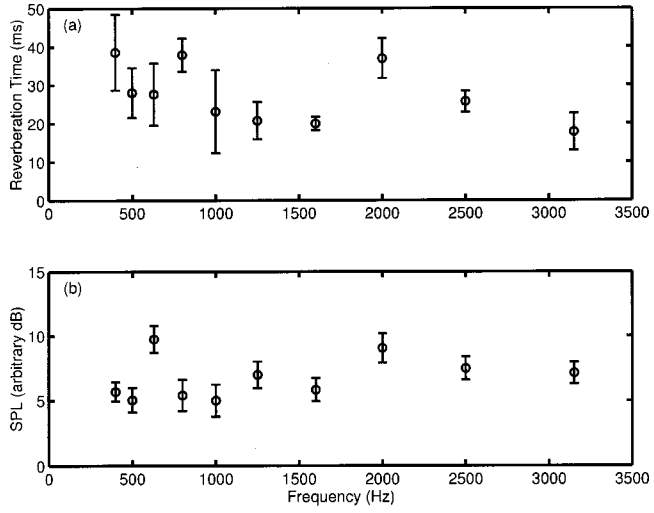


FIG. 9. An anechoic chamber measurement with the low-frequency (48 cm) array, the same as Fig. 8. (a) $\frac{1}{3}$-octave reverberation time, (b) $\frac{1}{3}$-octave sound pressure level. The microphone sensitivities are matched to within 0.1 dB over the operating range of the array (300–1000 Hz), and within 0.2 dB from 1000–3300 Hz.

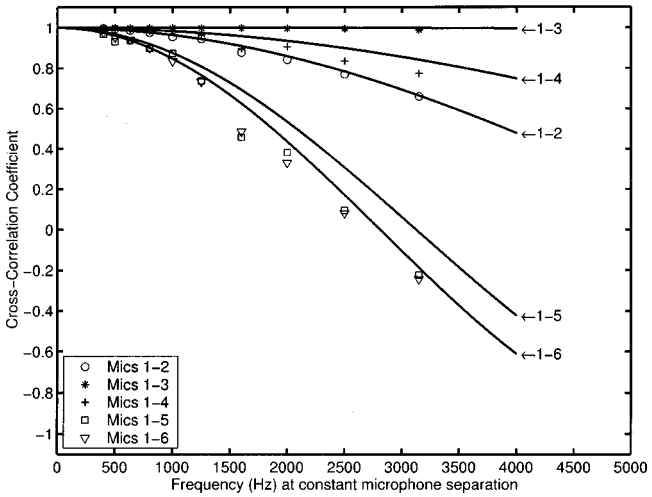


FIG. 10. Third-octave cross-correlation coefficient of high-frequency array microphone #1 with nearest neighbors, measured in an anechoic chamber. The solid curves are the expected theoretical relations, determined from Eq. (14) for a constant separation of $r = 5.1$ cm.

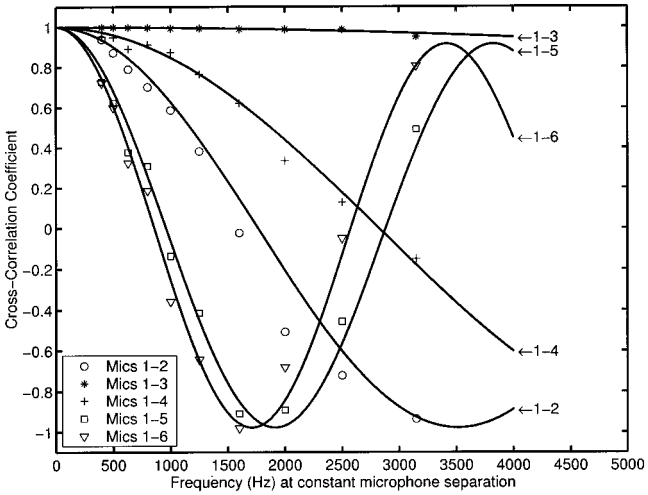


FIG. 11. Third-octave cross-correlation coefficient of low-frequency array microphone #1 with nearest neighbors, measured in an anechoic chamber. The solid curves are the expected theoretical relations, determined from Eq. (14) for a constant separation of $r = 15.4$ cm.

expected for a single point source. There is, in fact, very good agreement with the theoretical curves.

VI. DEMONSTRATION OF CAPABILITIES IN SIMPLE SOUND FIELDS

To demonstrate the directional discrimination capabilities of the system, several experiments were performed in progressively more complex sound fields in an anechoic chamber. Initially, a single reflecting panel was added to the anechoic setup to generate a single reflection. Next, a second panel was added to generate additional reflections, arriving at different times from widely separated directions. Finally, two reflectors were oriented parallel to each other, with the source and array both on the midline between them. This generates two symmetric reflections that arrive simultaneously at the array position.

In each of these cases, the received energy over the entire decay time of the sound field is plotted and the quantifiers relating to the variation of the field at the receiving point are computed. Since there is not much difference between the plots for the high-frequency array and for the low-frequency array, in the interest of conciseness, only the low-frequency results shall be discussed. Additionally, the quantifiers relating to the sound field decay and to the point-to-point variations turn out to be not very interesting in these artificially simple sound fields, and are also not discussed.

A. Single reflection

With the array and loudspeaker set up as in the anechoic chamber measurements of Sec. V, a single planar reflector was added to generate a second sound wave arrival. The reflector measured 183×76 cm, was located 100 cm away from both source and receiver, and had a surface of varnished wood. The results for the arriving energy in this case are shown in Fig. 12 for the low-frequency array measurement, the layout of the figure the same as Fig. 6.

As can be seen by the location of the dark gray and elongated lobes on the solid, both the direct sound (72° azi-

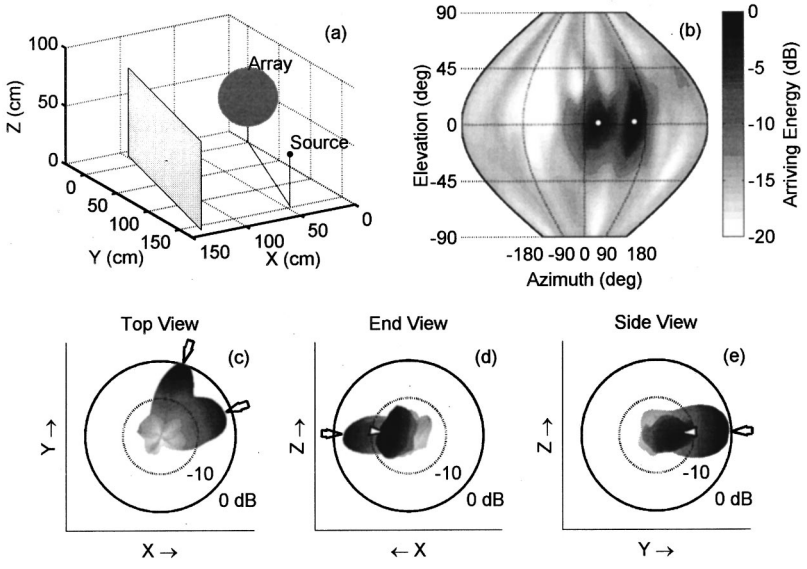


FIG. 12. Single reflector in an anechoic chamber; a measurement with the low-frequency array (300–1000 Hz), the same as Fig. 6. $d = 24\%$, DI = 10.6 dB, $T_{60,d} = 38 \pm 13$ ms.

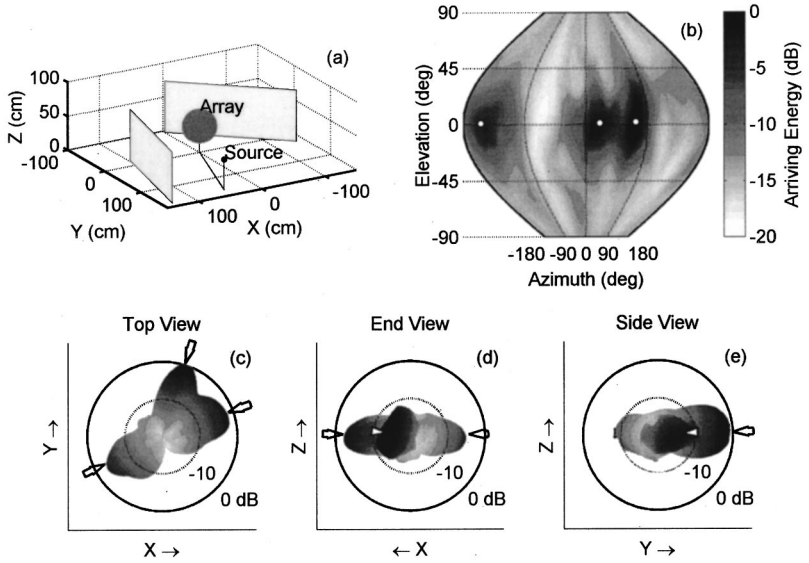


FIG. 13. Two skew reflectors in an anechoic chamber; a measurement with the low-frequency array (300–1000 Hz); the same as Fig. 6. $d=33\%$, DI$=9.3$ dB, $T_{60,d}=34\pm12$ ms.

muth, 2° elevation) and the reflection (20° azimuth, 2° elevation) are correctly located.

The value of d (24%) is higher than for the anechoic case (17%), and the value of DI (10.6 dB) is lower than for the anechoic case (12.8 dB), both of which are as expected.

B. Separate reflections

An additional planar reflector was added to the setup above to generate a second reflection, arriving at a later time than the first. A third reflection is also generated, reflecting off both panels, and arriving from the same direction as this second (first-order) reflection. The energy-arrival results for the low-frequency array measurement are shown in Fig. 13, with the same layout as Fig. 6.

Notice that the direct sound (72° azimuth, 2° elevation), first reflection (20° azimuth, 2° elevation), and the second and third reflections (−154° azimuth, 2° elevation) are all resolved and located correctly.

The values of the directional diffusion (33%) and of the directivity index (9.3 dB) are incrementally farther away from their anechoic values (17%, 12.8 dB): d being increased, and DI being lowered by the additional reflection.

C. Simultaneously arriving reflections

Two sound waves arriving simultaneously from different directions can lead to ambiguities with some measurement procedures, in particular those computing net intensity.[16] To investigate how the present system handles this situation, both reflectors above were oriented parallel to the line joining the source to the array, equally spaced on either side. This geometry results in two reflections arriving simultaneously at the array position.

As can be seen from Fig. 14, both reflections are clearly resolved. The directions of the direct sound (72° azimuth, 2° elevation) and of both reflections (20° azimuth, 2° elevation) and (124° azimuth, 2° elevation) are quite well discerned. Knowledge of the instantaneous intensity in this case would only reveal the net energy transport in the direction of the direct sound: the reflections would not be resolvable.

The directional diffusion in this case (32%) is not differ-

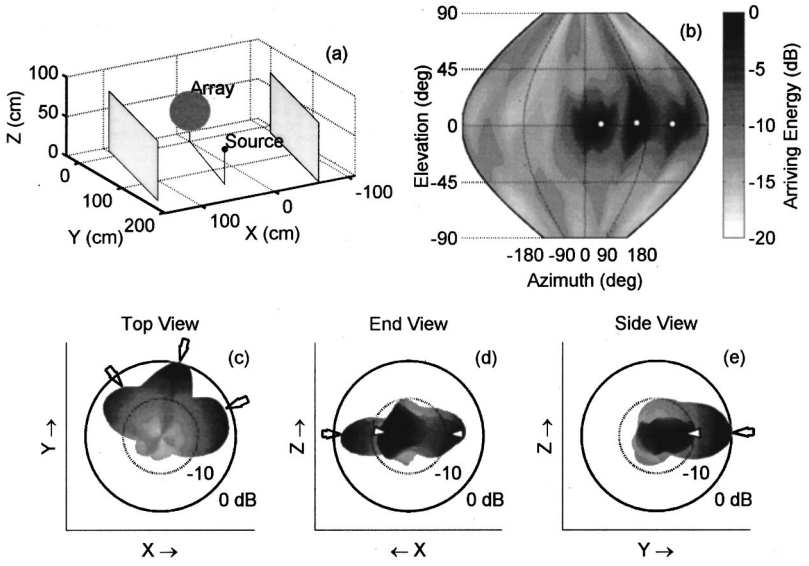


FIG. 14. Two parallel reflectors in an anechoic chamber; a measurement with the low-frequency array (300–1000 Hz), the same as Fig. 6. $d=32\%$, DI$=8.6$ dB, $T_{60,d}=70\pm20$ ms.

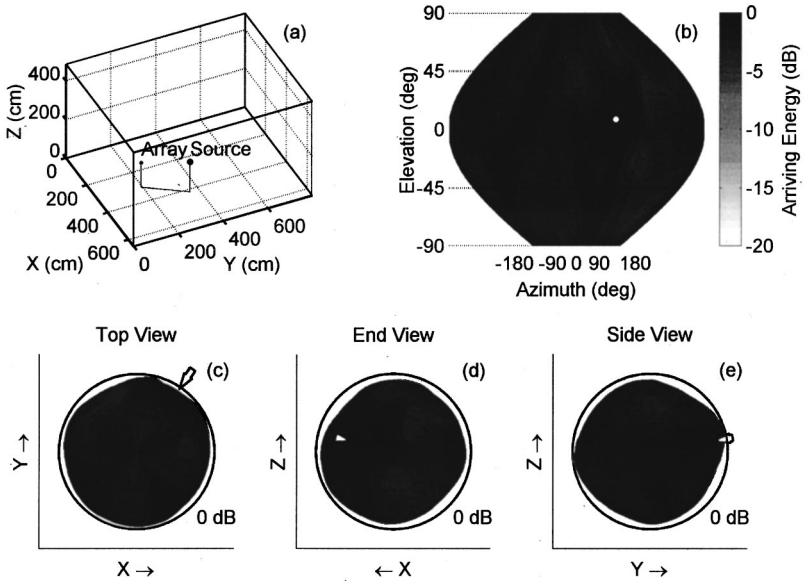


FIG. 15. A reverberation chamber measurement with the high-frequency array (1000–3300 Hz), the same as Fig. 6. Diffuser panels in the chamber are not shown. $d=91\%$, DI$=1.5$ dB, $T_{60,d}=4.85\pm0.48$ s.

ent than for the preceding case with the separate reflections, however, the directivity index (8.6 dB) is lower. This latter result appears consistent with the fact that the second reflection is closer (in angle) to the direct sound than in the previous case. The intermediate directions each have more arriving energy, due to the finite beamwidth of the array.

VII. REVERBERATION CHAMBER MEASUREMENT

Having previously investigated the behavior of the measuring system in the case of single-wave incidence, a characterization of its performance in the other extreme case, namely as diffuse a field as possible, was subsequently undertaken. To this end, the system was tested in a reverberation chamber.

The chamber used was rectangular ($7.96\times6.48\times4.88$ m), with 13 diffuser panels, none of which was moving. The source and array were positioned approximately 2 m apart, away from any line of symmetry. The geometry and the results for the total energy arrival with direction are shown in Fig. 15 for the high-frequency array measurement, and in Fig. 16 for the low-frequency array measurement. The layout of these figures is the same as for Fig. 6. The total energy is summed over the entire 5.92 s measured impulse response, to be sure to include the entire decay of the sound field. As is immediately obvious from the gray map, the sound field is very nearly isotropic, as expected. The reverberant field is so strong, in fact, that the direction of direct sound incidence is not even the peak energy direction. This is not unexpected since the reverberation distance[3] for this room is about 45 cm, considerably less than the distance between source and array. The directivity index values of 1.5 and 1.7 dB, the directional diffusion values of 91%, and the low percentage standard deviations for the directional decay rates (10% and 4%) do indeed confirm the isotropy of the sound field in this reverberation chamber.

The homogeneity of the sound field over the extent of the arrays is confirmed by the small standard deviations of reverberation time and of sound pressure level, shown, as in

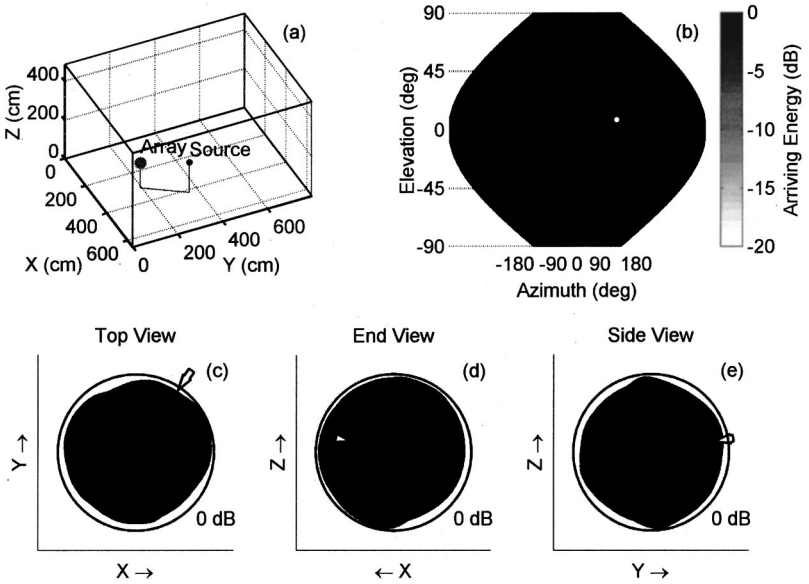


FIG. 16. Reverberation chamber measurement with the low-frequency array (300–1000 Hz), the same as Fig. 6. Diffuser panels in the chamber are not shown. $d=91\%$, DI$=1.7$ dB, $T_{60,d}=4.29\pm0.16$ s.

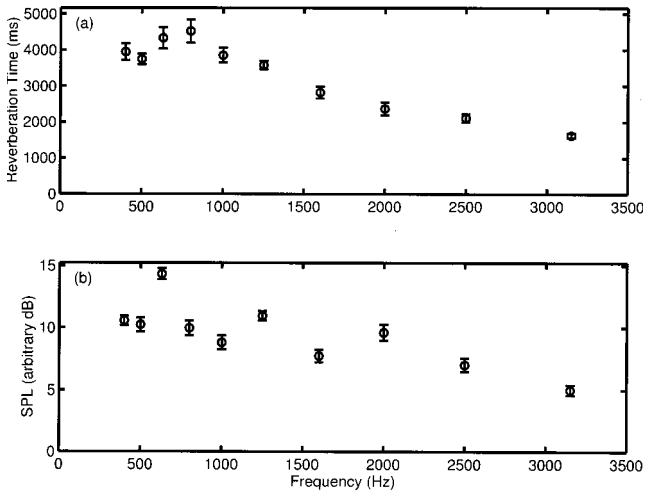


FIG. 17. Reverberation chamber measurement with the high-frequency (16 cm) array; the same as Fig. 8. (a) $\frac{1}{3}$-octave reverberation time; (b) $\frac{1}{3}$-octave sound pressure level. The microphone sensitivities are matched to within 0.1 dB over the operating range of the array (1000–3300 Hz), and within 0.2 dB from 300–1000 Hz.

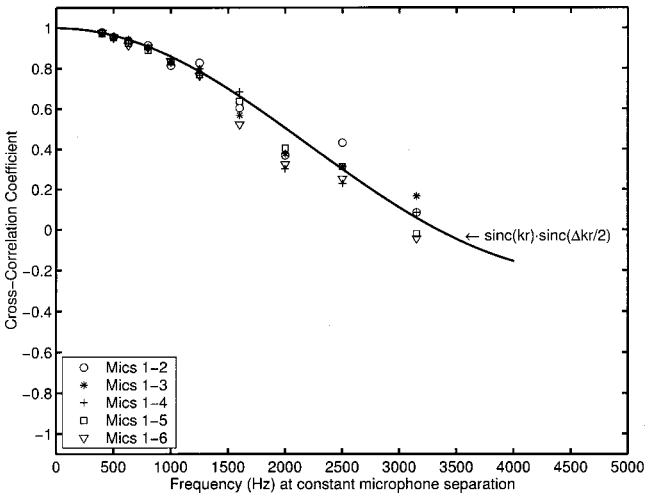


FIG. 19. Third-octave cross-correlation coefficient of high-frequency array microphone #1 with nearest neighbors, measured in reverberation chamber. The solid curve is the theoretical relation for a diffuse field, determined from Eq. (13) for a constant separation of $r=5.1$ cm.

Fig. 8, in Fig. 17 for the high-frequency array and in Fig. 18 for the low-frequency array. The variations for the low-frequency array are somewhat larger than for the high-frequency array, again as expected, due to the larger separation between microphone positions.

The point-to-point cross-correlation coefficients of microphone #1 with its nearest neighbors are shown in Fig. 19 and Fig. 20 for the high- and low-frequency arrays, respectively. Observe in both cases that there is very little deviation of the coefficient with direction, again implying isotropy, and that both sets of measurements follow the theoretical curve for a spherically diffuse sound field, as computed from Eq. (13).

VIII. SUMMARY

The microphone array measurement system described herein is a promising acoustical analysis tool. The isotropy of a sound field can be analyzed over a range of frequency from 300–3000 Hz by measuring both the arriving energy, and its rate of decay, with direction. The system is also capable of measuring the local point-to-point variation of reverberation time and sound pressure level, as well as the dependence of the spatial cross-correlation coefficient on direction, all of which are indicators of diffuseness of the sound field.

Furthermore, by restricting the time range over which the arriving energy is counted, the distribution of the sound field can be investigated on a temporal basis. It may be of interest, for instance, to examine the so-called ''early'' and ''late'' times of a room response separately, or to study the evolution of the sound decay by computing the arriving energy at each time sample.

The system is also capable of being expanded to handle nonreverberant fields, such as noise or speech, with the addition of 24 more channels of data acquisition hardware. This

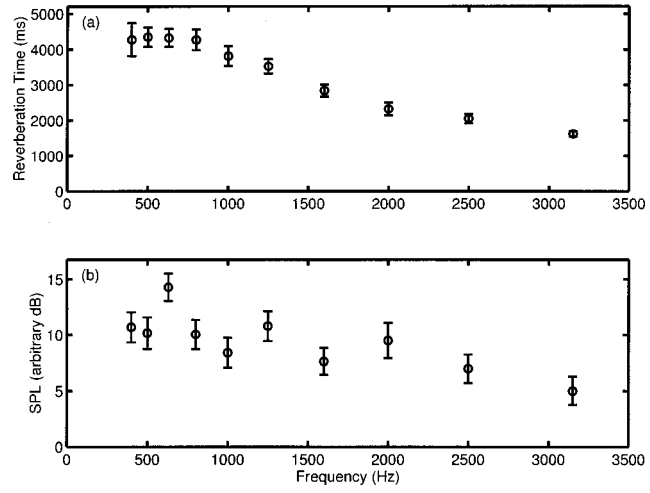


FIG. 18. Reverberation chamber measurement with the low-frequency (48 cm) array; the same as Fig. 8. (a) $\frac{1}{3}$-octave reverberation time; (b) $\frac{1}{3}$-octave sound pressure level. The microphone sensitivities are matched to within 0.1 dB over the operating range of the array (300–1000 Hz), and within 0.2 dB from 1000–3300 Hz.

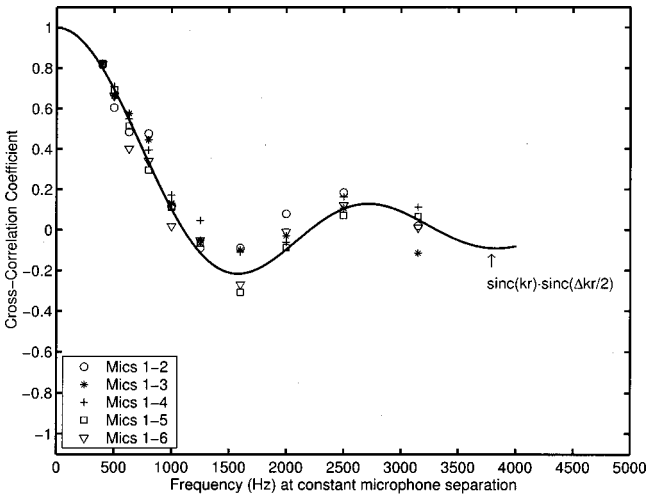


FIG. 20. Third-octave cross-correlation coefficient of low-frequency array microphone #1 with nearest neighbors, measured in reverberation chamber. The solid curve is the theoretical relation for a diffuse field, determined from Eq. (13) for a constant separation of $r=15.4$ cm.

would presumably enable a detailed analysis of sound fields that arise in such applications as hands-free teleconferencing, potentially leading to an improved speech-pickup system. Work in these areas will continue.

ACKNOWLEDGMENTS

Thanks to René St. Denis and John Quaroni at NRC for invaluable technical help, and to Harry Dickson and Allan Horner in the NRC M-36 Machine Shop for building the spherical microphone stands. Thanks also to Dr. John Bradley of the NRC's Institute for Research in Construction for arranging use of their reverberation chamber. One author (BG) would like to thank Dr. John Vanderkooy at the University of Waterloo for financial support and for many helpful discussions.

[1] "Acoustical terminology," ANSI Standard, S1.1-1994, 1994.
[2] T. J. Schultz, "Diffusion in reverberation rooms," J. Sound Vib. **16**(1), 17–28 (1971).
[3] L. Cremer and H. A. Müller, *Principles and Applications of Room Acoustics*, translated by T. J. Schultz (Applied Science Publishers, Essex, UK, 1982), Vol. I.
[4] A. Abdou and R. W. Guy, "A review of objective descriptors for sound diffuseness," Can. Acoust. **22**, 43–44 (1994).
[5] H. Nélisse and J. Nicolas, "Characterization of a diffuse field in a reverberant room," J. Acoust. Soc. Am. **101**, 3517–3524 (1997).
[6] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson, "Measurement of correlation coefficients in reverberant sound fields," J. Acoust. Soc. Am. **27**, 1072–1077 (1955).
[7] C. T. Morrow, "Point-to-point correlation of sound pressures in reverberation chambers," J. Sound Vib. **16**, 29–42 (1971).
[8] G. C. J. Bart, "Spatial crosscorrelation in anisotropic sound fields," Acustica **28**, 45–49 (1973).
[9] "Acoustics—Determination of sound power levels of noise sources—Precision methods for discrete-frequency and narrow-band sources in reverberation rooms," ISO Standard 3742, 1988.
[10] R. Thiele, "Richtungsverteilung und Zeitfolge der Schallrückwurfe in Räumen," Acustica **3**, 291–302 (1953).
[11] E. Meyer and R. Thiele, "Raumakustische Untersuchungen in zahlreichen Konzertsälen und Rundfunkstudios unter Anwendung neuerer Messverfahren," Acustica **6**, 425–444 (1956).
[12] M. R. Schroeder, "Measurement of sound diffusion in reverberation chambers," J. Acoust. Soc. Am. **31**, 1407–1414 (1959).
[13] A. D. Broadhurst, "An acoustic telescope for architectural acoustic measurements," Acustica **46**, 299–310 (1980).
[14] K. Sekiguchi, S. Kimura, and T. Hanyuu, "Analysis of sound field on spatial information using a four-channel microphone system based on regular tetrahedron peak point method," Appl. Acoust. **37**, 305–323 (1992).
[15] Y. Yamasaki and T. Itow, "Measurement of spatial information in sound fields by closely located four point microphone method," J. Acoust. Soc. Jpn. (E) **10**, 101–110 (1989).
[16] R. W. Guy and A. Abdou, "A measurement system and method to investigate the directional characteristics of sound fields in enclosures," Noise Control Eng. J. **42**, 8–18 (1994).
[17] B. N. Gover, "Measurement of directional information in sound fields," J. Acoust. Soc. Am. **106**, 2218(A) (1999).
[18] B. N. Gover, "Development and use of a spherical microphone array for measurement of spatial properties of reverberant sound fields," Ph.D. thesis, Department of Physics, University of Waterloo, Waterloo, Canada, 2001.
[19] F. I. Tseng and D. K. Cheng, "Gain optimization for arbitrary antenna arrays subject to random fluctuations," IEEE Trans. Antennas Propag. **15**, 356–366 (1967).
[20] D. K. Cheng and F. I. Tseng, "Optimum spatial processing in a noisy environment for arbitrary antenna arrays subject to random errors," IEEE Trans. Antennas Propag. **16**, 164–171 (1968).
[21] E. N. Gilbert and S. P. Morgan, "Optimum design of directive antenna array subject to random variations," Bell Syst. Tech. J. **34**, 637–663 (1955).
[22] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," IEEE Trans. Acoust., Speech, Signal Process. **35**, 1365–1376 (1987).
[23] T. W. Parks and C. S. Burrus, *Digital Filter Design* (Wiley, New York, 1987).
[24] See EPAPS Document No. E-JASMAN-112-015211 for the beamforming coefficients used in this work. 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.
[25] G. W. Elko, "Superdirectional microphone arrays," Chap. 10 of *Acoustic Signal Processing for Telecommunication*, edited by S. L. Gay and J. Benesty (Kluwer Academic, Norwell, MA, 2000).
[26] D. D. Rife and J. Vanderkooy, "Transfer-function measurement with maximum-length sequences," J. Audio Eng. Soc. **37**, 419–444 (1989).
[27] H. Kuttruff, *Room Acoustics*, 3rd ed. (Elsevier Applied Science, London, 1991).

Sweet spot size of virtual acoustic imaging systems at asymmetric listener locations[a)]

John Rose, Philip Nelson, Boaz Rafaely, and Takashi Takeuchi[b)]
Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

(Received 11 December 2001; revised 29 July 2002; accepted 3 August 2002)

Virtual acoustic imaging systems are effective when the listener's head location is close to the head location assumed when the system was designed. The ''sweet spot'' refers to the spatial bubble of head location in which the system is still effective. Some of the previous work investigating the ''stereo dipole'' acoustic imaging system shows that for the traditional on-axis listener location the ''sweet spot'' is about ± 5 cm for lateral head translations. Larger head movements than this require an update of the virtual acoustic imaging filters. The interest here is the ''sweet spot'' size at off-axis asymmetric listener locations or an understanding of how often one needs to update the filters to ensure the listener perceives a stable virtual image as they move. The examination of the off-axis ''sweet spot'' size comprises a theoretical acoustic analysis, computer simulations, and a subjective study. The simulations and subjective evaluation both demonstrate that the width of tolerable lateral head translations is comparable for the symmetric on-axis listener location and asymmetric listener locations that are as far as 25 cm off-axis.
[DOI: 10.1121/1.1510532]

PACS numbers: 43.38.Md, 43.66.Qp, 43.60.Pt, 43.66.Pn [SLE]

I. INTRODUCTION

The principal idea of binaural technology[1] is to present listeners with sound signals at their two ears that coincide with the signals of a real auditory environment. This gives listeners the perception of virtual acoustic images existing where there are no real sound sources. Virtual acoustic imaging systems can use headphones to present the binaural signals[1,2] or loudspeakers.[3–5] One problem with using loudspeakers is cross-talk. This occurs when the right ear hears the signal meant for the left ear and vice versa. Proper filtering of the signals' input into the loudspeakers overcomes this difficulty.[3–6] The design of these filters incorporates knowledge of the impulse responses from the loudspeakers to the listener's ears. These impulse responses depend on the listener's head location and so filter designs typically assume a listener location. Without physically restraining the listener's head to a single location, the listener's actual position will often stray from the optimal listening location. Virtual acoustic imaging systems tolerate some head displacement before the listener begins to lose the desired perception of the virtual acoustic images.[7] The ''sweet spot'' or equalization zone is the name given to the spatial bubble of head location, in which the virtual acoustic imaging system is effective. The geometrical arrangement of the system affects the ''sweet spot'' size.[7] This paper concentrates on one particular virtual acoustic imaging system known as the ''stereo dipole,'' which has been shown to be particularly robust to head movements.[5–10] The ''stereo dipole'' utilizes an arrangement of two loudspeakers subtending an angle of 10° with the listener's head,[10] as shown in Fig. 1. If one is given a way to track the listener's movements, such as by employing a video camera and an image processing head tracking algorithm,[11] the knowledge of the varying head position can be used to continuously select appropriate virtual acoustic imaging filters so that the listener location is always close to the optimum location. An adaptive system needs only to update the virtual acoustic imaging filters before the desired subjective impression begins to degrade. Therefore, it is important to know the ''sweet spot'' size at off-axis locations before implementing such a system.

Takeuchi and Nelson[7] examine the ''sweet spot'' size for the traditional symmetric on-axis listener location. The attempt in this paper is to understand where the boundary of the ''sweet spot'' lies for both the traditional symmetric on-axis listener location and asymmetric off-axis listener locations. The study is restricted to lateral head translations. Figure 1 illustrates this type of head movement. The results suggest that asymmetric listener locations, up to 25 cm off-axis, are as robust to head displacements as the traditional on- axis listener location.

The paper's presentation is as follows. Section II provides a brief overview of the virtual acoustic imaging filter design procedure. Section III first reviews three methods used to model the path responses. New work in Sec. III employs these models to gain insight into the physical nature of the sound field and calculates ''sweet spot'' boundaries using a cross-talk cancellation performance criterion and a strict just noticeable interaural time difference criterion. Section IV presents a subjective experiment that substantiates these boundaries. Section V is the conclusion.

[a)]Portions of this work were presented in ''Variance of sweet spot size with head location for virtual audio,'' at the 110th Audio Engineering Society Convention, 12–15 May 2001, Amsterdam, The Netherlands, AES preprint 5391.

[b)]Now at the Kaima Technical Research Institute, 2-19-1 Tobitakyu, Chofu-shi, Tokyo 182-0036, Japan.

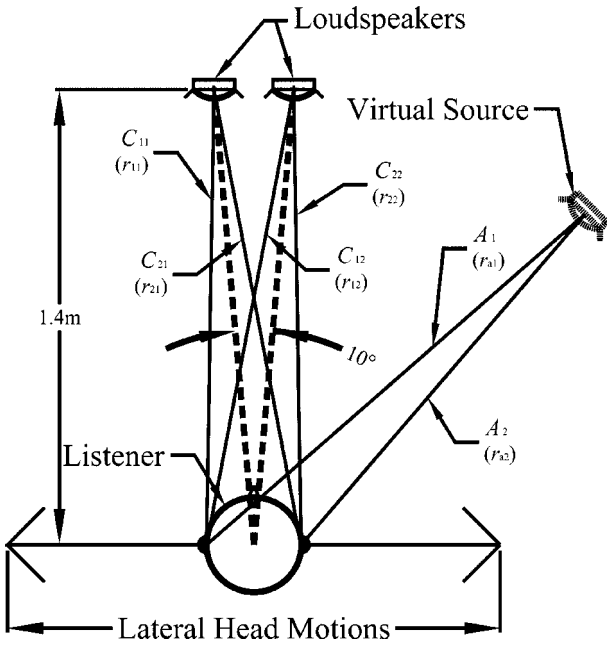


FIG. 1. Plan view depicting lateral head translations and the ''stereo dipole'' geometrical arrangement (Ref. 5). The elements of the electroacoustic path plant matrix **C** and the virtual paths **a** are shown as well as the corresponding distances used in Eqs. (6), (7).

II. FILTER DESIGN

This section presents an overview of well-established material.[4,7–12] Figure 2 shows a block diagram of the implementation of the virtual acoustic imaging system in the z-domain.[7] The virtual paths are the vector of filters $\mathbf{a}(z)$ that represent the transfer functions from a virtual source to the listener's ears. The plant matrix $\mathbf{C}(z)$ represents the transfer functions of the paths from the two real sources to the two ears:

$$\mathbf{a}(z)=\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \tag{1a}$$

$$\mathbf{C}(z)=\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}. \tag{1b}$$

Figure 1 provides an illustration of the elements in Eq. (1). The goal of the system is to produce at the ears of the listener a vector of desired binaural signals $\mathbf{d}(z)$. One acquires the desired signals $\mathbf{d}(z)$ by filtering the source signal $S(z)$ with $\mathbf{a}(z)$. The method of presenting the listener with the sound through two loudspeakers ensures a modification of $\mathbf{d}(z)$ by

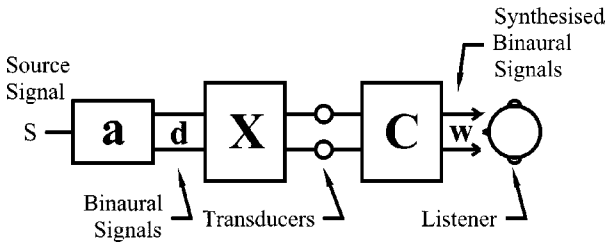


FIG. 2. Block diagram of the implemented virtual acoustic imaging system (Ref. 7).

the plant matrix $\mathbf{C}(z)$. The problem is to design a matrix of filters $\mathbf{X}(z)$ that will cancel the effect of $\mathbf{C}(z)$ i.e.,

$$\mathbf{X}(z)=\mathbf{C}(z)^{-1}. \tag{2}$$

This criterion is relaxed by allowing for some error and by employing a modeling delay Δ to ensure a realizable causal stable inverse so that Eq. (2) becomes

$$\mathbf{C}(z)\mathbf{X}(z)\approx z^{-\Delta}\mathbf{I}, \tag{3}$$

where $\mathbf{I}$ is the identity matrix. The filter matrix $\mathbf{X}(z)$ is the cross-talk cancellation filter matrix, and its job is to cancel the effect of the plant matrix $\mathbf{C}(z)$. Section III employs a control performance matrix $\mathbf{R}$ to help evaluate cross-talk cancellation performance:[7]

$$\mathbf{R}=\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}=\mathbf{CX}. \tag{4}$$

In the frequency domain, the vector of virtual acoustic imaging filters $\mathbf{h}(z)$ is given by the matrix multiplication of the cross-talk cancellation filters $\mathbf{X}(z)$ and the virtual path vector $\mathbf{a}(z)$, i.e.,

$$\mathbf{h}(z)=\mathbf{X}(z)\mathbf{a}(z). \tag{5}$$

III. COMPUTER SIMULATIONS

This section considers the basic physics of the problem to gain some understanding into how designing virtual acoustic imaging filters at off-axis locations affect the ''sweet spot'' throughout the audible frequency range. Subsection A reviews three head related transfer function (HRTF) models that the computer simulations employ. Subsection B considers the concept of the ringing frequency and its relation to geometry and robustness of the virtual acoustic imaging system. Subsection C considers the condition number of plant **C** as a function of head position and its relation to robustness. Subsection D considers the performance of cross-talk cancellation at different head positions and uses a performance criterion to derive a ''sweet spot'' size. Finally, subsection E uses an interaural time difference criterion to derive a ''sweet spot'' size.

A. HRTF approximations

It is possible to measure directly the path responses, shown in Fig. 1, on each individual listener.[13] This approach is impractical if there are many different or unknown people that will use the virtual acoustic imaging system. This subsection reviews three alternative approaches to modeling HRTFs. As the models progress in realism, their complexities also increase. Comparing results obtained using the different models makes it possible to attribute certain aspects of the sound field to specific model characteristics. The free field model provides insight into geometrical effects. In addition to geometrical effects, the spherical head model well approximates the shadowing effects of the listener's head. The dummy head model includes these effects and the effects of the listener's pinnae, neck, and torso.

1. Free field approximation

The simplest approximation is to remove the physical head and replace it with two monopole receivers at the position of the ears. In addition, point monopole sources model the loudspeakers and the environment for this entire arrangement is anechoic. This conceptual situation has the advantage of having a simple analytical solution. The acoustic complex pressure p due to a point monopole source at a distance r from the source is[14]

$$p(r)=\frac{j\omega\rho_0 q e^{-jkr}}{4\pi r}, \tag{6}$$

where an $e^{j\omega t}$ time dependence is assumed and where k, ρ_0, c_0, ω, and q are the wave number ω/c_0, density of the medium, sound speed, angular frequency, and effective complex source strength, respectively. By designating complex acoustic pressure p as the output and complex source volume acceleration $j\omega q$ as the input the transfer functions of Eqs. (1a), (1b) become

$$\mathbf{a}(j\omega)=\frac{\rho_0}{4\pi}\begin{bmatrix}\dfrac{e^{-jkr_{a1}}}{r_{a1}}\\ \dfrac{e^{-jkr_{a2}}}{r_{a2}}\end{bmatrix}, \tag{7a}$$

$$\mathbf{C}(j\omega)=\frac{\rho_0}{4\pi}\begin{bmatrix}\dfrac{e^{-jkr_{11}}}{r_{11}} & \dfrac{e^{-jkr_{12}}}{r_{12}}\\ \dfrac{e^{-jkr_{21}}}{r_{21}} & \dfrac{e^{-jkr_{22}}}{r_{22}}\end{bmatrix}, \tag{7b}$$

where Fig. 1 shows the distances r_{11}, r_{12}, r_{21}, r_{22}, r_{a1}, and r_{a2}. In this case, the inverse of the plant matrix $\mathbf{C}^{-1}$ has a simple analytical solution and Eq. (2) becomes

$$\mathbf{X}(j\omega)=\mathbf{C}(j\omega)^{-1}=\frac{4\pi}{\rho_0}\left(\frac{1}{\dfrac{e^{-jk(r_{11}+r_{22})}}{r_{11}r_{22}}-\dfrac{e^{-jk(r_{12}+r_{21})}}{r_{12}r_{21}}}\right)\times\begin{bmatrix}\dfrac{e^{-jkr_{22}}}{r_{22}} & -\dfrac{e^{-jkr_{12}}}{r_{12}}\\ -\dfrac{e^{-jkr_{21}}}{r_{21}} & \dfrac{e^{-jkr_{11}}}{r_{11}}\end{bmatrix}. \tag{8}$$

The free field computer simulations discussed below employ Eqs. (7), (8). The receivers (ears) are always set at 18 cm apart. The head position changes by varying the distances (r_{11}, r_{12}, etc.) appropriately. The low computation required for this type of simulation provides quick results that provide insight into the basic effects of the geometry.

2. Spherical head approximation

A common approach to improve the approximation is to model the head as a perfectly rigid sphere where the two ears are on the surface at ends of a diameter. By again assuming an anechoic environment and modeling the sound sources as point monopole sources, this model yields an analytical solution. Taking the complex source volume acceleration $j\omega q$ of a point monopole source as the input and acoustic complex pressure p on the surface of the sphere as the output, the total frequency response transfer function $C_t(j\omega)$ is[5]

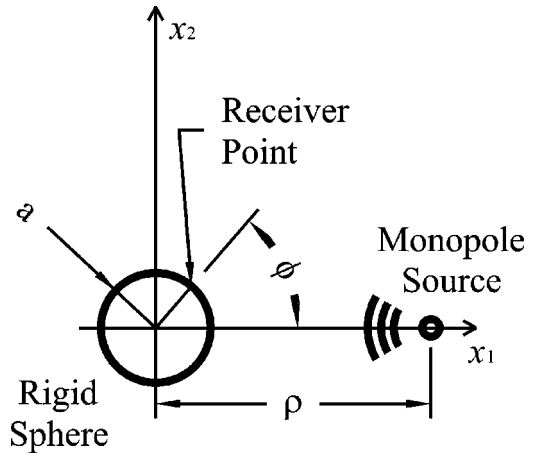


FIG. 3. Variables used when calculating the sound field scattered from a rigid sphere (Ref. 5).

$$C_t(j\omega)=C_{ff}(j\omega)+C_s(j\omega). \tag{9}$$

The two components of Eq. (9) correspond to the incident $C_{ff}(j\omega)$ and scattered $C_s(j\omega)$ sound fields. Equations (10)–(12) calculate these elements:[5]

$$C_{ff}(j\omega)=-\frac{j\rho_0 k}{4\pi}\sum_{m=0}^{\infty}(2m+1)j_m(ka)\times[j_m(kr)-jn_m(kr)]P_m(\cos\phi), \tag{10}$$

$$C_s(j\omega)=\frac{\rho_0 k}{4\pi}\sum_{m=0}^{\infty}b_m[j_m(ka)-jn_m(ka)]P_m(\cos\phi), \tag{11}$$

$$b_m=j(2m+1)\frac{j_m(kr)-jn_m(kr)}{1-\dfrac{jn_m'(ka)}{j_m'(ka)}}. \tag{12}$$

In these equations r is the distance from the source to the center of the sphere, a is the sphere's radius and the angle ϕ is defined in Fig. 3. The functions j_m and n_m are mth-order spherical Bessel functions of the first and second kind, respectively, and P_m is the mth-order Legendre polynomial.

At frequencies near and below 1 kHz, this model well approximates the shadowing effects of a human head. The model's predictions of the binaural localization cue, interaural time delay (ITD), agree closely with measurements made on man-shaped dummies[4,15] and on subjects.[16] The spherical head model's prediction of the binaural localization cue, interaural level difference (ILD), differs by 1–2 dB from measurements made on subjects due to the sphere's lack of a neck.[15] A practical disadvantage of this model is the slow convergence of the Bessel functions making the calculations time consuming. In the simulations that follow, the sphere was set to have a radius of 9 cm.

3. Dummy head

The previous two models made no effort to model the effects of the listener's pinnae, torso, and neck. The last method reviewed here makes use of direct measurements on a KEMAR (Knowles Electronics Manikin for Acoustic Research) dummy.[17] The KEMAR dummy has median human adult dimensions including its pinnae. The acoustic behavior of KEMAR's ear canals and eardrum simulators matches that of real ears. The MIT Media Lab measured HRTFs of a KEMAR dummy in an anechoic chamber at a sampling frequency of 44.1 kHz.[18] This database is available for downloading.[19] The measurements at 0° elevation at a radial distance of 1.4 m for the full 360° of azimuth sampled in 5° increments are employed below with the loudspeaker responses deconvolved.

To match the angle and distance of interest some manipulation of the data is necessary. The first manipulation is linear interpolation in the frequency domain for angle locations between those measured. The next manipulation is to correct for distance by scaling the magnitude by the inverse of distance and by applying a fractional delay FIR filter to account for the time difference. The fractional delay filter is a shifted truncated sinc function.[20] Error introduced by this filter is concentrated around the Nyquist frequency (22.05 kHz) due to the Gibbs phenomenon. The actual time delay added in the simulations is determined by assuming a 344 m/s sound speed.

B. Ringing frequency

Employing the relationship for a geometric series of the type

$$\frac{1}{1-x}=\sum_{n=0}^{\infty} x^n \quad (\text{for } |x|<1) \tag{13}$$

to Eq. (8) yields

$$\mathbf{X}(j\omega)=\frac{4\pi}{\rho_0}\begin{bmatrix} r_{11}e^{jkr_{11}} & -\dfrac{r_{11}r_{22}}{r_{12}}e^{-jk(r_{12}-r_{11}-r_{22})} \\ -\dfrac{r_{11}r_{22}}{r_{21}}e^{-jk(r_{21}-r_{11}-r_{22})} & r_{22}e^{jkr_{22}} \end{bmatrix} \sum_{n=0}^{\infty}\left(\frac{r_{11}r_{22}}{r_{12}r_{21}}\right)^n e^{-jk(r_{12}+r_{21}-r_{11}-r_{22})n}. \tag{14}$$

Equation (14) reveals cross-talk cancellation as inherently a recursive process with an associated ''ringing'' frequency f_r related to the path length differences:

$$f_r=\frac{c_0}{r_{12}+r_{21}-r_{11}-r_{22}}. \tag{15}$$

Employing Eq. (13) to the infinite summation in Eq. (14) yields

$$\frac{1}{1-\dfrac{r_{11}r_{22}}{r_{12}r_{21}}e^{-jk(r_{12}+r_{21}-r_{11}-r_{22})}} \tag{16}$$

which approaches infinity at frequencies $f=0, f_r, 2f_r, 3f_r, \cdots$ for $(r_{11}r_{22}/r_{21}r_{12})\approx 1$. The next subsection shows these frequencies to be ill-conditioned and therefore not very robust.

Figure 4(a) shows the ringing frequency for listener positions 0–2 m off-axis for the 10° loudspeaker arrangement shown in Fig. 1. At the on-axis location, f_r is about 11 kHz. As the listener moves off-axis, the ringing frequency increases monotonically. At about 110 cm off-axis f_r passes beyond the audible frequency range.

C. Condition number

The condition number with respect to matrix inversion is the ratio of the matrix's maximum to minimum singular value. This is a measure (for a system of linear equations) of the solution's sensitivity or ''vulnerability'' to small perturbations or errors. A well-conditioned matrix has a condition number equal or close to one. An ill-conditioned matrix has a large condition number. In this case, the interest is in the sensitivity to error of system inversion of an assumed plant matrix **C**. A source of error in plant matrix **C** is in differences of the HRTF model used and the actual listener's HRTF. These differences include head displacements from the assumed listener location. Therefore, **C**'s condition number reflects the system's robustness to head displacements.[21]

Figures 4(b)–(d) show the condition number of **C** as a function of frequency for head locations 0–2 m off-axis for the free field approximation [Fig. 4(b)], spherical head model [Fig. 4(c)], and for the KEMAR dummy [Fig. 4(d)]. In these figures, the abscissa represents the head position relative to the inter-source axis and the ordinate represents frequency. The gray scale represents the condition number with black and white corresponding to low and high condition numbers, respectively. The free field and sphere models both show ill conditioning at the ringing frequency and at very low frequencies. The dummy HRTF is ill conditioned at very low frequencies but also at other frequencies especially at farther off-axis positions. The dummy is also ill conditioned around 8 kHz for head positions greater than 1 m off-axis.

The condition number calculations using the free field and spherical head approximations predict off-axis positions to be more robust over a greater frequency range than on-axis. When including pinna and torso effects the condition number of the dummy predicts a more robust system on-axis.

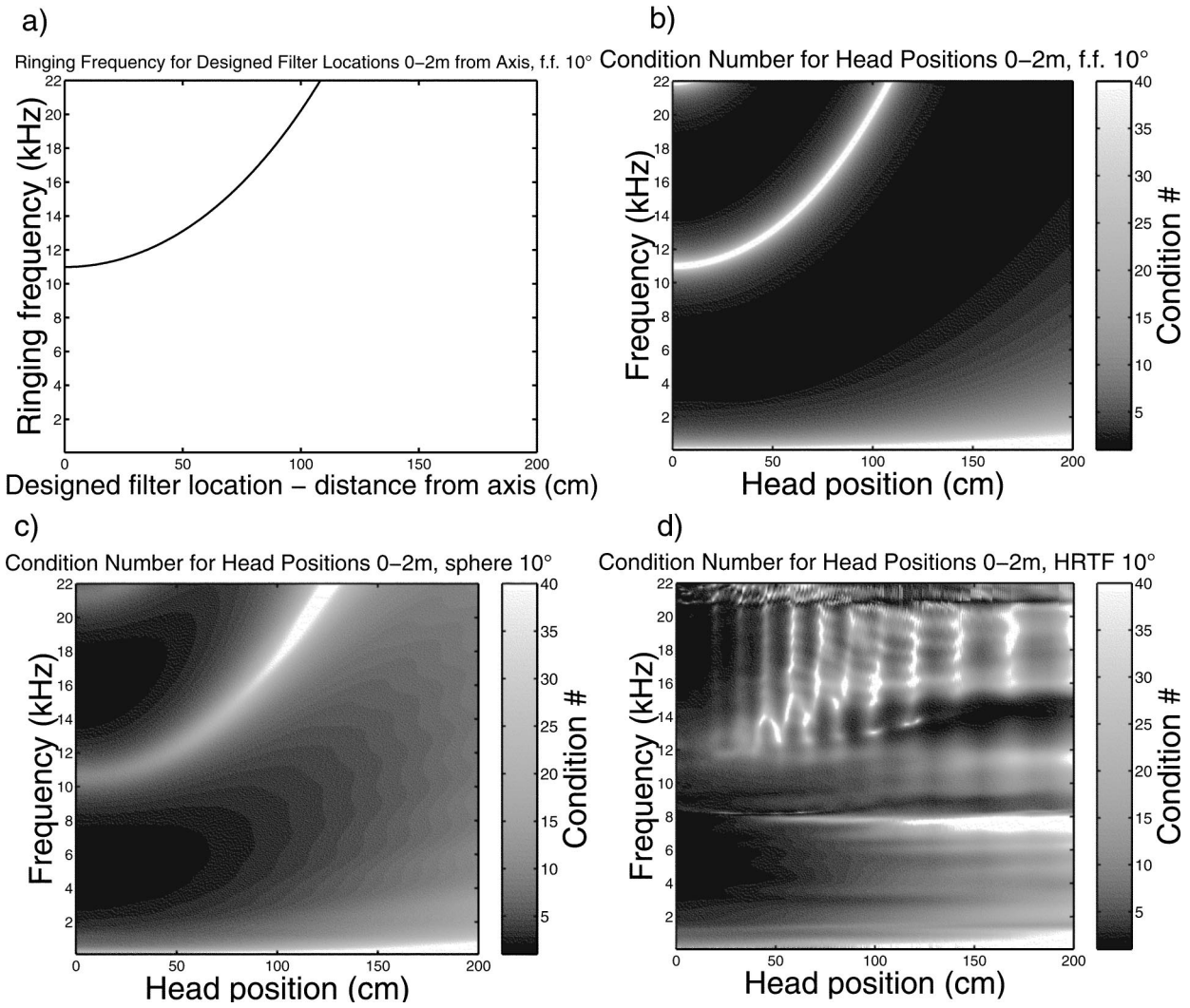


FIG. 4. Calculated results for the 10° loudspeaker arrangement at head positions 0–2 m off-axis: (a) free field ringing frequency; (b) free field condition number (maximum values of 280 000 at 1 Hz and 1100 at 22 kHz); (c) rigid sphere model condition number (maximum values of 13 000 at 1 Hz and 74 at 22 kHz); and (d) KEMAR dummy condition number (maximum values of 50 000 at 1 Hz, 2900 at 8 kHz, and 1200 at 16 kHz).

D. Cross-talk cancellation effectiveness

The effect of the listener's location on cross-talk cancellation performance of the system is now considered. Recall matrix **R** defined by Eq. (4) in Sec. II. Perfect cross-talk cancellation occurs when the diagonal terms of **R**, or the direct paths, are time shifted delta functions in the time domain and the off-diagonal terms of **R**, or the cross-talk paths, are zero. Channel separation refers to a ratio of two of these responses to one ear (e.g., R_{12}/R_{11} or R_{21}/R_{22}).[4] This is essentially a measure of the effectiveness of the cross-talk cancellation process.

Figure 5 shows the magnitudes of the frequency responses R_{11} and R_{12} for head displacements 0.1 mm and 5 cm from optimal locations on-axis and 1 m off-axis with the free field approximation. Optimally, $|R_{11}|$ should be 0 dB over the whole frequency range and $|R_{12}|$ should be minus infinity.

Figures 5(a), (b) show the results for the on-axis optimal location. For a 0.1 mm head displacement [Fig. 5(a)] $|R_{11}|$ is very flat over the whole frequency range and $|R_{12}|$ is predominantly below -40 dB. A peak in $|R_{12}|$ occurs at the ringing frequency at about 11 kHz. A head displacement of 5 cm degrades performance greatly. At this head displacement, Fig. 5(b) shows about 10 dB cross-talk cancellation below 5 kHz but poor performance above this frequency.

Figures 5(c), (d) show the results for the 1 m off-axis optimal location. The ringing frequency is now at about 20 kHz due to a smaller path length difference at this listener location. This has an effect of expanding the frequency range of the cross-talk cancellation performance. The ''sweet spot'' includes higher frequencies at off-axis listener locations because of a higher ringing frequency. Rose *et al.*[22] look at this in more detail.

Designating at least 10 dB cross-talk cancellation performance in the frequency range 0.3–3 kHz at the left ear as the criteria for the boundary of the ''sweet spot'' yields Fig. 6. This range of frequencies corresponds with the subjective experiment discussed in the next section. The authors feel the choice of 10 dB cross-talk cancellation to be a reasonable estimate of the amount needed in order to ensure a desired subjective perception. In Fig. 6, the abscissas are the designed filter location relative to on-axis. The ordinates are the

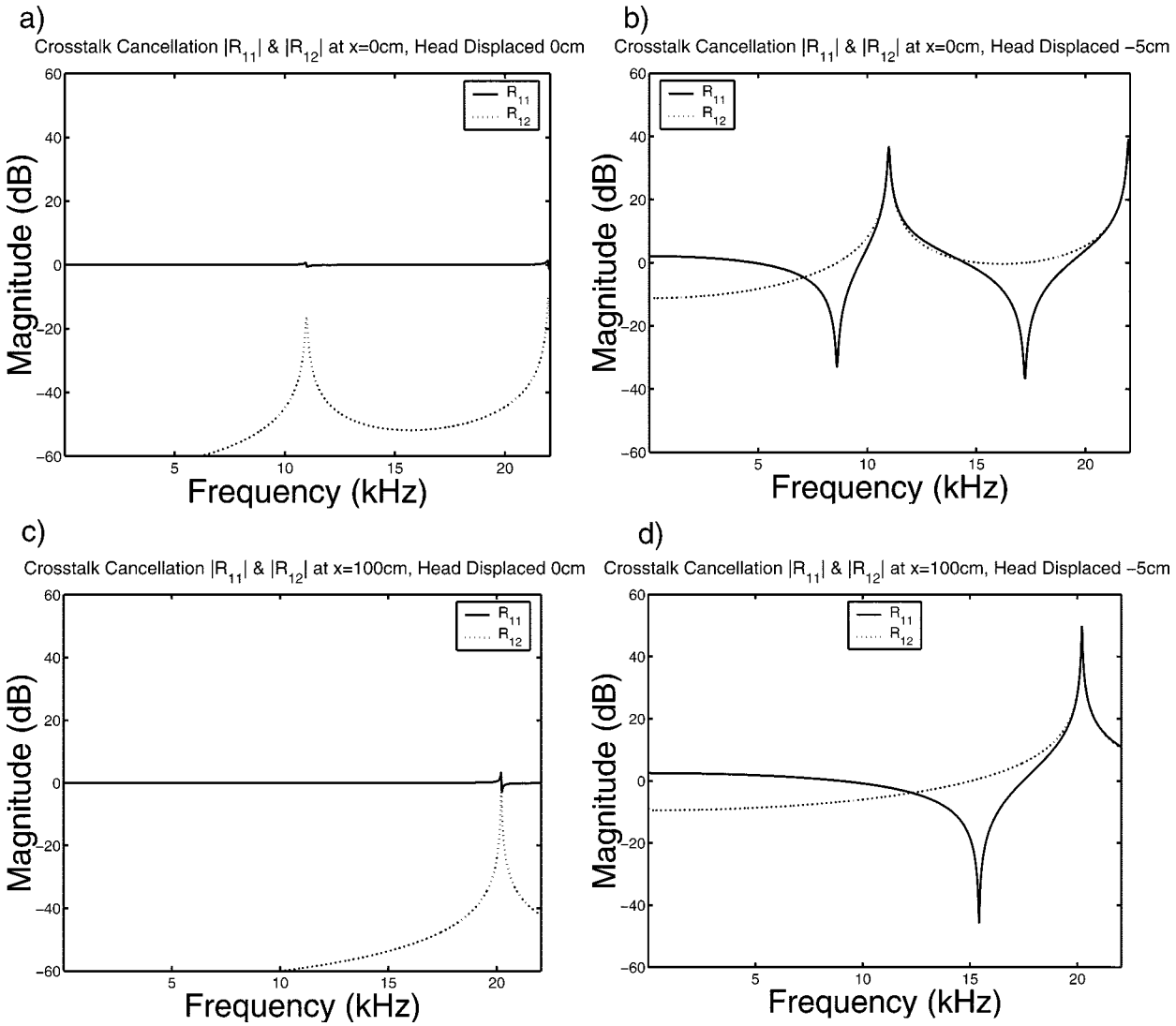


FIG. 5. Elements $|R_{11}|$ and $|R_{12}|$ of control performance matrix **R** for filters designed at (a) the on-axis head location when the head is displaced 0.1 mm from optimum, (b) the on-axis head location when the head is displaced 5 cm from optimum, (c) 1 m off-axis head location when the head is displaced 0.1 mm from optimum, and (d) 1 m off-axis head location when the head is displaced 5 cm from optimum. All the calculations are based on the free field approximation.

head movements from the optimal position. The solid lines give the ''sweet spot'' size for the range of head positions ± 2 m off-axis as determined by an achievement of at least 10 dB cross-talk cancellation performance. Section IV discusses the subjective examination of the designed filter positions within the dashed lines of these figures.

The basis for Fig. 6(a) is the free field model. The ''sweet spot'' size for the ipsi-lateral ear (ear closer to the sources) is almost constant at about ± 4 cm$-\pm 6$ cm from the optimal position. The ''sweet spot'' size for the contra-lateral ear (ear farthest from the sources) tolerates a little more movement away from the sources at farther off-axis positions. For example, at positions greater than 1 m off-axis the contra-lateral ear has a ''sweet spot'' size of 10 cm away from the sources and 5 cm toward the sources.

The basis for Fig. 6(b) is the spherical head model. The effect of head shadowing greatly increases the robustness of cross-talk cancellation at the contra-lateral ear, especially at farther off-axis positions. The ipsi-lateral ear becomes less robust at farther off-axis positions, until it reaches about ± 3 cm at positions farther than 50 cm off-axis. The limiting case might be the ipsi-lateral ear, making the expectation to be a decrease in ''sweet spot'' size for off-axis listening.

The basis for Fig. 6(c) is the KEMAR dummy HRTFs. The ''sweet spot'' has the same general shape for this model as for the spherical head model. The measurements of the dummy with the inherent complexities of its shape introduce some randomness to the ''sweet spot'' and generally reduce its size by 1–2 cm throughout.

The ''sweet spot's'' frequency range increases at off-axis listening positions because of a smaller path length difference at off-axis positions. Head shadowing counters this benefit by decreasing the ''sweet spot'' size for the ipsi-lateral ear (ear closer to the sources). However, head shadowing increases the ''sweet spot'' size for the contra-lateral ear (ear farthest from the sources) at off-axis listener locations. In practice, the expansion of the frequency range of the ''sweet spot'' to include higher frequencies might not matter if the sound includes low frequencies (below 1.5 kHz) due to the dominance of the low frequency interaural time difference

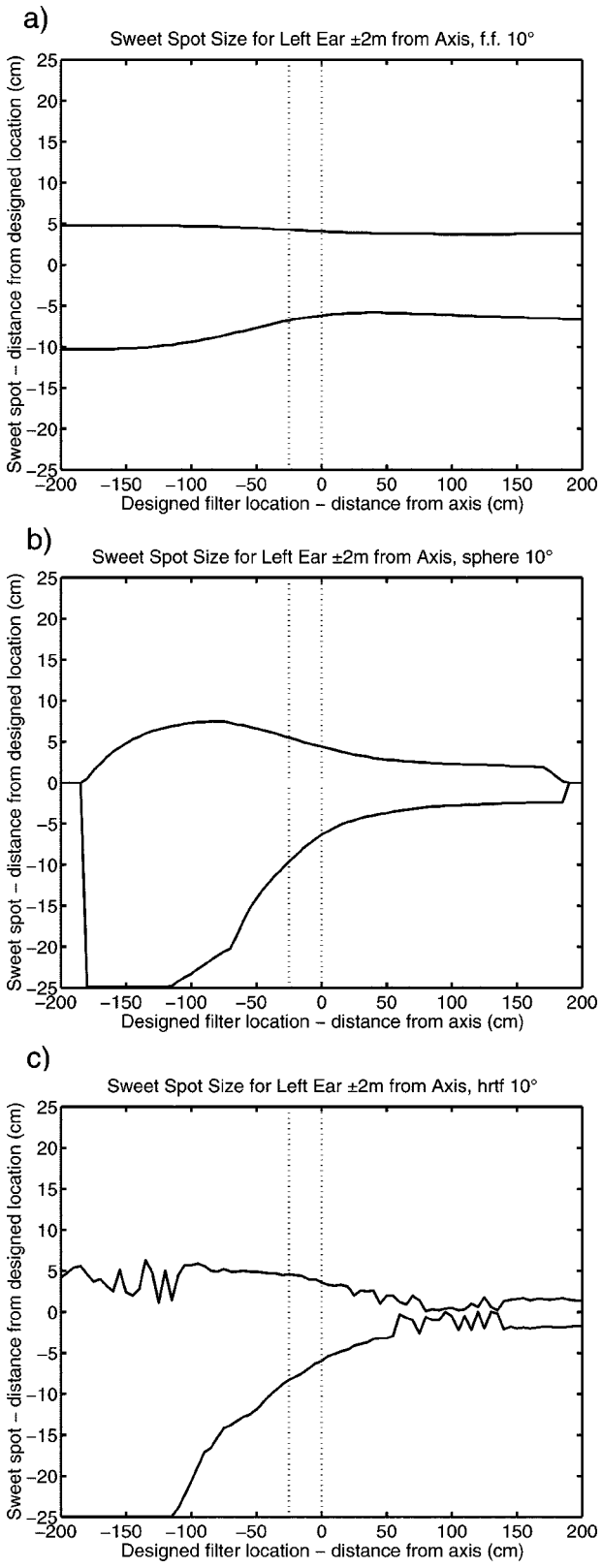


FIG. 6. Boundaries of the sweet spot for a range of head locations ±2 m off-axis as calculated from a 10 dB cross-talk cancellation performance criterion in the frequency range 0.3–3 kHz at the left ear when using (a) the free field model, (b) spherical head model, and (c) the KEMAR dummy. Dashed lines show the positions examined subjectively in Sec. IV.

cue for sound localization. Within the dashed lines in Fig. 6, the ''sweet spot'' size is fairly constant and allows for head movements up to about 5 cm to one side. This range of positions corresponds to the listener locations examined subjectively in Sec. IV.

E. Calculating the sweet spot boundary from just noticeable interaural time difference

The predominant horizontal localization cue at low frequencies (below 1.5 kHz) is the interaural time difference (ITD) of sound arriving at the two ears.[1,23–25] There is convincing evidence that the brain extracts ITD by performing a running cross-correlation of the two-ear signals.[26–29] Computer models of the human auditory system generally first implement a band-pass filter bank to simulate the frequency selectivity of the cochlea and then perform an interaural cross-correlation (IACC) in each of the frequency bands.[28,30,31] The time lag at which the peaks in the IACC functions occur is then the estimate of ITD. Therefore, ITD is generally a function of critical frequency bandwidth. For the following simulations, a simple low-pass filtering of the two ear signals replaces the filter bank. A single cross-correlation calculation then follows. The reasoning behind this scheme is because humans utilize ITD at only low frequencies and ITD is fairly constant as a function of frequency. The cut-off frequency of the low pass filter is 4 kHz. This has an effect of averaging the results for the critical frequency bandwidths below 4 kHz.

A vector of synthesized HRTFs $\mathbf{q}(z)$ in the z domain is defined as[7]

$$\begin{bmatrix} Q_1(z) \\ Q_2(z) \end{bmatrix} = \mathbf{q}(z) = \mathbf{C}(z)\mathbf{X}(z)\mathbf{a}(z) = \mathbf{C}(z)\mathbf{h}(z). \tag{17}$$

The computation of these synthesized HRTFs $\mathbf{q}(z)$ are for one set of designed virtual acoustic imaging filters $\mathbf{h}(z)$ and initially with the plant matrix $\mathbf{C}(z)$ corresponding to the optimal head location. Inverse fast Fourier transforms are performed on the two HRTFs yielding the discrete time domain synthesized head related impulse responses (HRIRs) $q_1(n)$ and $q_2(n)$. These signals are passed through a low-pass filter with a cut-off frequency of 4 kHz yielding $q_{1L}(n)$ and $q_{2L}(n)$. These two signals are cross-correlated yielding an interaural cross-correlation function $\Psi_{\mathbf{q}_L}(m)$,

$$\Psi_{\mathbf{q}_L}(m) = \sum_{n=0}^{N-m-1} q_{1L}(n) q_{2L}(n+m) \quad 0 \leqslant m < N, \tag{18a}$$

$$\Psi_{\mathbf{q}_L}(-m) = \sum_{n=0}^{N-m-1} q_{1L}(n+m) q_{2L}(n) \quad 0 \leqslant m < N. \tag{18b}$$

The estimate of ITD is the amount of lag where the maximum value of this function occurs,[7] i.e.,

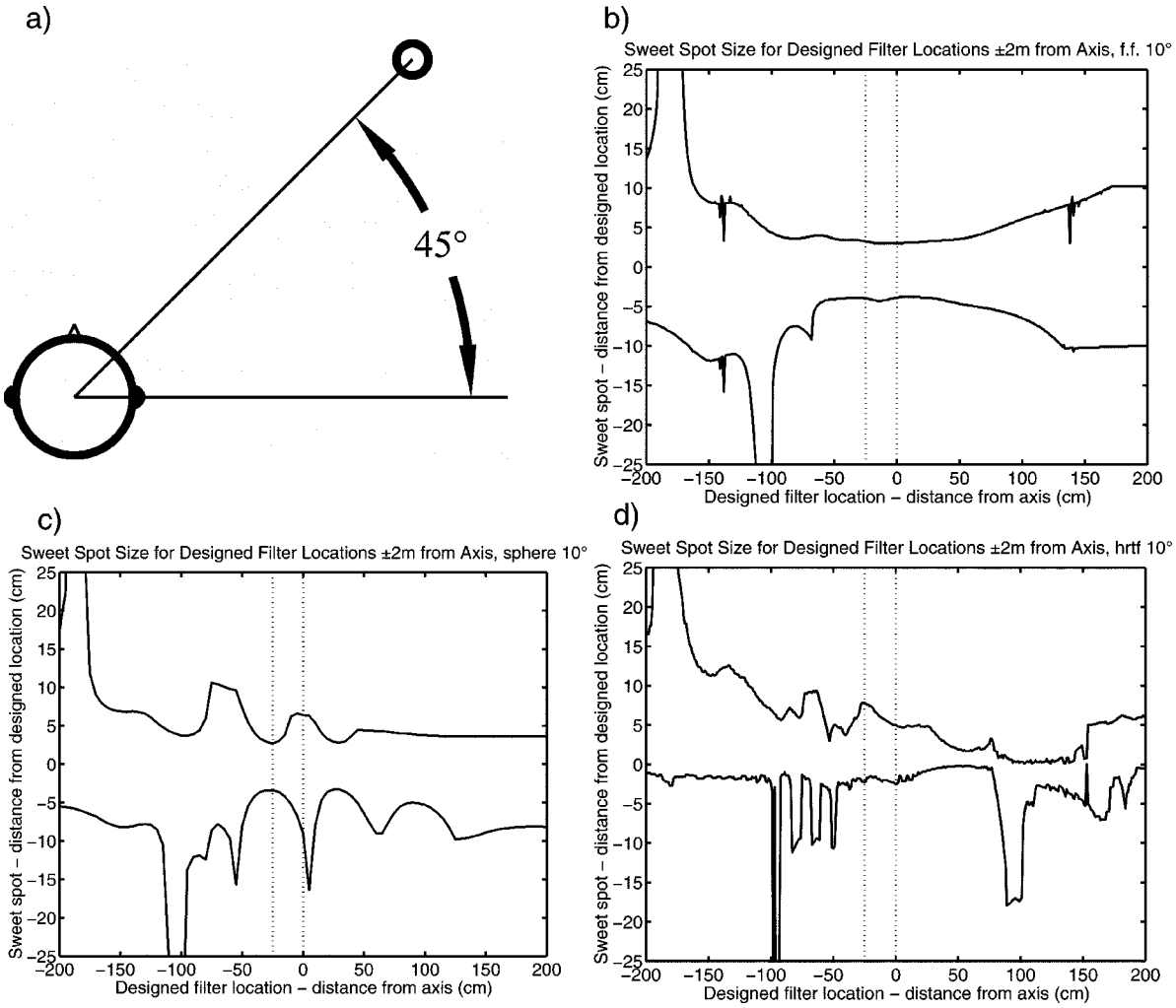


FIG. 7. Boundaries of the sweet spot for a range of head locations ±2 m off-axis as calculated from a 10 μs shift in ITD of the single virtual source represented by the small circle in (a) when using (b) the free field model, (c) spherical head model, and (d) the KEMAR dummy. Dashed lines show the positions examined subjectively in Sec. IV.

$$ITD = (m|_{\max(\Psi_{\mathbf{q}_L})}) \frac{1}{f_s}, \tag{19}$$

where f_s is the sampling frequency.

The initial calculation finds ITD with the listener at the optimal location. Observations of the change in ITD due to incremental head displacements from the optimal listener location then follow. The head position is changed by changing the plant matrix $\mathbf{C}(z)$ in Eq. (17).

Klumpp and Eady have suggested that the just noticeable ITD for humans is about 10 μs.[32] The following designations of ''sweet spot'' boundary correspond to head displacements that introduce this amount of ITD as found by the above method. It is reasonable to think that head displacements less than the size of the ''sweet spot'' as defined by this method should not introduce enough ITD to change the listener's perception of virtual image location.

Figure 7 shows the ''sweet spot'' size when only considering the ITD shift of a single virtual source 45° to the listener's right as predicted by the three HRTF models. The virtual source location corresponds to the subjective experiment of the next section. These figures include results for head positions through the range ±2 m off-axis. Two striking peaks in the ''sweet spot'' size occur at 190 cm and 100 cm to the left of the loudspeakers. The peaks correspond to the listener moving to a position where the real sources are at the location of the virtual source (45° to the listener's right). When considering more than one virtual source these peaks disappear as shown by Rose *et al.*[22] Adding more virtual sources into the calculations has the effect of decreasing the ''sweet spot'' size.

Generally, the free field model predicts a larger ''sweet spot'' for farther off-axis locations while the sphere and dummy models predict a constant ''sweet spot'' size. The free field model predicts a size of about ±4 cm for locations close to on-axis and about ±10 cm farther off-axis. The dummy model predicts the smallest size ''sweet spot'' of about ±2 cm. The sphere model predicts a size of about ±4 cm. Section IV discusses the subjective examination of the designed filter positions within the dashed lines of Fig. 7. The free field model predicts a constant ''sweet spot'' size of about ±3.5 cm at these listener locations. The shadowing effect of the sphere causes the size to vary within the dashed lines but the dummy again predicts a fairly constant size of about ±3 cm.

IV. SUBJECTIVE EXPERIMENT

This section describes a subjective experiment undertaken in an anechoic chamber to investigate the ''sweet spot'' size at a range of head positions in a static case. The subjective results compare to the results of the calculated ''sweet spot'' size of the previous section.

A. Procedure

Band passed white noise with a pass-band of 0.3–3 kHz was used as the source signal. This band of frequencies was chosen to include frequencies which humans localize predominantly by utilizing ITD (below about 1.5 kHz[1,23,25]) and also to include the first ear canal resonance at about 2 kHz[1] but not include pinnae effects.

The subjects sat inside a 2-m-diam sphere covered with a black curtain to hide the position of the loudspeakers. A small headrest supported the subject's head and the subject was asked to limit their movement as much as possible. The two loudspeakers were mounted on a moveable slide outside of the sphere, 1.4 m away from the subject's head. The two loudspeakers subtend an angle of 10° when the listener is in the symmetric head position.[10] The motion of the slide and thereby the loudspeakers was controlled via a computer in the anechoic chamber. The motion of the loudspeakers was limited to the horizontal direction parallel to a line joining the receiver's ears. When asked, the subjects voiced their perception of the direction of the noise. Responses were limited to the horizontal plane at angle locations marked inside the sphere.

The virtual acoustic imaging filters were designed off line with the path responses approximated by monopole sources impinging sound on a perfectly rigid sphere. The designed listener locations examined were when the inter-source axis exactly coincides with the inter-receiver axis and when it is offset 5 cm, 10 cm, 15 cm, 20 cm, and 25 cm to the right. These six filters were examined at loudspeaker locations as far as 15 cm from optimal locations. There were 14 different subjects taking part in the experiment in all. Each subject listened to at least 2 different designed filter positions, although most of the subjects listened to 3 designed filter positions. On average then, each set of virtual acoustic imaging filters were examined by seven different subjects. The virtual source was designed to be at 45° to the front right of the subject in all cases.

B. Results

Figure 8 shows the subjects' responses for the 5 cm off-axis designed virtual acoustic imaging filters. The results for this listener location give an example of the types of results obtained at all six positions examined. This figure shows the difference in the angle perceived from the angle perceived when the loudspeakers are at the optimal location, plotted against the displacement of the loudspeakers from the 5 cm off-axis optimal listener location. The numbers 1–14 designate the different subjects and different symbols represent their individual responses in the plots. Solid lines in these graphs show the mean responses. In addition, error bars show the standard deviation of the data. The standard deviation seems considerable at points but the means show clear trends. Generally, the standard deviation is greater for the greater loudspeaker displacements. Front–back confusions are resolved in Fig. 8. That is, the plots show the listeners' perceptions of the noise originating behind them at the corresponding mirrored angles in front.

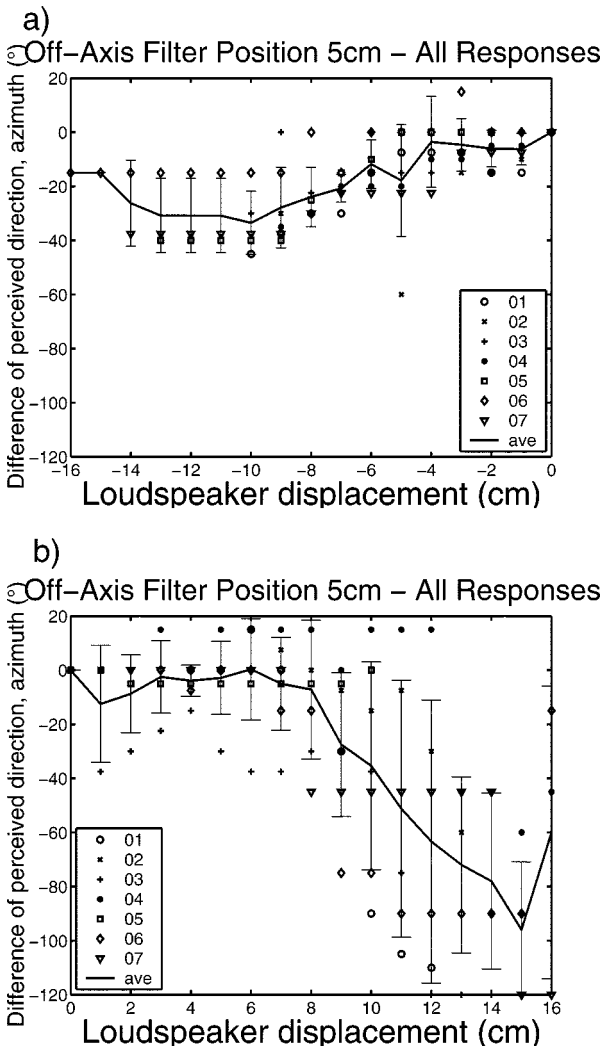


FIG. 8. Subjects' responses for the 5 cm off-axis designed virtual acoustic imaging filter position. Different symbols denote different subjects. Responses are shown with (a) loudspeakers moving toward the designed filter location, from the left and (b) loudspeakers moving toward the designed filter location, from the right.

Figure 9 shows the mean responses for all six of the examined designed filter positions. The six designed filter positions yield similar results. At the optimal listener locations, the subjects generally tended to perceive the direction to be slightly in front of the target location of 45° (e.g., 30°). Other evaluations of virtual sound systems have also found this to occur.[28] Front–back confusions are resolved in Fig. 9.

Figures 8(a), 9(a), show the results with the loudspeakers left of the designed filter listener location while Figs. 8(b), 9(b) show the responses with the loudspeakers to the right of the designed listener location. Comparing the responses with the loudspeakers to the left and right suggest that the virtual sound image is more stable with the loudspeakers to the right. In fact, displacements in this direction

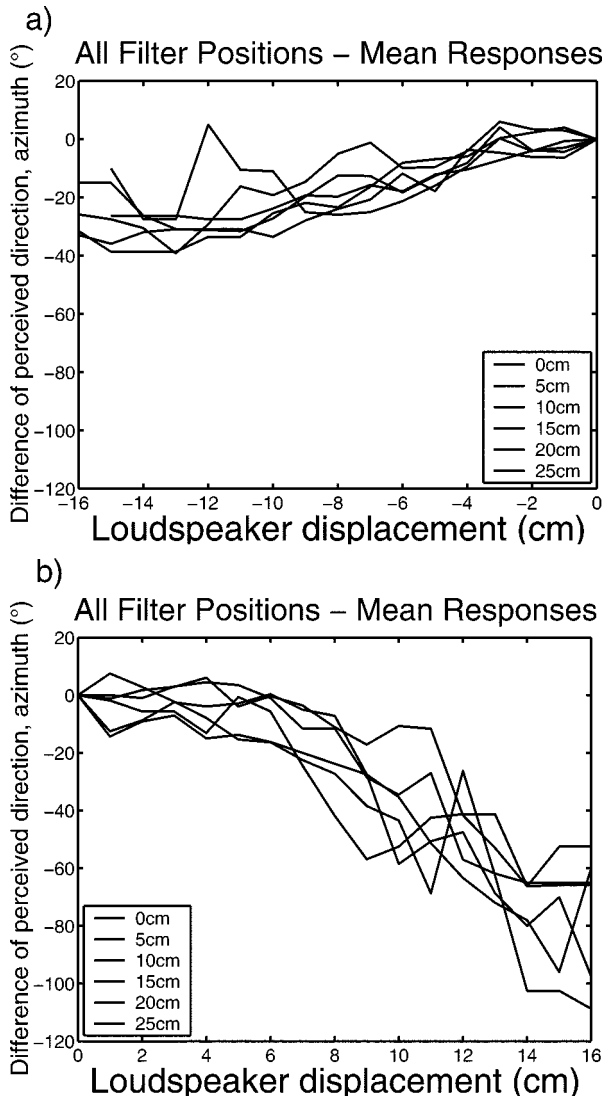


FIG. 9. Average subject responses for the designed virtual acoustic imaging filter positions 0–25 cm off-axis. Different lines represent different designed filter positions. Responses are shown with (a) loudspeakers moving toward the designed filter location, from the left and (b) loudspeakers moving toward the designed filter location, from the right.

correspond to the loudspeakers closer to the virtual source direction. This implies that the system is more robust for head movements away from the virtual source or loudspeaker movements toward the virtual source.

One might interpret these figures to imply a ''sweet spot'' size of about 3 cm to one side of the designed filter listener location (6 cm overall). The virtual image collapses somewhat gradually to the spot directly in front of the listener as the listener moves away from the optimal listening position. As a result, one must decide on how much error is tolerable in order to state the size of the ''sweet spot.'' A 3 cm ''sweet spot'' boundary to one side of the optimal listener location corresponds with the results of the previous section. Rose *et al.*[22] provide further results from this experiment.

C. Limitations of results

While the results of the experiment are useful and provide insight into the problem, this subsection discusses some observations made of the experimental procedure and possible limitations.

(1) The virtual acoustic imaging filters were designed with the path responses approximated by monopole sources impinging sound on a perfectly rigid sphere. Therefore, the subjects' pinnae were not taken into consideration in the filter design. During the experiment, subjects often reported elevation movements. Many subjects localized the source behind them at times. These types of responses may have been avoided if the design of the filters had incorporated the pinnae responses. Figures 8, 9 show the listeners' perceptions from behind them at their corresponding angles in the front.

(2) Toward the end of the experiment, some of the subjects expressed a feeling of fatigue in their right ear. This is because the intended virtual source location was invariably 45° to the subject's right. It would make for a more comfortable and interesting experiment by including some variation in the virtual source angle. This would complicate matters due to a dependence of ''sweet spot'' size on the number and locations of virtual sources as shown by Rose *et al.*[22]

V. CONCLUSION

Virtual acoustic imaging systems with the geometrical arrangement shown in Fig. 1, are as robust to lateral head translations at asymmetric listener locations, that are offset up to 25 cm from the inter-source axis, as the traditional symmetric on-axis listener location. A static subjective experiment suggests that virtual acoustic imaging filter adaptation should occur with lateral head translations of about 3 cm from the designed optimum listener location. This value is in agreement with calculations made using a 10 μs just noticeable interaural time difference threshold and a 10 dB cross-talk cancellation performance threshold. Both the ITD and cross-talk cancellation simulations predict close to the same size ''sweet spot'' making it difficult to define the limiting factor of the ''sweet spot'' size. Farther off-axis listening has an advantage of a higher ringing frequency due to smaller path length differences. Head-shadowing effects decreasing the ''sweet spot'' size at low frequencies offset this advantage. A result worth mentioning is that both the simulations and the subjective experiment find the system to be more robust when the real sources are close to the virtual source location.

[1] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (The MIT Press, Cambridge, MA, 1997).
[2] F. L. Wightman and D. J. Kistler, ''Headphone simulation of free-field listening: II: Psychophysical validation,'' J. Acoust. Soc. Am. **85**, 868–878 (1989).
[3] J. Bauck and D. H. Cooper, ''Generalized transaural stereo and applications,'' J. Audio Eng. Soc. **44**, 683–705 (1996).
[4] W. G. Gardner, ''3-D audio using loudspeakers,'' Ph.D. thesis, MIT Media Laboratory, Cambridge, MA, 1997.
[5] O. Kirkeby, P. A. Nelson, and H. Hamada, ''Local sound field reproduction using two closely spaced loudspeakers,'' J. Acoust. Soc. Am. **104**, 1973–1981 (1998).
[6] B. S. Atal and M. R. Schroeder, ''Apparent sound source translator,'' U.S. Patent 3,236,949 (1962).
[7] T. Takeuchi and P. A. Nelson, ''Robustness to head misalignment of virtual sound imaging systems,'' J. Acoust. Soc. Am. **109**, 958–971 (2001).
[8] O. Kirkeby, P. A. Nelson, and H. Hamada, ''Stereo dipole,'' Patent Application, PCT/GB97/00415, 1997.

[9] P. A. Nelson, O. Kirkeby, T. Takeuchi, and H. Hamada, ''Sound fields for the production of virtual acoustic images,'' J. Sound Vib. **204**, 386–396 (1997).
[10] O. Kirkeby, P. A. Nelson, and H. Hamada, ''The ''stereo dipole'' a virtual source imaging system using two closely spaced loudspeakers,'' J. Audio Eng. Soc. **46**, 387–395 (1998).
[11] J. F. W. Rose, ''A visually adaptive virtual sound imaging system,'' MSc. thesis, Institute of Sound and Vibration Research, Southampton, U.K., 1999.
[12] P. A. Nelson, F. Orduna-Bustamante, and D. Engler, ''Experiments on a system for the synthesis of virtual acoustic sources,'' J. Audio Eng. Soc. **44**, 990–1007 (1996).
[13] F. L. Wightman and D. J. Kistler, ''Headphone simulation of free-field listening. I: Stimulus synthesis,'' J. Acoust. Soc. Am. **85**, 858–867 (1989).
[14] P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, London, 1992), p. 26.
[15] F. A. Firestone, ''The phase difference and amplitude ratio at the ears due to a source of pure tone,'' J. Acoust. Soc. Am. **2**, 260–270 (1930).
[16] W. E. Feddersen, T. T. Sandel, D. C. Teas, and L. A. Jeffress, ''Localization of high-frequency tones,'' J. Acoust. Soc. Am. **29**, 988–991 (1957).
[17] M. D. Burkhard and R. M. Sachs, ''Anthropometric manikin for acoustic research,'' J. Acoust. Soc. Am. **58**, 214–222 (1975).
[18] W. G. Gardner and K. D. Martin, ''HRTF measurements of a KEMAR,'' J. Acoust. Soc. Am. **97**, 3907–3908 (1995).
[19] W. G. Gardner and K. D. Martin, *KEMAR HRTF Measurements* (MIT's Media Lab through http://sound.media.mit.edu/KEMAR.html, 1994).
[20] T. I. Laakso, V. Valimaki, M. Karjalainen, and U. K. Laine, ''Splitting the unit delay: Tools for fractional delay filter design,'' IEEE Signal Process. Mag. **13**, 30–60 (1996).
[21] D. B. Ward and G. W. Elko, ''Effect of loudspeaker position on the robustness of acoustic crosstalk cancellation,'' IEEE Signal Process. Lett. **6**, 106–108 (1999).
[22] J. Rose, P. A. Nelson, B. Rafaely, and T. Takeuchi, ''A study of virtual acoustic imaging systems for asymmetric listener locations,'' ISVR Technical Report No. 295 (2001).
[23] J. C. Middlebrooks and D. M. Green, ''Sound localization by human listeners,'' Annu. Rev. Psychol. **42**, 135–159 (1991).
[24] F. L. Wightman and D. J. Kistler, ''The dominant role of low-frequency interaural time differences in sound localization,'' J. Acoust. Soc. Am. **91**, 1648–1661 (1992).
[25] W. M. Hartmann, ''How we localize sound,'' Phys. Today **52**, 24–29 (1999).
[26] J. C. R. Licklider, ''Auditory frequency analysis,'' in *Information Theory: Third London Symposium*, edited by C. Cherry (Butterworths Scientific, London, 1956), pp. 253–268.
[27] B. McA. Sayers and E. C. Cherry, ''Mechanism of binaural fusion in the hearing of speech,'' J. Acoust. Soc. Am. **29**, 973–987 (1957).
[28] J. C. R. Licklider, ''Three auditory theories,'' in *Psychology: A Study of a Science, Study I. Conceptual and Systematic, Volume 1. Sensory, Perceptual, and Physiological Formulations*, edited by S. Koch (McGraw-Hill, New York, 1959), pp. 41–144.
[29] B. McA. Sayers, ''Acoustic-image lateralization judgments with binaural tones,'' J. Acoust. Soc. Am. **36**, 923–926 (1964).
[30] W. Lindemann, ''Extension of a binarural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals,'' J. Acoust. Soc. Am. **80**, 1608–1622 (1986).
[31] W. Gaik, ''Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling,'' J. Acoust. Soc. Am. **94**, 98–110 (1993).
[32] R. G. Klumpp and H. R. Eady, ''Some measurements of interaural time difference thresholds,'' J. Acoust. Soc. Am. **28**, 859–860 (1956).

Fundamental frequency degeneracy of standing surface acoustic waves under metallic gratings on piezoelectric substrates

A. N. Darinskii
Institute of Crystallography, Academy of Sciences of Russia, Leninskii pr. 59, Moscow 117333, Russia

S. V. Biryukov
Mints Radiotechnical Institute, Moscow 125083, Russia and Institute of Solid State and Materials Research, Dresden, D-01171 Dresden, Germany

M. Weihnacht
Institute of Solid State and Materials Research, Dresden, D-01171 Dresden, Germany

(Received 10 February 2002; revised 3 July 2002; accepted 29 July 2002)

The paper studies specific features of surface acoustic waves on piezoelectric substrates under infinite periodic grating arrays. The grating is made up of identical metallic electrodes of finite thickness and has one electrode per cell. It has been shown via the analysis of numerical algorithms and the coupling-of-modes equations that one of the fundamental frequencies of the standing wave under short-circuited grating necessarily coincides with one of the fundamental frequencies corresponding to the open-circuited grating if the substrate assumes particular orientations connected with its crystallographic symmetry. The existence of degeneracy does not depend on the material constants of the substrate and the electrode, the thickness of the electrode, and the specific electrode shape provided that the electrode is uniform along the wave normal and shaped symmetrically in the same direction. The proof applies both to ordinary and leaky surface waves. [DOI: 10.1121/1.1509433]

PACS numbers: 43.38.Yn [DEC]

I. INTRODUCTION

When a wave travels across a medium with periodically varying properties, its spectrum acquires a band structure, splitting into allowed and forbidden frequency ranges.[1–3] The band edges are the frequencies of the Bragg resonance reflection from the periodic structure. The upper and lower limiting frequencies $\omega^{(\pm)}$ are also the fundamental frequencies of standing modes supported by the periodic structure; the wave vectors of these modes lie on the boundary of the Brillouin zones.

We are concerned with the surface acoustic wave (SAW) propagation under periodic metallic gratings deposited on the surface of a piezoelectric medium. The wave is assumed to move perpendicular to the strips (or electrodes) forming the grating. Such one-dimensional structures have found a wide practical application in modern SAW devices.[4,5]

The grating affects SAW mechanically (mass loading plus the effect of electrode elasticity) and electrically by shorting locally the electric field that accompanies SAW in piezoelectric media. In turn, SAW induces current and voltage on the electrodes so that the SAW characteristics are dependent on how the electrodes are interconnected.

One commonly considers two limiting cases, namely, the open-circuited grating and the short-circuited grating, where the electrodes are completely isolated and short-circuited, respectively.[5] Accordingly, one distinguishes between the upper and lower edges $\omega_{\mathrm{oc}}^{(\pm)}$ and $\omega_{\mathrm{sc}}^{(\pm)}$ of the stop-bands corresponding to the two electrical states of the grating.

The $\omega_{\mathrm{oc,sc}}^{(\pm)}$ values can be estimated using the coupling-of-modes (COM) equations if the relations between the COM parameters and the material constants of the substrate and electrodes as well as geometrical characteristics of the grating are known.[6–11] Apparently the most thorough and self-consistent derivation of the COM parameters has been performed in Ref. 11; however, likewise other papers, Ref. 11 estimates the electrode reflection coefficient only with accuracy of the linear term in electrode thickness–wavelength ratio.

Alternatively, the $\omega_{\mathrm{oc,sc}}^{(\pm)}$ values can be determined numerically by solving the appropriate dispersion equations or finding the zeros and poles of the so-called harmonic admittance Y_h (Ref. 12) employing the fact that Y_h is equal to the ratio of the above dispersion equations (functions).[13] Knowing $\omega_{\mathrm{oc,sc}}^{(\pm)}$ allows the solution of the inverse problem—to estimate the COM parameters and, hence, the characteristics of IDT and reflectors.[5,14]

In what follows, we will be considering a grating formed of identical equidistant strips shaped symmetrically with respect to their centers. When such a grating is on the substrate oriented arbitrarily, the four $\omega_{\mathrm{oc,sc}}^{(\pm)}$ values are different, that is, any frequency edges of the forbidden bands for the open-circuited and short-circuited gratings in a general manner do not coincide. On the other hand, it has been found out that a limiting frequency for the open-circuited grating coalesces with a limiting frequency for the short-circuited grating if the substrate assumes certain orientations. Three possibilities can take place, namely, $\omega_{\mathrm{oc}}^{(-)}=\omega_{\mathrm{sc}}^{(-)}$ or $\omega_{\mathrm{oc}}^{(-)}=\omega_{\mathrm{sc}}^{(+)}$ or $\omega_{\mathrm{oc}}^{(+)}=\omega_{\mathrm{sc}}^{(+)}$. It is apparent that the common standing wave must not induce either voltage or current at the electrodes.

The objective of the present paper is to study the relation between the crystallographic symmetry of the substrate and

the occurrence of the frequency-edge degeneracy. Both pure (nonleaky) SAW and leaky SAW (LSAW) will be considered.

Apart from being interesting scientifically, the degeneracy problem is of importance for technical applications. If two frequencies coalesce, then the cut is naturally bidirectional and IDTs, unless specially designed, will generate the right- and left-propagating waves with equal amplitude. Otherwise, when no degeneracy occurs, the cut is known as naturally unidirectional and any IDT, even fabricated on the basis of the above-mentioned simplest "symmetric" grating, launches the waves towards the right and left with unequal amplitudes.[4,5]

It is noteworthy that the relation between the IDT characteristics and the symmetry of the substrate has been discussed in Ref. 15. Using a simple model of the wave propagation under grating arrays the phase difference between the normal component of displacement and the electric potential has been estimated and some conclusions have been made regarding the position of the transduction center.

Our paper is organized as follows. First we shall discuss the COM equations and general conditions for the frequency degeneracy that this approach yields. Afterwards, in studying the degeneracy problem within the frame of COM theory, we derive a general relation between the squares of the polarization vectors of right- and left-propagating SAW (LSAW) and show that for certain families of cuts this relation becomes such that the degeneracy of $\omega_{\mathrm{oc,sc}}^{(\pm)}$ occurs. In the next two sections, the problem is treated on more rigorous footing. We shall prove that the $\omega_{\mathrm{oc,sc}}^{(\pm)}$ values found numerically also appear to be degenerate for the same cuts.

II. THE COM EQUATIONS

Within the COM theory, our study of the degeneracy problem will be based on the COM equations borrowed from Ref. 11, where these equations have been derived in parallel with the analytical expressions for the parameters they involve.

Let us write the electric potential of SAW on the surface under grating array as a combination of two waves traveling towards the right and left with slowly varying amplitude $R(x)$ and $S(x)$, respectively,

$$\varphi(x)=R(x)e^{i\pi x/l}+S(x)e^{-i\pi x/l}, \tag{1}$$

where l is the period of the grating and the common time-factor $\exp[-i\omega t]$ is omitted. Making use of the impedance method and of the exact solution to the appropriate electrostatic problem[16–18] after a fairly lengthy averaging procedure one obtains[11]

$$\begin{aligned} dR/dx &= iA_1R+iA_2S+i\,DU,\\ dS/dx &= -iA_2^*R-iA_1S-i\,DU,\\ J(x) &= -i\omega[\eta(R+S)+CU], \end{aligned} \tag{2}$$

where U is voltage applied to the electrode as shown in Fig. 1 and J is current. Explicitly, the COM parameters are given in Ref. 11.

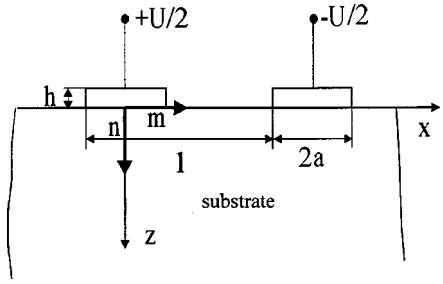


FIG. 1. The geometry of the problem.

For our purpose, it is of importance that the parameter η is real and $D=(\Delta v/v_0)(\eta/\varepsilon_\infty)$, where $\Delta v=v_0-v_m$ is the difference between the wave velocity on the nonmetallized and metallized free surface, respectively, ε_∞ is the limiting value of the effective permittivity of the substrate. In addition, of our concern are A_2 and its complex conjugate A_2^*. They characterize the coupling between the right- and left-propagating waves; $A_2=r/l$, where r is the reflection coefficient from a single electrode.

Equations (2) as obtained in Ref. 11 apply only to SAWs. However, the examination of the derivation procedure reveals that the generalization of (2) to the LSAW problem yields minor modifications in the COM parameters.

In particular, η remains real. The relation between D and η also does not change but D becomes complex because of the factor $\Delta v/v_0$; v_0 and v_m are the LSAW complex velocity on the nonmetallized and metallized free surface, respectively. Like in the case of SAWs, the coefficients before R and S in the first and the second equations, respectively, [A_1 in (2)] remain equal.

The coefficient A_2 goes to $A_2^{(-)}=r^{(-)}/l$ and A_2^* to $A_2^{(+)}=r^{(+)}/l$, where $r^{(-)}$ and $r^{(+)}$ are the reflection coefficients from a single electrode of the left-propagating and right-propagating waves, respectively. Explicitly, using the notation from Ref. 11,

$$r^{(\mp)}=R_e(\alpha)+R_m^{(\mp)}(\alpha)\frac{H}{2l}, \tag{3}$$

where $R_e(\alpha)$ is the common term depending only on the metallization ratio α, H is the electrode thickness,

$$\begin{aligned} R_m^{(\mp)}(\alpha)=\frac{2\pi\sin(\pi\alpha)}{\varepsilon_\infty}\frac{\Delta v}{v_0}\{&[\rho v_0^2+4\mu(1-\chi)]e_x^{(\mp)2}\\ &+(\rho v_0^2+\mu)e_y^{(\mp)2}+\rho v_0^2 e_z^{(\mp)2}\} \end{aligned} \tag{4}$$

(similar expression has been derived in Ref. 10; see also Ref. 19). In (4) $e_{x,y,z}^{(\mp)}$ are the components of the vectors $\mathbf{e}^{(\mp)}=\mathbf{u}^{(\mp)}/\varphi^{(\mp)}$, where $\mathbf{u}^{(-)}$ and $\mathbf{u}^{(+)}$ are the polarization vectors of the left- and right-traveling waves, $\varphi^{(-)}$ and $\varphi^{(+)}$ are the potentials that these waves produce on the surface, respectively, $\chi=\mu/(\lambda+2\mu)$, λ and μ are the Lame coefficients and ρ is the density of the electrode. Equation (4) has been derived assuming that the electrode has a rectangular shape; note that only the factor outside the brackets is shape dependent. [For SAW, $\mathbf{e}^{(-)}=\mathbf{e}^{(+)*}$ and Eq. (4) goes to Eq. (73) in Ref. 11.] If the substrate is taken to be nonpiezoelectric, then (4) reduces to the expression derived in Ref. 20.

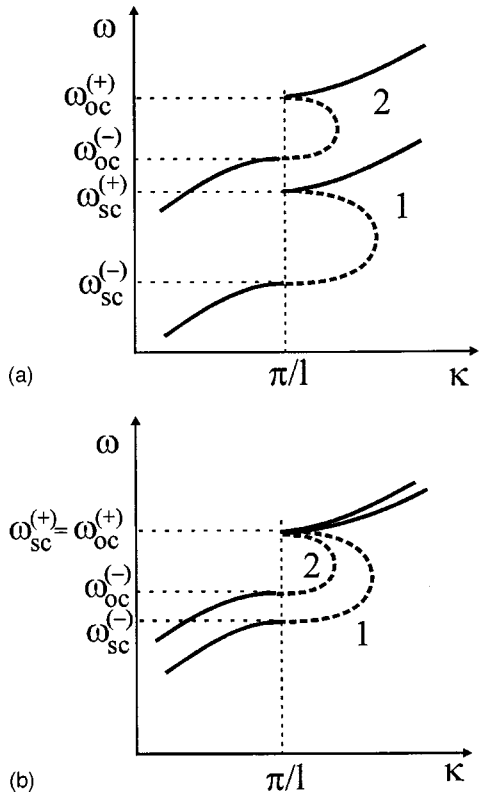


FIG. 2. Dispersion curve obtained using the COM theory for SAW. (1) short-circuited grating, (2) open-circuited grating. Dashed bold line shows the imaginary component of the wave number k within the forbidden zone. (a) All the frequencies are distinct; (b) a degenerate case.

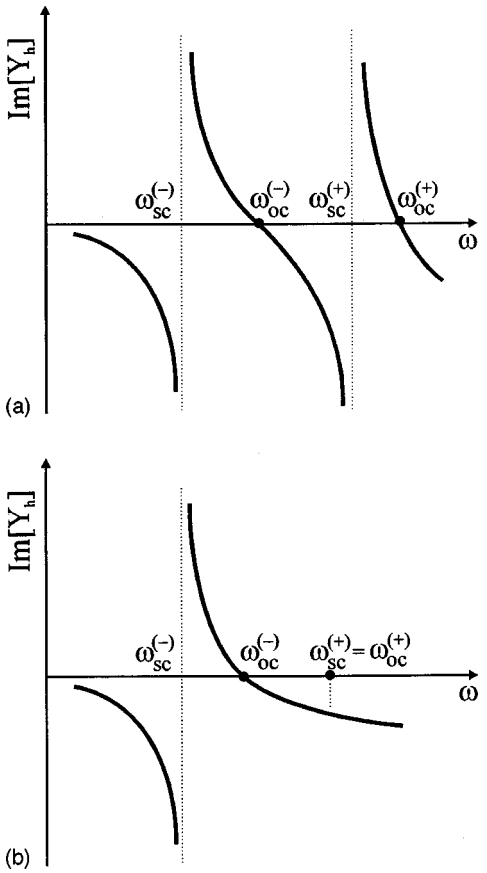


FIG. 3. Frequency dependence of the imaginary part of the harmonic admittance constructed using the COM theory for SAW ($\mathrm{Re}[Y_h]\equiv 0$). (a) No degeneracy; (b) a degenerate case.

The COM equations allow a simple analytical derivation of the wave spectrum under grating arrays (see, e.g., Ref. 5). Figure 2 shows representative dispersion curves that the COM theory yields. Using the COM equation one can also derive an expression for the harmonic admittance $Y_h=I/U$ of the structure per unit cell to obtain $Y_h\propto(\omega-\omega_{oc}^{(+)})(\omega-\omega_{oc}^{(-)})/(\omega-\omega_{sc}^{(+)})(\omega-\omega_{sc}^{(-)})$. In Fig. 3 the typical frequency dependence of Y_h is schematically given for the case when the fundamental frequencies for the short- and open-circuited grating are distinct (a) and degenerate (b). In the latter situation, from the explicit expression it follows that the zeros of the numerator and denominator in Y_h merges so that the right-most discontinuity and zero points in Fig. 3(a) disappear.

The analysis of the dispersion dependencies derived from (2) for open-circuited ($J=0$) and short-circuited ($U=0$) gratings yields that an edge of the forbidden band of the open-circuited grating coalesces with a band edge of the short-circuited one provided $A_2^{(-)}=A_2^{(+)}$ or, accounting for (3),

$$R_m^{(-)}(\alpha)=R_m^{(+)}(\alpha). \tag{5}$$

To summarize, the COM theory yields that the conditions for the frequency degeneracy are those securing the fulfillment of equality (5).

Inspection of (4) reveals that in principle (5) is not a stringent requirement. When the SAW problem is being considered, $R_m^{(-)}(\alpha)\equiv R_m^{(+)*}(\alpha)\equiv R_m(\alpha)$ and (5) reduces to the condition $\mathrm{Im}(R_m)=0$. The reflection coefficient is a function of three orientation angles specifying the geometry of propagation, that is, the orientation of the substrate and the direction of propagation in it. Hence, changing the orientation angles one can fulfill the condition $\mathrm{Im}(R_m)=0$. Since the three angles are subject to one relation, two of them can be viewed as free parameters while the third angle is to be found by solving the equation $\mathrm{Im}(R_m)=0$. Therefore the set of orientations at which $\mathrm{Im}(R_m)$ vanishes is represented as a surface in the three-dimensional space of orientation angles. If one of the angles steadily has the same value, then only two angles are changeable and one obtains a line where the frequencies are degenerate.

We have calculated such a line of degeneracy for the cut family $(\varphi,38°,\psi)$ of $LiNbO_3$ (Fig. 4); the Euler θ angle is fixed and equal to $\theta=38°$ and two other angles are changed [notice that (0,38°,0) cut, or 128°-XY cut, is widely used in SAW devices]. The periodicity of the line with the period $\Delta\varphi=120°$ is due to the crystallographic symmetry of $LiNbO_3$ ($3m$). It will be recalled that to arrive at a cut $(\varphi,38°,\psi)$ the rotation through φ is performed about the z axis that at this stage coincides with the three-fold symmetry axis. The $\pm 60°$ portions of the line are center symmetric with respect to the point (0,0) on the plane (φ,ψ). This is related with the fact that the center of symmetry additionally appears in acoustic problems.

Within the broken regions ($\varphi=60°$, $\psi=90°$ and similar points) the SAW velocity approaches very closely the limit-

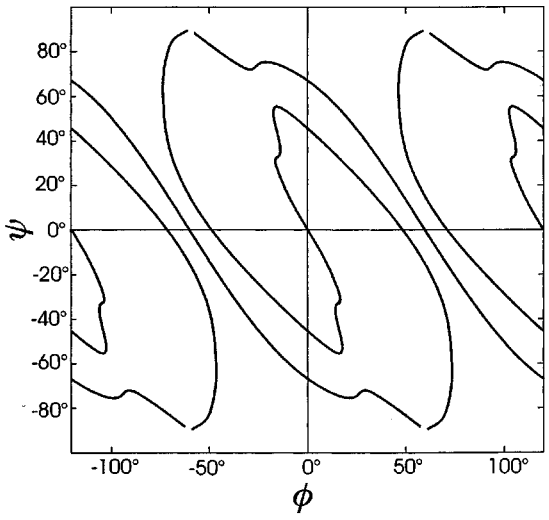


FIG. 4. Line on which reflection coefficient (4) becomes real; (ϕ, 38, ψ) $LiNbO_3$ cut family.

ing velocity of the slow transverse bulk mode and the numerical procedure becomes poor convergent. In addition, it is worth noting that the COM equations as given by (1) and, correspondingly, the expression for $R_m^{\pm}$ do not secure a good approximation near the limiting velocity. In this case, the COM equations derived in Refs. 21 and 22 (see also Ref. 23) are more adequate. However, in these equations the parameters are left unknown and supposed to be found numerically.

As applied to LSAW, the degeneracy of fundamental frequencies happens to be more severely constrained. Since the coefficients $R_m^{(\mp)}(\alpha)$ are not complex conjugate in this case, Eq. (5) implies two conditions so that the set of orientations allowing the degeneracy is at most a line in the three-dimensional space rather than a surface. For ''two-dimensional'' case (like shown in Fig. 4) one has only a point instead of a line.

However, in any of the above cases the degeneracy occurs only approximately with the accuracy to H/l and is lifted off if the higher-order terms are taken into account.

We will focus our attention on a special case: relation (5) holds, because the squared components of $\mathbf{e}^{(+)}$ and $\mathbf{e}^{(-)}$ coincide, $e_q^{(-)2}=e_q^{(+)2}$, $q=x,y,z$. We want to show that the three equalities hold true provided the substrate assumes certain particular orientations which are listed in Sec. V. It is remarkable that the higher-order terms do not lift off the frequency degeneracy at those orientations (Sec. VII).

Subsequent considerations will largely involve the relations given in the next section.

III. BASIC RELATIONS

In studying SAW or LSAW problems it is important to know the decay factor p_α of plane partial modes, the polarization vector $\mathbf{A}_\alpha$, the electric potential Φ_α, the traction $\mathbf{f}_\alpha$, and the normal projection of electric displacement $D_{n\alpha}$; the subscript α labels plane modes. According to Refs. 24 and 25, the above quantities can be found from the eigenvalue problem for an 8×8 matrix $\hat{\mathbf{N}}$,

$$\hat{\mathbf{N}}\boldsymbol{\xi}_\alpha = p_\alpha \boldsymbol{\xi}_\alpha, \quad \alpha=1,...,8, \tag{6}$$

where $\boldsymbol{\xi}_\alpha=(\mathbf{U}_\alpha,\mathbf{V}_\alpha)^t$ are eight-component column vectors constructed from two four-component column vectors

$$\mathbf{U}_\alpha=(\mathbf{A}_\alpha,\Phi_\alpha)^t,$$

$$\mathbf{V}_\alpha=(\mathbf{L}_\alpha,D_\alpha)^t\equiv ik^{-1}(\mathbf{f}_\alpha,D_{n\alpha})^t\equiv ik^{-1}\mathbf{V}'_\alpha$$

with $\mathbf{f}_\alpha=-ik\mathbf{L}_\alpha$ and $D_{n\alpha}=-ikD_\alpha$. The symbol $(\,)^t$ signifies transposition, k is the wave number along the surface.

Explicitly $\hat{\mathbf{N}}$ reads as

$$\hat{\mathbf{N}}=-\begin{Bmatrix} (nn)^{-1}(nm) & (nn)^{-1} \\ (mn)(nn)^{-1}(nm)-(mm) & (mn)(nn)^{-1} \end{Bmatrix}. \tag{7}$$

The symbol (ab) stands for the 4×4 matrices with components $(ab)_{IJ}=a_kE_{kIJl}b_l$, $I,J=1,...,4$, where **a** and **b** is a pair of three-component vectors and $E_{kIJl}=c_{kIJl}-\rho v^2 m_k m_l\delta_{IJ}$, $I,J=1,2,3$, $E_{k4Jl}=e_{kJl}$, $J=1,2,3$, $E_{kI4l}=e_{lIk}$, $I=1,2,3$, $E_{k44l}=-\varepsilon_{kl}$. Here c_{ijkl} are elastic constants, e_{ijk} are piezoelectric constants, ε_{ij} is permittivity, ρ is density, and $v=\omega/k$ is the velocity of a wave field along the surface. The unit vectors **m** and **n** are directed along the x and z axes, respectively (Fig. 1). The matrix $(nn)^{-1}$ is the inverse of (nn).

Note that $\hat{\mathbf{N}}$ depends on v^2 that enters (mm); the other matrices are velocity independent. Since $\hat{\mathbf{N}}$ does not depend on the sign of v, the set of eigensolutions to (6) is the same at v and $-v$. We assign the subscripts such that $\alpha=1,...,4$ label the eigensolutions associated with the right-propagating ($\propto\exp\{i[k(x+p_\alpha z)-\omega t]\}$, $k>0$) nonuniform decaying modes or uniform waves carrying the energy into the substrate (reflected modes); the mode classification at real v is meant. The four remaining eigensolutions $\alpha=5,...,8$ then correspond to the right-propagating nonuniform growing modes or the uniform modes with energy flux directed towards the surface (incident modes). At the same time it is seen that these four eigensolutions are also associated with the left-propagating ($\propto\exp\{i[k(x+p_\alpha z)-\omega t]\}$, $k<0$) decaying or reflected modes. Thus to construct the right- and left-propagating SAW or LSAW the partial modes $\alpha=1,...,4$ and $\alpha=5,...,8$ are available, respectively.

Let us denote the amplitude of a partial mode α by $b_\alpha^{(\pm)}$. The sign $(+)$ or $(-)$ indicates that the mode propagates to the right or the left, respectively. The resultant vectors

$$\mathbf{U}^{(+)}=\sum_{\alpha=1}^{4} b_\alpha^{(+)}\mathbf{U}_\alpha, \quad \mathbf{V}^{(+)}=ik^{-1}\mathbf{V}^{(+)\prime}=\sum_{\alpha=1}^{4} b_\alpha^{(+)}\mathbf{V}_\alpha, \tag{8}$$

$$\mathbf{U}^{(-)}=\sum_{\alpha=5}^{8} b_\alpha^{(-)}\mathbf{U}_\alpha, \quad \mathbf{V}^{(-)}=ik^{-1}\mathbf{V}^{(-)\prime}=\sum_{\alpha=5}^{8} b_\alpha^{(-)}\mathbf{V}_\alpha, \tag{9}$$

can be connected with the help of the acoustic admittances $\hat{\mathbf{Y}}^{(\pm)}$,

$$\mathbf{U}^{(\pm)}=\pm i\hat{\mathbf{Y}}^{(\pm)}\mathbf{V}^{(\pm)}=\mp k^{-1}\hat{\mathbf{Y}}^{(\pm)}\mathbf{V}^{(\pm)\prime}. \tag{10}$$

In terms of $\mathbf{U}_\alpha$ and $\mathbf{V}_\alpha$ the admittances read as

$$\hat{\mathbf{Y}}^{(+)}=-i\sum_{\alpha=1}^{4}\mathbf{U}_\alpha\otimes\mathbf{V}_\alpha^{-1}, \quad \hat{\mathbf{Y}}^{(-)}=i\sum_{\alpha=5}^{8}\mathbf{U}_\alpha\otimes\mathbf{V}_\alpha^{-1}, \tag{11}$$

where the symbol $\otimes$ stands for dyadic multiplication [$\mathbf{a}\otimes\mathbf{b}$ is a matrix with components $(\mathbf{a}\otimes\mathbf{b})_{ij}=a_ib_j$], $\mathbf{V}_\alpha^{-1}$ is the "inverse" of $\mathbf{V}_\alpha$, that is, $\mathbf{V}_\alpha^{-1}\cdot\mathbf{V}_\beta=\delta_{\alpha\beta}$ for $\alpha,\beta=1,...,4$ and $\alpha,\beta=5,...,8$. For instance, $\mathbf{V}_1^{-1}=[\mathbf{V}_2\mathbf{V}_3\mathbf{V}_4]/|\mathbf{V}_1\mathbf{V}_2\mathbf{V}_3\mathbf{V}_4|$; the notation $[\mathbf{V}_\alpha\mathbf{V}_\beta\mathbf{V}_\gamma]$ designates a vector with the components

$$[\mathbf{V}_\alpha\mathbf{V}_\beta\mathbf{V}_\gamma]_i=\varepsilon_{ijkl}(V_\alpha)_j(V_\beta)_k(V_\gamma)_l,\quad i,j,k,l=1,...,4 \tag{12}$$

and $|\mathbf{V}_\alpha\mathbf{V}_\beta\mathbf{V}_\gamma\mathbf{V}_\mu|$ is a determinant,

$$|\mathbf{V}_\alpha\mathbf{V}_\beta\mathbf{V}_\gamma\mathbf{V}_\mu|=\varepsilon_{ijkl}(V_\alpha)_i(V_\beta)_j(V_\gamma)_k(V_\mu)_l, \tag{13}$$

where ε_{ijkl} is the unit 4×4 completely antisymmetric tensor, $\varepsilon_{1234}=1$, and $(V_\gamma)_i$, $i=1,...,4$, are components of the vector $\mathbf{V}_\gamma$.

It follows from the definition of $\hat{\mathbf{N}}$ that the product $\hat{\mathbf{T}}\hat{\mathbf{N}}$ is a symmetric matrix; $\hat{\mathbf{T}}$ is an 8×8 matrix with the elements $T_{ij}=T_{i+4,j+4}=0$, $T_{i,j+4}=T_{i+4,j}=\delta_{ij}$, $i,j=1,...,4$. This property allows the introduction of the eigenvectors $\boldsymbol{\xi}_\alpha$ such that

$$\boldsymbol{\xi}_\alpha\cdot\hat{\mathbf{T}}\boldsymbol{\xi}_\beta=\mathbf{U}_\alpha\mathbf{V}_\beta+\mathbf{V}_\alpha\mathbf{U}_\beta=\delta_{\alpha\beta}. \tag{14}$$

If normalization (14) is adopted, then

$$\sum_{\alpha=1}^{8}\boldsymbol{\xi}_\alpha\otimes\hat{\mathbf{T}}\boldsymbol{\xi}_\alpha=\begin{pmatrix}\hat{\mathbf{I}} & \hat{\mathbf{O}}\\ \hat{\mathbf{O}} & \hat{\mathbf{I}}\end{pmatrix},$$

or

$$\begin{aligned}&\sum_{\alpha=1}^{8}\mathbf{U}_\alpha\otimes\mathbf{V}_\alpha=\hat{\mathbf{I}},\\ &\sum_{\alpha=1}^{8}\mathbf{U}_\alpha\otimes\mathbf{U}_\alpha=\sum_{\alpha=1}^{8}\mathbf{V}_\alpha\otimes\mathbf{V}_\alpha=\hat{\mathbf{O}},\end{aligned} \tag{15}$$

where $\hat{\mathbf{I}}$ and $\hat{\mathbf{O}}$ are the unit and zero 4×4 matrices, respectively.

IV. POLARIZATION VECTORS

We wish to evaluate and compare the squared components of the polarization vectors

$$\mathbf{e}^{(\pm)}=\mathbf{A}^{(\pm)}/\Phi^{(\pm)}=\left(\sum_\alpha b_\alpha^{(\pm)}\mathbf{A}_\alpha\right)\Big/\left(\sum_\alpha b_\alpha^{(\pm)}\Phi_\alpha\right) \tag{16}$$

associated with SAW (LSAW) propagating on the free nonmetallized surface to the right and left with the same velocity v. The boundary conditions require the vanishing of traction $(\mathbf{L}^{(\pm)}=0)$. In addition the normal component of electric displacement and the electrical potential must be continuous. Combining the latter two conditions yields $D_{n\alpha}^{(\pm)}=\mp k\varepsilon_0\Phi^{(\pm)}$ and it follows from (10) that

$$e_q^{(\pm)}=\varepsilon_0 Y_{q4}^{(\pm)}, \tag{17}$$

where the correspondence $1\Leftrightarrow x$, $2\Leftrightarrow y$, $3\Leftrightarrow z$ is assumed.

We can make use of (11) to represent the components $Y_{q4}^{(\pm)}$ of the admittances in the form

$$\begin{aligned}Y_{q4}^{(+)}&=-i\frac{|\mathbf{P}_1^{(q)}\mathbf{P}_2^{(q)}\mathbf{P}_3^{(q)}\mathbf{P}_4^{(q)}|}{|\mathbf{V}_1\mathbf{V}_2\mathbf{V}_3\mathbf{V}_4|},\\ Y_{q4}^{(-)}&=i\frac{|\mathbf{P}_5^{(q)}\mathbf{P}_6^{(q)}\mathbf{P}_7^{(q)}\mathbf{P}_8^{(q)}|}{|\mathbf{V}_5\mathbf{V}_6\mathbf{V}_7\mathbf{V}_8|},\end{aligned} \tag{18}$$

where $\mathbf{P}_\alpha^{(q)}=(A_{\alpha q},L_{\alpha x},L_{\alpha y},L_{\alpha z})^t$ are four-component vectors. Note that the third identity in (15) entails

$$|\mathbf{V}_1\mathbf{V}_2\mathbf{V}_3\mathbf{V}_4|^2=|\mathbf{V}_5\mathbf{V}_6\mathbf{V}_7\mathbf{V}_8|^2. \tag{19}$$

Hence, the values of the determinants in the numerator of expressions (18) decide whether the equalities $e_q^{(+)2}=e_q^{(-)2}$ hold.

Combining identities (15) yields

$$\sum_{\alpha=1}^{8}\mathbf{P}_\alpha^{(q)}\otimes\mathbf{P}_\alpha^{(q)}=\hat{\mathbf{I}}^{(q)\prime}, \tag{20}$$

where $\hat{\mathbf{I}}^{(q)\prime}$ is the 4×4 matrix with two nonzero elements $I_{1,q+1}^{(q)\prime}=I_{q+1,1}^{(q)\prime}=1$.

Taking advantage of (20), after some evaluations we arrive at the following relations:

$$\begin{aligned}|\mathbf{P}_1^{(q)}\mathbf{P}_2^{(q)}\mathbf{P}_3^{(q)}\mathbf{P}_4^{(q)}|^2&=\det\left[\sum_{\alpha=1}^{4}\mathbf{P}_\alpha^{(q)}\otimes\mathbf{P}_\alpha^{(q)}\right]\\ &=\det\left[-\sum_{\alpha=5}^{8}\mathbf{P}_\alpha^{(q)}\otimes\mathbf{P}_\alpha^{(q)}\right]+\Delta^{(q)}\\ &=|\mathbf{P}_5^{(q)}\mathbf{P}_6^{(q)}\mathbf{P}_7^{(q)}\mathbf{P}_8^{(q)}|^2+\Delta^{(q)}.\end{aligned} \tag{21}$$

In (21),

$$\Delta^{(q)}=\frac{(-1)^{q+1}}{4i}\det\begin{pmatrix}S_{q1} & B_{p1} & B_{r1}\\ S_{q2} & B_{p2} & B_{r2}\\ S_{q3} & B_{p3} & B_{r3}\end{pmatrix},$$

$$q,p,r=1,2,3;\quad p,r\neq q;\quad p<r, \tag{22}$$

is the determinant of the 3×3 matrix formed of the components B_{ij} and S_{ij} of the 4×4 matrices

$$\begin{aligned}\hat{\mathbf{B}}&=2i\left[\sum_{\alpha=1}^{4}\mathbf{V}_\alpha\otimes\mathbf{V}_\alpha\right]=i\left[\sum_{\alpha=1}^{4}\mathbf{V}_\alpha\otimes\mathbf{V}_\alpha-\sum_{\alpha=5}^{8}\mathbf{V}_\alpha\otimes\mathbf{V}_\alpha\right],\\ \hat{\mathbf{S}}&=i\left[2\sum_{\alpha=1}^{4}\mathbf{U}_\alpha\otimes\mathbf{V}_\alpha-\hat{\mathbf{I}}\right]\\ &=i\left[\sum_{\alpha=1}^{4}\mathbf{U}_\alpha\otimes\mathbf{V}_\alpha-\sum_{\alpha=5}^{8}\mathbf{U}_\alpha\otimes\mathbf{V}_\alpha\right],\end{aligned} \tag{23}$$

in (22) and (23) the symbol i is imaginary unity. The matrices $\hat{\mathbf{B}}$ and $\hat{\mathbf{S}}$ have been introduced in connection with the existence theorems for SAW.[24,25]

V. IMPACT OF SYMMETRY

Let us show that all the three $\Delta^{(q)}$ vanish, rendering $e_{x,y,z}^{(+)2}=e_{x,y,z}^{(-)2}$, if the orientation of the substrate pertains to one of the families listed below. Namely,

case 1, the direction of propagation is perpendicular to a plane of symmetry;
case 2, the direction of propagation is parallel to an evenfold symmetry axis;
case 3, the surface of the substrate coincides with a plane of symmetry;
case 4, the surface of the substrate is perpendicular to an evenfold symmetry axis. (24)

In particular, $(0,\theta,0)$ and $(90,90,\psi)$ cuts of $3m$ crystals, such as $LiNbO_3$ and $LiTaO_3$, fall into cases 1 and 3, respectively. The same cuts of crystals with symmetry 32 (e.g., SiO_2) refer to cases 2 and 4, respectively.

The above-listed orientations have the feature that the decay parameters appear in pairs $(p_\alpha, -p_\alpha)$ as it is seen straightforward from the Green–Christoffel equations.[26,27] The eight quantities—the non-normalized components of $\mathbf{A}_\alpha$ and $\mathbf{L}_\alpha$ and Φ_α and D_α—become then split into two four-element groups, depending on their evenness with respect to p_α. When normalized according to (14), the vectors $\boldsymbol{\xi}_\alpha$ acquire the structure as follows:[28]

$$\begin{aligned} &p_\alpha = -p_\beta, \quad \boldsymbol{\xi}_\alpha = i(A_x, A_y, A_z, \Phi, L_x, L_y, L_z, D)^t; \\ &\text{case 1,} \quad \boldsymbol{\xi}_\beta = (A_x, -A_y, -A_z, -\Phi, -L_x, L_y, L_z, D)^t; \\ &\text{case 3,} \quad \boldsymbol{\xi}_\beta = (A_x, A_y, -A_z, \Phi, -L_x, -L_y, L_z, -D)^t. \end{aligned} \tag{25}$$

For cases 2 and 4 (24), the only difference from cases 1 and 3, respectively, is in interchanged "behavior" of Φ and D; $\boldsymbol{\xi}_\beta = (\ldots, \Phi, \ldots, -D)^t$, case 2, and $(\ldots, -\Phi, \ldots, D)^t$, case 4.

From Eq. (25) it follows that some of the components of $\hat{\mathbf{B}}$ and $\hat{\mathbf{S}}$ (23) vanish identically at orientations (24). Indeed, letting for the sake of definiteness $p_\alpha = -p_{\alpha+4}$, $\alpha = 1, 2, 3, 4$, we can insert the relevant components of $\boldsymbol{\xi}_\alpha$ and $\boldsymbol{\xi}_\beta \equiv \boldsymbol{\xi}_{\alpha+4}$ (25) into (23) to obtain

$$B_{ij} = 0, \quad i = 1, \quad j = 2,3,4; \tag{26}$$
$$S_{ij} = 0, \quad i = j = 1, \quad i,j = 2,3,4 \ \text{(case 1)};$$

$$B_{12} = B_{13} = B_{24} = B_{34} = 0, \quad S_{ii} = 0, \quad i = 1,2,3,4, \tag{27}$$
$$S_{23} = S_{32} = S_{14} = S_{41} = 0 \ \text{(case 2)};$$

$$B_{13} = B_{23} = B_{34} = 0, \quad S_{ii} = 0, \quad i = 1,2,3,4, \tag{28}$$
$$S_{12} = S_{21} = S_{14} = S_{41} = S_{24} = S_{42} = 0 \ \text{(case 3)};$$

$$B_{13} = B_{14} = B_{23} = B_{24} = 0, \tag{29}$$
$$S_{ij} = 0, \quad i,j = 1,2; \ i,j = 3,4 \ \text{(case 4)}.$$

One sees that $S_{ij} = 0$ if $B_{ij} \neq 0$, and vice versa.

Now, it can be checked that the determinants $\Delta^{(q)}$, accounting for (26)–(29), become zero.

As a result, $|\mathbf{P}_1^{(q)}\mathbf{P}_2^{(q)}\mathbf{P}_3^{(q)}\mathbf{P}_4^{(q)}|^2$ equals $|\mathbf{P}_5^{(q)}\mathbf{P}_6^{(q)}\mathbf{P}_7^{(q)}\mathbf{P}_8^{(q)}|^2$ and, hence, $e_{x,y,z}^{(+)2} = e_{x,y,z}^{(-)2}$ when the substrate is oriented as indicated in (24). Accordingly, equality (7) holds, suggesting the coalescence of two frequencies of the four $\omega_{\mathrm{oc,sc}}^{(\pm)}$'s.

Notice that on curve in Fig. 4 the equalities $e_{x,y,z}^{(+)2} = e_{x,y,z}^{(-)2}$ are fulfilled at $(\varphi = \psi = 0)$ and $(\varphi = 60°, \psi = 0)$ and similar orientations; they fall into case 1 [due to symmetry in $LiNbO_3$ $(60, \theta, 0)$ cut is equivalent to $(0, -\theta, 0)$ one].

VI. NUMERICAL PROCEDURE

Using approximate expressions for the COM parameters, we have shown that cuts (24) ensure the frequency degeneracy. Now we wish to prove that, when found numerically, the frequencies $\omega_{\mathrm{oc,sc}}^{(\pm)}$ also fall into degeneracy at orientations (24).

In this section, we outline a method conventionally utilized to compute fundamental frequencies and dispersion dependences for infinite gratings. Emphasis is put on the features of the matrices and functions involved that are of importance for our subsequent discussions of the frequency degeneracy.

Note that the algorithm under discussion yield results that prove to be in a good agreement with measurements, see, e.g., Ref. 29. In addition, the data thus obtained are widely used in designing SAW devices.

Numerical computations use the Green matrix-function $\hat{\mathbf{G}}$ of the substrate (see, e.g., Refs. 30, 31, 32 and 33). This matrix connects the force and the surface density of free charges with the displacement and potential on the surface. In the k-representation, $\hat{\mathbf{G}}$ is known to obey the symmetry relation $\hat{\mathbf{G}}(-k) = [\hat{\mathbf{G}}(k)]^t$ whether $v = \omega/k$ is less, or greater, than the limiting velocity v_L of the slowest bulk wave, or complex valued. Separating the symmetric $\hat{\mathbf{G}}^{(s)}$ and antisymmetric $\hat{\mathbf{G}}^{(a)}$ matrices, we write

$$\begin{aligned} &\hat{\mathbf{G}}(k>0) \equiv \hat{\mathbf{G}}^{(+)}(k) = \hat{\mathbf{G}}^{(s)} + i\hat{\mathbf{G}}^{(a)}, \\ &\hat{\mathbf{G}}(-k) \equiv \hat{\mathbf{G}}^{(-)}(k) = \hat{\mathbf{G}}^{(s)} - i\hat{\mathbf{G}}^{(a)}. \end{aligned} \tag{30}$$

If v is less than v_L and real, $\hat{\mathbf{G}}^{(s),(a)}$ are real matrices and $\hat{\mathbf{G}}$ is an Hermitian matrix.

The electrical boundary conditions are formulated as if the electrodes were infinitesimally thin (the finite thickness of electrodes is taken into account only in the mechanical "part" of the problem). One assumes that the potentials in the substrate and the exterior half-space are equal on the interface. The normal components of electric displacement differ by the free charge density on the surface of the substrate.

As a result, the electrical boundary equations for the right-propagating modes can be cast into the relation $k\Sigma_{\alpha=1}^4 b_\alpha^{(+)}[-iD_\alpha + \varepsilon_0\Phi_\alpha] = \sigma$, where σ denote the k-Fourier component of the free charge density on the surface. Combining the above relation with Eq. (10) one obtains an expression for $\hat{\mathbf{G}}^{(+)}$ in terms of the admittance $\hat{\mathbf{Y}} = \hat{\mathbf{Y}}^{(+)}$ (11).

$$\hat{\mathbf{G}}^{(+)}(k) = \frac{1}{k(\varepsilon_0 Y_{44} - 1)} \times \begin{bmatrix} (1 - \varepsilon_0 Y_{44})Y_{ij} + \varepsilon_0 Y_{i4}Y_{4j}, & Y_{i4} \\ Y_{4j}, & Y_{44} \end{bmatrix}, \tag{31}$$

where $i, j = 1,2,3$.

To compute the fundamental frequencies, the response of the grating-substrate structure is studied to an external ac

voltage $V(m)=V\exp[i\pi m]$ applied to the electrodes; here m is the electrode number (see Fig. 1). One introduces the so-called periodic Green's function[34–36]

$$\hat{\mathbf{G}}(x)=\frac{1}{l}\sum_{n=-\infty}^{\infty}\hat{\mathbf{G}}(k^{(n)})e^{ik^{(n)}x},$$
$$k^{(n)}=\frac{2\pi}{l}(n+1/2). \tag{32}$$

Due to periodicity one restricts the x value to the range $|x|\leqslant 0.5l$ and a relation is then exploited,

$$\mathbf{U}(x)=\int_{-a}^{a}\hat{\mathbf{G}}(x-x')\mathbf{V}'(x')dx', \tag{33}$$

where a is half-electrode width, $\mathbf{U}(x)=(\mathbf{A}(x),\Phi(x))^t$ and $\mathbf{V}'(x)=(\mathbf{f}(x),\sigma(x))^t$ are constructed from the displacement $\mathbf{A}(x)$, the potential $\Phi(x)$, the force $\mathbf{f}(x)$, and the charge density $\sigma(x)$. The latter two quantities are nonzero only within the under-electrode interval.

The vector $\mathbf{V}'(x)$ is sought for in the form

$$\mathbf{V}'(x)=\sum_{n}\mathbf{C}^{(n)}P^{(n)}(x), \tag{34}$$

where $P^{(n)}(x)$, $n=0,1,...$ is a set of orthogonal polynoms of degree n and the vector $\mathbf{C}^{(n)}=(C_x^{(n)},C_y^{(n)},C_z^{(n)},C_\sigma^{(n)})^t$ involves the three components $C_{x,y,z}^{(n)}$ of force and the charge $C_\sigma^{(n)}$.

Due to (30) and the relation $k^{(n)}=-k^{(-n-1)}$ the matrix-function $\hat{\mathbf{G}}(x)$ decomposes into the symmetric x-even matrix $\hat{\mathbf{G}}^{(s)(e)}(x)$ and the antisymmetric x-odd matrix $\hat{\mathbf{G}}^{(a)(o)}(x)$,

$$\hat{\mathbf{G}}(x)=\hat{\mathbf{G}}^{(s)(e)}(x)+\hat{\mathbf{G}}^{(a)(o)}(x). \tag{35}$$

We can substitute (34) into (33) and integrate to obtain, accounting for (35) and the fact that $P^{(2n)}(x)$ and $P^{(2n+1)}(x)$ are x-even and x-odd functions, respectively,

$$\mathbf{U}(x)=\sum_{n}[\hat{\mathbf{G}}_e^{(s)(2n)}(x)\mathbf{C}^{(2n)}+\hat{\mathbf{G}}_e^{(a)(2n+1)}(x)\mathbf{C}^{(2n+1)}]$$
$$+\sum_{n}[\hat{\mathbf{G}}_o^{(s)(2n+1)}(x)\mathbf{C}^{(2n+1)}$$
$$+\hat{\mathbf{G}}_o^{(a)(2n)}(x)\mathbf{C}^{(2n)}]. \tag{36}$$

Here $\hat{\mathbf{G}}_e^{(s)(2n)}(x)$ and $\hat{\mathbf{G}}_o^{(s)(2n+1)}(x)$ are the convolutions of $\hat{\mathbf{G}}^{(s)(e)}(x)$ with $P^{(2n)}(x)$ and $P^{(2n+1)}(x)$, respectively. $\hat{\mathbf{G}}_o^{(a)(2n)}(x)$ and $\hat{\mathbf{G}}_e^{(a)(2n+1)}(x)$ stand for the convolutions of $\hat{\mathbf{G}}^{(a)(o)}(x)$ with $P^{(2n)}(x)$ and $P^{(2n+1)}(x)$, respectively. The superscripts "(s)" and "(a)" still label symmetric and antisymmetric matrices, respectively.

It is worth having in mind for further use that $\hat{\mathbf{G}}_e^{(s)(2n)}(x)$ and $\hat{\mathbf{G}}_e^{(a)(2n+1)}(x)$ are x-even functions while $\hat{\mathbf{G}}_o^{(s)(2n+1)}(x)$ and $\hat{\mathbf{G}}_o^{(a)(2n)}(x)$ are x-odd functions.

Within $|x|\leqslant a$ the electrical and mechanical boundary conditions must be fulfilled,

$$\Phi(x)=V=\text{const},\quad \mathbf{A}(x)=\mathbf{A}^{(\text{el})}(x),\quad \mathbf{f}(x)=\mathbf{f}^{(\text{el})}(x), \tag{37}$$

where $\mathbf{A}^{(\text{el})}(x)$ and $\mathbf{f}^{(\text{el})}(x)$ are the displacement and traction that the elastic vibrations in the electrode produce at the electrode–substrate interface. In addition, the traction on the free surface of the electrode must vanish.

A convenient method of calculating acoustic fields in the electrode has been implemented in Ref. 37. Following Ref. 37 we seek for $\mathbf{A}^{(\text{el})}$ as

$$\mathbf{A}^{(\text{el})}=\sum_{n,m}\boldsymbol{\lambda}_{nm}\varphi_n(x)\psi_m(z), \tag{38}$$

where $\varphi_n(x)$ and $\psi_m(z)$ are sets of basis functions and $\boldsymbol{\lambda}_{nm}$ are such that functions (38) minimize the Lagrangian[38]

$$\mathcal{L}=0.5\int[\Sigma_{ij}^{(\text{el})}A_{ij}^{(\text{el})}-\rho\omega^2A_i^{(\text{el})2}]dS+\int_{-a}^{a}f_iA_i^{\text{el}}\,dx, \tag{39}$$

where $\Sigma_{ij}^{(\text{el})}$ are the mechanical stresses produced by strains $A_{ij}^{(\text{el})}$, $f_i(x)$ are the components of the force at the interface, ρ is the density of the electrode. The integration in the first term is performed over the electrode cross section.

Consider that the electrode is elastically isotropic and shaped symmetrically. In this instance, it can be checked that the x-odd x-component $A_x^{(\text{el})}(x,z)$ is coupled with the x-even z-component $A_z^{(\text{el})}(x,z)$ and the x-even $A_x^{(\text{el})}(x,z)$ with the x-odd $A_z^{(\text{el})}(x,z)$. Therefore the appropriate combinations of $A_{x,z}^{(\text{el})}(x,z)$ can be sought for as two independent solutions.

The $\boldsymbol{\lambda}_{nm}$'s are found from the appropriate systems of linear equations with $\psi_m(0)\int_{-a}^{a}f_i(x)\varphi_n(x)dx$ on the right-hand side. Having in mind how x-odd and x-even displacements are "mixed" allows the following observations to be made. The x-even forces f_x (or $C_x^{(2n)}$) contribute to x-even $A_x^{(\text{el})}$ and x-odd $A_z^{(\text{el})}$; the x-odd forces f_x (or $C_x^{(2n+1)}$) contribute to x-odd $A_x^{(\text{el})}$ and x-even $A_z^{(\text{el})}$. The x-even forces f_z (or $C_z^{(2n)}$) contribute to x-odd $A_x^{(\text{el})}$ and x-even $A_z^{(\text{el})}$; the x-odd forces f_z (or $C_z^{(2n+1)}$) contribute to x-even $A_x^{(\text{el})}$ and x-odd $A_z^{(\text{el})}$. At last, the x-even forces f_y (or $C_y^{(2n)}$) contribute to x-even $A_y^{(\text{el})}$; the x-odd forces f_y (or $C_y^{(2n+1)}$) contribute to x-odd $A_y^{(\text{el})}$.

To summarize, the resultant $A_{x,y,z}^{(\text{el})}(x)$ on the interface can be written as

$$A_x^{(\text{el})}(x)=\sum_{n}[K_o^{(2n+1)}(x)C_x^{(2n+1)}+K_o^{(2n)}(x)C_z^{(2n)}]$$
$$+\sum_{n}[K_e^{(2n)}(x)C_x^{(2n)}+K_e^{(2n+1)}(x)C_z^{(2n+1)}], \tag{40}$$

$$A_y^{(\text{el})}(x)=\sum_{n}[J_e^{(2n)}(x)C_y^{(2n)}+J_o^{(2n+1)}(x)C_y^{(2n+1)}], \tag{41}$$

$$A_z^{(\text{el})}(x)=\sum_{n}[W_e^{(2n+1)}(x)C_x^{(2n+1)}+W_e^{(2n)}(x)C_z^{(2n)}]$$
$$+\sum_{n}[W_o^{(2n)}(x)C_x^{(2n)}+W_o^{(2n+1)}(x)C_z^{(2n+1)}], \tag{42}$$

where the subscripts "o" and "e" denote x-odd and x-even functions, respectively.

Explicit expressions for $K_o^{(2n+1)}(x)$, etc., are of no concern for us. Of importance is how $C_{x,y,z}^{(2n)}$'s and $C_{x,y,z}^{(2n+1)}$'s are involved in x-odd and x-even components of displacements.

Inserting (40)–(42) into (36) and equating left-hand x-even and x-odd parts to their right-hand counterparts lead to

$$\sum_n [\hat{\mathbf{E}}^{(2n)}(x)\mathbf{C}^{(2n)}+\hat{\mathbf{E}}^{(2n+1)}(x)\mathbf{C}^{(2n+1)}]+\hat{\mathbf{I}}^{(v)}V$$
$$=\sum_n [\hat{\mathbf{G}}_e^{(s)(2n)}(x)\mathbf{C}^{(2n)}+\hat{\mathbf{G}}_e^{(a)(2n+1)}(x)\mathbf{C}^{(2n+1)}],$$
$$\sum_n [\hat{\mathbf{O}}^{(2n)}(x)\mathbf{C}^{(2n)}+\hat{\mathbf{O}}^{(2n+1)}(x)\mathbf{C}^{(2n+1)}]$$
$$=\sum_n [\hat{\mathbf{G}}_o^{(s)(2n+1)}(x)\mathbf{C}^{(2n+1)}+\hat{\mathbf{G}}_o^{(a)(2n)}(x)\mathbf{C}^{(2n)}], \tag{43}$$

where $\hat{\mathbf{E}}^{(n)}(x)$ and $\hat{\mathbf{O}}^{(n)}(x)$, are 4×4 x-even and x-odd matrices, respectively, with elements

$$E_{11}^{(2n)}=K_e^{(2n)},\quad E_{22}^{(2n)}=J_e^{(2n)},\quad E_{33}^{(2n)}=W_e^{(2n)},$$
$$E_{13}^{(2n+1)}=K_e^{(2n+1)},\quad E_{31}^{(2n+1)}=W_e^{(2n+1)},$$
$$O_{13}^{(2n)}=K_o^{(2n)},\quad O_{31}^{(2n)}=W_o^{(2n)}, \tag{44}$$
$$O_{11}^{(2n+1)}=K_o^{(2n+1)},\quad O_{22}^{(2n+1)}=J_o^{(2n+1)},$$
$$O_{33}^{(2n+1)}=W_o^{(2n+1)},$$

the other components are zero. The 4×4 matrix $\hat{\mathbf{I}}^{(v)}$ has only one nonzero component $I_{44}^{(v)}=1$.

Finally, using the method of moments or letting $x=x_0,\ldots,x_n$ in (43) (collocation method), one arrives at a system of $8(n+1)$ equations to determine $\mathbf{C}^{(m)}$, $m=0,\ldots,2n+1$ as functions of the voltage amplitude V.

One computes the harmonic admittance $Y_h=I/V$ as a function of the frequency ω. Here $I=-i\omega Q^{(0)}C_\sigma^{(0)}$ is the current across the electrode and $Q^{(0)}$ is $P^{(0)}(x)$ integrated over the electrode width.

A proof has recently been given by one of the authors in Ref. 13 that Y_h simplifies to the ratio of two determinants $\mathcal{D}^{\rm oc}$ and $\mathcal{D}^{\rm sc}$,

$$Y_h=-i\omega\mathcal{D}^{\rm oc}/\mathcal{D}^{\rm sc}; \tag{45}$$

equating these determinants to zero gives the dispersion equations for open-circuited ($\mathcal{D}^{\rm oc}=0$) and short-circuited ($\mathcal{D}^{\rm sc}=0$) gratings.[13] Indeed, expressing $C_\sigma^{(0)}$ in terms of V via the set of equations that (43) implies, substituting it into I, taking the ratio I/V, we see that the numerator and the denominator are none other than $\mathcal{D}^{\rm oc}$ and $\mathcal{D}^{\rm sc}$, respectively.

Notice that if only inhomogeneous modes are involved ($\pi/l>\omega/v_L$), i.e., the SAW problem is being considered, the functions on the right-hand side of (43) are real and, hence, both $\mathcal{D}^{\rm oc}$ and $\mathcal{D}^{\rm sc}$ are real.

VII. FREQUENCY COALESCENCE

The numerically calculated frequency dependence of Im$[Y_h]$ for SAW generally appears to be of the form as shown in Fig. 3(a). However, if the substrate assumes one of orientations (24), then one obtains a curve of the type depicted in Fig. 3(b). The representative frequency dependencies of Im$[Y_h]$ and Re$[Y_h]$ for LSAW are shown in Fig. 5. The analogous specific feature is observed. The right-most anomaly existing in Fig. 5(a) disappears in Fig. 5(b) (the curve Im$[Y_h]$ is continuous and Re$[Y_h]$ does not vanish identically and exhibits peaks since LSAW attenuates due to radiation).

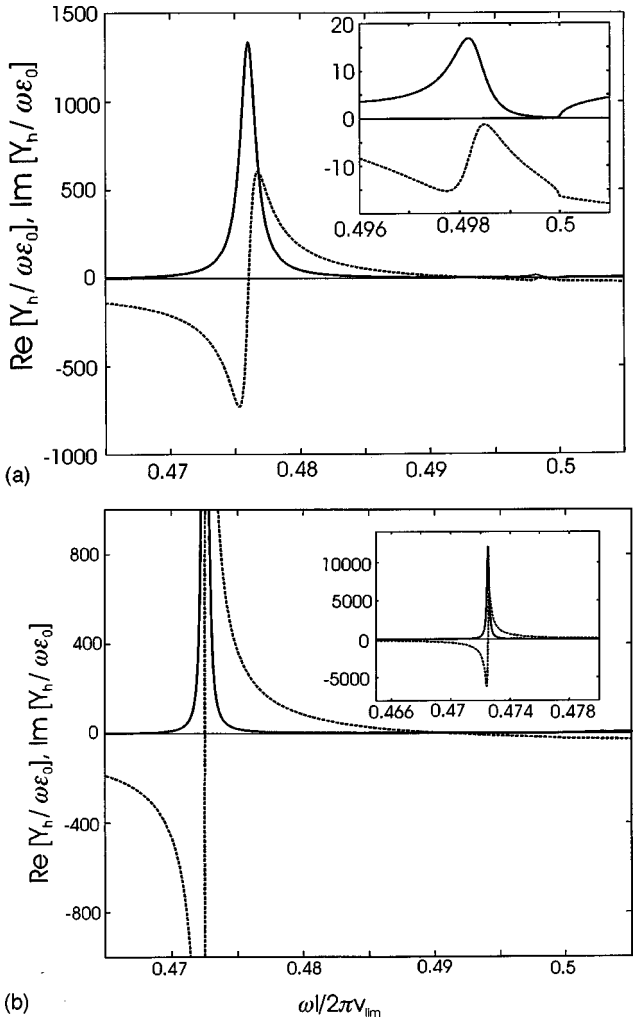


FIG. 5. Frequency dependence of Im$[Y_h/\omega\varepsilon_0]$ (dashed line) and Re$[Y_h/\omega\varepsilon_0]$ (solid line) for LSAW on $(-7.5,-54,0)$ (a) and $(0,-54,0)$ (b) $LiTaO_3$ cuts; ω, frequency; ε_0, vacuum permittivity; 1, period of the grating; $v_{\rm lim}$, limiting velocity of the fast transverse bulk wave [$v_{\rm lim}=4162.47$ m/s in (a) and $v_{\rm lim}=4226.5$ m/s in (b)].

The annihilation of singularities is due to $\mathcal{D}^{\rm oc}=0$ and $\mathcal{D}^{\rm sc}=0$ having a common root; this root is real for SAW and complex for LSAW. We want to clarify why such a degeneracy occurs at the orientations of families (24) (or at least must occur with increasing accuracy of computations). Our discussion is equally valid for SAW and LSAW.

First of all, we need a more expedient expression for the acoustic admittance than as given by (11). Invoking the definition of $\hat{\mathbf{B}}$ and $\hat{\mathbf{S}}$ (23), using (14), one can check that

$$\hat{\mathbf{B}}\mathbf{U}_\alpha+\hat{\mathbf{S}}^t\mathbf{V}_\alpha=i\mathbf{V}_\alpha,\quad \alpha=1,\ldots,4, \tag{46}$$

where $\hat{\mathbf{S}}^t$ is the transpose of $\hat{\mathbf{S}}$. From (46) the desired relation follows:

$$\hat{\mathbf{Y}}=\hat{\mathbf{B}}^{-1}(\hat{\mathbf{I}}+i\hat{\mathbf{S}}^{t}). \tag{47}$$

The matrix $\hat{\mathbf{B}}^{-1}$ is symmetric, because $\hat{\mathbf{B}}$ is symmetric. Making use of (15) permits to check that $\hat{\mathbf{S}}^{t}\hat{\mathbf{B}}$ and, hence, $\hat{\mathbf{B}}^{-1}\hat{\mathbf{S}}^{t}$ are antisymmetric matrices.

Let the notation $\hat{\mathbf{X}}\mapsto\hat{\mathbf{Z}}$ indicate that the positions of zero and nonzero elements in a matrix $\hat{\mathbf{Z}}$ are the same as in $\hat{\mathbf{X}}$, i.e., $Z_{ij}=0$ if $X_{ij}=0$ and $Z_{ij}\neq 0$ if $X_{ij}\neq 0$.

It can be checked that at orientations (24)

$$\hat{\mathbf{B}}\mapsto\hat{\mathbf{B}}^{-1},\quad \hat{\mathbf{S}}\mapsto\hat{\mathbf{B}}^{-1}\hat{\mathbf{S}}^{t} \tag{48}$$

in the corresponding cases (26)–(29).

Inserting $\hat{\mathbf{Y}}$ (47) into (31), taking into account (48), we discover by direct calculations that $\hat{\mathbf{B}}\mapsto\hat{\mathbf{G}}^{(s)(e)}$ and $\hat{\mathbf{S}}\mapsto\hat{\mathbf{G}}^{(a)(o)}$ [see Eq. (35)].

Using the fact that $\hat{\mathbf{G}}^{(s)(e)}\mapsto\hat{\mathbf{G}}_{e}^{(s)(2n)}$ and $\hat{\mathbf{G}}^{(s)(e)}\mapsto\hat{\mathbf{G}}_{o}^{(s)(2n+1)}$ while $\hat{\mathbf{G}}^{(a)(o)}\mapsto\hat{G}_{o}^{(a)(2n)}$ and $\hat{\mathbf{G}}^{(a)(o)}\mapsto\hat{\mathbf{G}}_{e}^{(a)(2n+1)}$ [see Eq. (36)] we eventually conclude that

$$\begin{aligned}&\hat{\mathbf{B}}\mapsto\hat{\mathbf{G}}_{e}^{(s)(2n)},\quad \hat{\mathbf{B}}\mapsto\hat{\mathbf{G}}_{o}^{(s)(2n+1)},\\ &\hat{\mathbf{S}}\mapsto\hat{\mathbf{G}}_{o}^{(a)(2n)},\quad \hat{\mathbf{S}}\mapsto\hat{\mathbf{G}}_{e}^{(a)(2n+1)}.\end{aligned} \tag{49}$$

From the examination of (43) with due regard to (44) and (49) we infer that (43) decomposes into two subsystems at orientations (24). The two subsystems possess two independent solutions (i) and (ii). Writing explicitly these subsystems reveals that solutions (i) and (ii) are arranged as follows:

$$\begin{aligned}&\text{case 1 (i)}\quad C_{x}^{(2n)},\ C_{y}^{(2n+1)},\ C_{z}^{(2n+1)},\ C_{\sigma}^{(2n+1)};\\ &\qquad\text{(ii)}\quad C_{x}^{(2n+1)},\ C_{y}^{(2n)},\ C_{z}^{(2n)},\ C_{\sigma}^{(2n)};\end{aligned} \tag{50}$$

$$\begin{aligned}&\text{case 2 (i)}\quad C_{x}^{(2n+1)},\ C_{y}^{(2n)},\ C_{z}^{(2n)},\ C_{\sigma}^{(2n+1)};\\ &\qquad\text{(ii)}\quad C_{x}^{(2n)},\ C_{y}^{(2n+1)},\ C_{z}^{(2n+1)},\ C_{\sigma}^{(2n)};\end{aligned} \tag{51}$$

$$\begin{aligned}&\text{case 3 (i)}\quad C_{x}^{(2n+1)},\ C_{y}^{(2n+1)},\ C_{z}^{(2n)},\ C_{\sigma}^{(2n+1)};\\ &\qquad\text{(ii)}\quad C_{x}^{(2n)},\ C_{y}^{(2n)},\ C_{z}^{(2n+1)},\ C_{\sigma}^{(2n)};\end{aligned} \tag{52}$$

$$\begin{aligned}&\text{case 4 (i)}\quad C_{x}^{(2n)},\ C_{y}^{(2n)},\ C_{z}^{(2n+1)},\ C_{\sigma}^{(2n+1)};\\ &\qquad\text{(ii)}\quad C_{x}^{(2n+1)},\ C_{y}^{(2n+1)},\ C_{z}^{(2n)},\ C_{\sigma}^{(2n)}.\end{aligned} \tag{53}$$

The coefficients $C_{x,y,z,\sigma}^{(n)}$ in (50)–(53) indicate that the corresponding components of force and the charge are constructed from the polynoms of degree n; e.g., $\sigma(x)=\Sigma_{n}\,C_{\sigma}^{(2n+1)}P^{(2n+1)}(x)$ in solution (i).

Since solution (i) involves only odd charge harmonics, no current flows in the electrodes in this case. In addition, (i) fulfills the subsystem not containing V and therefore is associated with the vanishing potential on the electrode. Hence, the frequencies of this solution cannot depend on the electrical state of the grating. This suggest that (i) is a common solution for open-circuited and short-circuited gratings.

The dispersion function describing solution (i) naturally becomes a common multiplier in $\mathcal{D}^{\mathrm{oc}}$ and $\mathcal{D}^{\mathrm{sc}}$. It cancels out, rendering the admittance Y_{h} equal to the ratio of the dispersion functions that determine the fundamental frequencies of solution (ii) under open-circuited and short-circuited gratings.

Note that the above proof applies irrespective of whether the wave velocity lies ''far'' or ''close'' to the limiting velocity of bulk waves, in contrast to our consideration based on the COM equations (2).

When the electrode thickness is totally ignored, $\mathbf{f}(x)\equiv 0$ and Eq. (33) simplifies to $\Phi(x)=\int_{-a}^{a}[\hat{\mathbf{G}}(x-x')]_{44}\sigma(x')dx'$. Hence, only two equations remain in Eq. (43):

$$V=\sum_{n}\,[\hat{\mathbf{G}}_{e}^{(s)(2n)}(x)]_{44}C_{\sigma}^{(2n)},$$

$$0=\sum_{n}\,[\hat{\mathbf{G}}_{o}^{(s)(2n+1)}(x)]_{44}C_{\sigma}^{(2n+1)},$$

since $[\hat{\mathbf{G}}^{(a)}(x)]_{44}\equiv 0$. The second equation determines the fundamental frequency which is independent of the electrical state of the grating. The COM equations yield the same result. Letting $H=0$ in Eq. (3), one obtains $r^{\pm}=R_{e}(\alpha)$, where $R_{e}(\alpha)$ describes the reflection because of the electrode short-circuiting electrical fields. In other words, as has been pointed out as early as in Refs. 39 and 40, the degeneracy occurs for any substrate if the mechanical loading is neglected.

VIII. CONCLUSION

We have indicated four families of substrate orientations at which two of the four fundamental frequencies characterizing simple short- and open-circuited gratings merge. The frequencies fall into equality not only to the lowest approximation with respect to the electrode thickness but also when they are determined through numerical computations that are believed to completely account for the ''mechanical'' effect of the electrode of finite thickness. The degeneracy occurs irrespective of the material constants of the substrate and the electrode and of the electrode shape. It is required that the electrodes should be of the same material and of identical symmetric shape. The electrode can be composed from horizontal slices of different materials.

The analysis developed in the paper allows one to monitor how specifically the crystallographic symmetry leads to the frequency coalescence. The key feature is that due to symmetry the decay factors of partial plane modes come about in pairs $\pm p_{\alpha}$.

In principle, it is not impossible that the degeneracy can occur at some other orientations. At present we cannot make any clear statements in this respect and leave the problem opened for further investigation. Note only that if the degeneracy does come about at orientations other than (24), then it seems reasonable to expect these orientations to be dependent on the thickness, shape, and material of electrodes as well as on the material constants of the substrate.

For the sake of simplicity the structure of the matrix on the left-hand side of Eq. (43) that accounts for the electrode elasticity has been derived on the basis of the so-called fast variation method (FVM).[37] The use of the finite element

method (FEM) to calculate elastic fields in the electrode does not change anything. Indeed, the structure of the matrices $\hat{\mathbf{E}}$ and $\hat{\mathbf{O}}$ (the position of zero and nonzero elements) does not depend on how they have been found. It is stipulated by isotropy of elastic properties of the electrode and its symmetric shape that lead to specific interrelations amid mechanical displacements of different directions and evenness. (FVM can be viewed as a simplified FEM—the whole electrode cross section is taken as a mesh; the method shortens time for programming work and computing and provides good accuracy, as has been found out using analytic formulas for the case of shear horizontal displacements.[37])

Our considerations are applicable if different sets of polynoms are used to interpolate $\mathbf{f}(x)$ and $\sigma(x)$ or, instead of polynoms, the Fourier series is used [FEM/SDA (spectral domain analysis) method[5,41]].

Note that the specificity of orientations (24) shows up in various respects. In particular, they secure the existence of nonattenuating surface waves that propagate on the substrate coated with layers faster than both transverse bulk waves in the substrate.[28,42–44] The same orientations support high-velocity waveguide modes propagating without leakage with the phase velocity exceeding that of the transverse bulk waves in the medium embedding the waveguide.[45]

The analogous degeneracy problem can be studied as applied to layered substrates. The corresponding procedure will also rely on the analysis of how the crystallographic symmetry affects the components of Green's function constructed from the admittance $\hat{\mathbf{Y}}$ of the layered structure. In principle, for layered substrates, $\hat{\mathbf{Y}}$ can be found from a matrix nonlinear equation derived in Ref. 18 (Chap. 3). On the other hand, for an homogeneous substrate coated with $n_{\max}$ piezoelectric layers (the perfect rigid contact on each interface is realized), $\hat{\mathbf{Y}}$ can also be expressed in terms of the vectors $\boldsymbol{\xi}_{h\alpha}=(\Pi_{n=1}^{n_{\max}}\hat{\mathbf{P}}_n)\boldsymbol{\xi}_{\alpha}$ and $\boldsymbol{\eta}_{h\alpha}=(\Pi_{n=1}^{n=n_{\max}}\hat{\mathbf{P}}_n^{-1})\boldsymbol{\xi}_{\alpha}$, where $\boldsymbol{\xi}_{\alpha}$ are $\boldsymbol{\xi}$ vectors of the substrate and $\hat{\mathbf{P}}_n=\exp[-ikh_n\hat{\mathbf{N}}_n]$ is the propagator of the nth layer of thickness h_n; the matrix $\hat{\mathbf{N}}_n$ (7) pertains to this layer. One then obtains an expression of the type (10) and, using the relation $\boldsymbol{\eta}_{h\alpha}\cdot\hat{\mathbf{T}}\boldsymbol{\xi}_{h\beta}=\delta_{\alpha\beta}$, an expression of the type (47), where $\hat{\mathbf{B}}$ and $\hat{\mathbf{S}}$ are found in terms of the components of $\boldsymbol{\xi}_{h\alpha}$ and $\boldsymbol{\eta}_{h\alpha}$ in a manner similar to (23). The thorough study of the degeneracy problem for layered substrates will be given elsewhere.

ACKNOWLEDGMENTS

A. Darinskii is grateful to Deutsche Forschungsgemeinschaft (Grant No. 436RUS 113/645/0) for financial support. A.D. also thanks the Russian Foundation for Basic Research for partial financial support. M. Weihnacht is grateful to Dr. R. Wobst for his programming work.

[1] C. Elachi, "Waves in active and passive periodic structures: a review," Proceedings of the IEEE, p. 297, 1976.
[2] L. M. Brekhovskikh and O. A. Godin, *Acoustics of Layered Media* (Springer-Verlag, New York, 1992).
[3] A. H. Nayfeh, *Wave Propagation in Layered Media* (North-Holland, Amsterdam, 1995).
[4] C. Campbell, *Surface Acoustic Wave Devices for Mobile and Wireless Communications* (Academic, San Diego, 1998).
[5] K. Hashimoto, *Surface Acoustic Wave Devices in Telecommunications* (Springer-Verlag, Berlin, 2000).
[6] M. S. Sandler and B. V. Sveshnikov, "An investigation of interdigital transducers of surface acoustic waves, taking into account the finite mass of the electrodes," Sov. Phys.-Radio Engng. Electron. **26**, 9–17 (1981).
[7] C. S. Hartmann, P. V. Wright, R. J. Kansy, and E. M. Gaber, "An analysis of SAW interdigital transducers with internal reflections and the application to the design of single-phase unidirectional transducers," Proceedings. 1982 IEEE Ultrasonics Symposium, pp. 40–45.
[8] D. P. Chen and H. A. Haus, "Analysis of metal strip SAW gratings and transducers," IEEE Trans. Sonics Ultrason. **SU-32**, 395–408 (1985).
[9] Z. H. Chen, M. Takeuchi, and K. Yamanouchi, "Analysis of the film thickness dependence of a single-phase unidirectional transducer using the coupling-of-mode theory and the finite element method," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 82–94 (1992).
[10] E. Danicki, "Generation and Bragg reflection of surface acoustic waves in nearly periodic system of elastic metal strips on piezoelectric half-space," J. Acoust. Soc. Am. **93**, 116–131 (1993).
[11] S. V. Biryukov, G. Martin, V. G. Polevoy, and M. Weihnacht, "Derivation of COM equations using the surface impedance method," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 602–611 (1995).
[12] Y. Zhang, J. Debois, and L. Boyer, "Characteristic parameters of surface acoustic waves in a periodic metal grating on a piezoelectric substrate," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, 183–193 (1993).
[13] V. P. Plessky, S. V. Biryukov, and J. Koskela, "Harmonic admittance and dispersion equation—the theorem," IEEE Trans. Ultrason., Ferroelect., Freq. Contr. **49**(4), 528–534 (2002).
[14] K. Hashimoto and M. Yamaguchi, "Derivation of coupling-of-modes parameters for SAW device analysis by means of boundary method analysis," Proceedings 1991 IEEE Ultrasonics Symposium, pp. 21–26.
[15] G. Martin, M. Weihnacht, and K. Franke, "Properties of interdigital transducers in relation to the substrate crystal symmetry," Proceedings 1995 IEEE Ultrasonics Symposium, pp. 291–295.
[16] S. V. Biryukov, "Equation for the surface impedance of an inhomogeneous elastic medium," Sov. Phys. Acoust. **31**, 174–178 (1985).
[17] S. V. Biryukov and L. L. Goryshnik, "Theory of interaction between surface waves in piezoelectric media and electrode structures," Sov. Phys. Tech. Phys. **25**, 960–964 (1980).
[18] S. V. Biryukov, Yu. V. Gulyaev, V. V. Krylov, and V. P. Plessky, *Surface Acoustic Waves in Inhomogeneous Media* (Springer-Verlag, Berlin, 1995).
[19] E. Danicki, "Spectral theory for IDTs," 1994 Proceedings Ultrasonics Symposium, pp. 213–222.
[20] E. Danicki, "Perturbation theory of surface acoustic wave reflection form a periodic structure with arbitrary angle of incidence," Arch. Mech. **36**, 623–638 (1984).
[21] V. Plessky, "Two parameter coupling-of-modes model for shear horizontal type SAW propagation in periodic gratings," Proceedings 1993 IEEE Ultrasonics Symposium, pp. 195–200.
[22] B. P. Abbott and K. Hashimoto, "A coupling-of-modes formalism for surface transverse wave devices," Proceedings 1995 IEEE Ultrasonics Symposium, pp. 239–245.
[23] E. Danicki, "Propagation of transverse surface acoustic waves in rotated Y-cut quartz substrates under heavy periodic metal electrodes," IEEE Trans. Sonics Ultrason. **SU-30**, 303–313 (1983).
[24] J. Lothe and D. M. Barnett, "Integral formalism for surface waves in piezoelectric crystals. Existence considerations," J. Appl. Phys. **47**, 1799–1807 (1976).
[25] J. Lothe and D. M. Barnett, "Further development of the theory for surface waves in piezoelectric crystals," Phys. Norv. **8**, 239–254 (1976).
[26] G. A. Maugin, *Continuum Mechanics of Electromagnetic Solids* (North-Holland, Amsterdam, 1988).
[27] D. Royer and E. Dieulesaint, *Elastic Waves in Solids I, II* (Springer-Verlag, Berlin, 1999).
[28] A. N. Darinskii, "Symmetry aspects of the existence of high-velocity SAW in layered composites," Phys. Lett. A **266**, 183–186 (2000).
[29] P. Ventura, J. Hode, J. Debois, and M. Solal, "Combined FEM and Green's function analysis of periodic SAW structure, application to the calculation of reflection and scattering parameters," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **48**, 1259–1274 (2001).
[30] S. V. Biryukov and M. Weihnacht, "Real-space field of surface sources and the problem of fast leaky wave generation in a piezoelectric half-space," J. Appl. Phys. **83**, 3276–3287 (1998).

[31] A. R. Baghai-Wadji, H. Reichinger, H. Zidek, and Ch. Mecklenbrauker, ''Green's function applications in SAW devices,'' Proceedings 1991 IEEE Ultrasonics Symposium, pp. 11–20.
[32] R. C. Peach, ''Green function analysis for SAW devices with arbitrary electrode structure,'' Proceedings 1997 IEEE Ultrasonics Symposium, pp. 99–103.
[33] J. M. Hode and J. Desbois, ''Original basic properties of the Green's functions of a semi-infinite piezoelectric substrate,'' Proceedings 1999 IEEE Ultrasonics Symposium, pp. 131–136.
[34] P. Ventura, J. Hode, and M. Solal, ''A new efficient combined FEM and periodic Green's function formalism for the analysis of periodic SAW structures,'' Proceedings 1995 IEEE Ultrasonics Symposium, pp. 263–268.
[35] V. P. Plessky and T. Thorvaldsson, ''Periodic Green's function analysis of SAW and leaky SAW propagation in a periodic system of electrodes,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 280–293 (1995).
[36] A. R. Baghai-Wadji and V. P. Plessky, ''Propagation of Rayleigh type SAW in periodic structures with resonating elements,'' Sov. Phys. Acoust. **38**, 806–810 (1992).
[37] S. V. Biryukov, ''Fast variation method for elastic strip calculation,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **49**, 635–642 (2002).
[38] E. P. EerNisse and R. Holland, ''On variational techniques for piezoelectric device analysis,'' Proc. IEEE **55**, 1524–1525 (1967).
[39] K. Blotekjaer, K. Ingebrigtsen, and H. Skeie, ''A method for analyzing waves in structures consisting of metal strips on dispersive media,'' IEEE Trans. Educ. **20**, 1133–1138 (1973).
[40] K. Blotekjaer, K. Ingebrigtsen, and H. Skeie, ''Acoustic surface waves in piezoelectric materials with periodic metal strip on the surface,'' IEEE Trans. Educ. **20**, 1139–1146 (1973).
[41] K.-Y. Hashimoto, T. Omori, and M. Yamaguchi, ''Analysis of SAW excitation under periodic metallic grating structures,'' in *Advances in Surface Acoustic Waves Technology, Systems and Applications*, edited by C. C. W. Ruppel and T. A. Fjeldly (World Scientific, Singapore, 2000), Vol. 1, pp. 133–182.
[42] I. S. Didenko, F. S. Hickernell, and N. F. Naumenko, ''The experimental and theoretical characterization of the SAW propagation properties for zinc oxide on silicon carbide,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 179–187 (2000).
[43] N. F. Naumenko and I. S. Didenko, ''High-velocity surface acoustic waves in diamond and sapphire with zinc oxide film,'' Appl. Phys. Lett. **75**, 3029–3031 (1999).
[44] A. N. Darinskii, I. S. Didenko, and N. F. Naumenko, ''Fast quasilongitudinal sagittally polarized surface waves in layer-substrate structures,'' J. Acoust. Soc. Am. **107**, 2351–2359 (2000).
[45] A. N. Darinskii and M. Weihnacht, ''High-velocity waveguide acoustic modes in solids,'' J. Appl. Phys. **88**, 471–477 (2000).

Modal analysis of a drumlike silencer

Lixi Huang[a)]
Department of Mechanical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong

(Received 23 January 2002; revised 15 July 2002; accepted 15 July 2002)

Low-frequency duct noise is difficult to deal with by passive methods such as porous duct lining. Reactive methods like expansion chamber are rather bulky, while compact resonators are too narrow banded. This study shows that a suitably stretched thin membrane backed by a slender cavity can achieve a satisfactory performance from low to medium frequencies over an octave band. The present paper focuses on the details of the modal behavior of the fully coupled membrane-cavity system, and examples are given with parameters set in a practical range. Typically, the membrane has a structure to air mass ratio of unity, and is stretched towards the elastic stress limit for a material like aluminum. The backing cavity has a depth equal to the duct height and a length five times the duct height. Three resonant peaks are found in the low to medium frequency range while the transmission loss between adjacent peaks remain above 10 dB. For the first peak, almost complete sound reflection occurs as a result of an out-of-phase combination of the first and second *in vacuo* modes of simply supported membranes. The second peak is solely contributed to by the first mode, while the third peak features mainly the second mode vibration. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1508778]

PACS numbers: 43.50.Gf, 43.20.Tb, 43.20.Ks [MRS]

I. INTRODUCTION

The objective of this study is to explore a mechanism of broadband, passive control of duct noise without using porous duct lining. Porous media have been the backbone of almost all dissipative noise abatement techniques. It is a very mature and reliable technique which works for a very broad frequency band. But there are two concerns, one acoustical and another environmental. Acoustically, existing sound absorption techniques are ineffective in the low frequency range, such as that below 200 Hz. This range is also often over-looked due to its low A-weighting. Nevertheless, the actual power of noise radiation from, say, a fan, is often highest in this frequency range. Environmentally, there has been increasing concern about the deposition and accumulation of dusts in the pores of the porous material. A periodical cleaning of the lining would be rather costly and indeed tedious. In fact, there are already public concerns of bacteria breeding in the centralized ventilation systems of ordinary commercial buildings. The use of porous material for noise or heat insulation purposes might have contributed to an indoor air quality which is often worse than outdoors. There are also places such as operation theatres, where high hygienic requirement forbids the use of such materials.

The need of controlling low frequency noise in an environment friendly manner calls for a fiberless approach. A team of acousticians led by Fuchs (2001a) have been very successful in achieving this goal, both technically and commercially. The building blocks for their devices include microperforated sheets for sound absorption, and impervious thin membranes for separating the harsh environment from acoustic elements like resonators (Fuchs, 2001b). Active control techniques are also implemented, albeit in its simplest and most practical fashion. Sharing exactly the same goal of fiberless noise control at low-frequencies, the present study sets out to tackle the specific problem of duct noise by yet another approach. Instead of using a thin membrane (such as aluminum of around 0.1 mm thickness) in its natural state, the use of such membranes under high tension is explored. It is found in this study that a grazing incident noise induces the tensioned membrane to vibrate in a low-order axial mode, and such vibration reflects low-frequency noise very effectively. The acoustic elements in the present theoretical model are entirely reactive without any sound absorption material, nor microperforation, although some mechanism of energy dissipation always exists in all experimental rigs. As it turns out, the level of tensile stress required to produce a broadband performance approaches the elastic limit of common materials like aluminum. For this reason, the device is tentatively called a drum silencer. Unlike normal vehicle exhaust silencer or splitter silencers commonly used in the ventilation systems, the drum silencer reflects noise without causing any extra pressure loss.

Focusing on the fundamental mechanism of air-membrane interaction, the present article is devoted to the analysis of membrane response expressed in terms of its *in vacuo* modes. The remaining paragraphs of the introduction are devoted to the theoretical modelling in studies related to the interaction of sound and flexible walls. Knowledge of such interaction has been steadily built up during the past three decades or so. Research in this area has been motivated by a diverse range of engineering problems which include, but are not limited to, aerospace, underwater and room acoustics applications. For example, Lyon (1963), Pretlove (1965), and Guy (1979) studied the effect of a plane flexible wall on the transmission of sound into rooms or cavities. Pan and Bies (1990), and Sum and Pan (1998) investigated the effect of flexible walls on the reverberation performance of rooms. Cheng (1994) investigated the coupling of sound with

a)Electronic mail: mmlhuang@polyu.edu.hk

cylindrical shells simulating aircraft cabins, the purpose being to identify the noisiest structural modes in the cabin. Dowell and Voss (1963), and Dowell *et al.* (1977) studied the stability of a cavity-backed panel exposed to external flows. In all these studies, the pressure on the internal panel surface, i.e., the side facing the cavity, is expanded in terms of the normal cavity modes, and the excitation pressure on the external panel surface is prescribed. In most cases, the cavity air modifies the panel vibration and the mode is said to be panel controlled. On the other hand, when the cavity is very shallow and the excitation frequency is low, the cavity becomes stiffer than the panel, the latter then acts merely as an added mass to the room acoustics mode. This scenario is often categorized as strong fluid-structure interaction when one chooses to regard such modified room modes as an eigenmode of the panel, see, for example, Pretlove (1965). The radiation pressure on the external panel surface is either ignored or lumped into a prescribed excitation force. The validity of such simplification, and the dominance of either fluid or structure in modified modes of resonance, are in fact characteristic of weak coupling in the context of the present studies. The structure used in the current model is a thin membrane with negligible bending stiffness. It is found here that the panel-controlled or cavity-controlled resonance tends to be ineffective as far as the reflection of grazing incident wave is concerned. Another departure of this study from the previous literature is that the radiation pressure on the membrane surface external to the cavity is fully coupled with the membrane motion. This fluid loading is found to be equally important as the acoustic response inside the cavity.

In the field of architectural acoustics, panels are also widely used in front of walls and under the ceilings to enhance the low-frequency sound absorption via the excitation of cavity or panel resonance. Often, these panels also serve optical purposes. To achieve sufficient sound absorption, porous material may be used behind the panel. Perforated panels and multilayer designs may also be employed. Analytical efforts have been made to predict the acoustic performance of such panels. For example, Kang and Fuchs (1999) successfully treated the problem of microperforated membrane as a parallel connection of the (impervious) membrane and apertures. This is a locally reactive model which nevertheless reveals most of the essential physics. Frommhold *et al.* (1994) described the acoustic performance of a splitter silencer made by a host of combined Helmholtz and plate resonators. The analysis is again carried out on the basis of locally reactive model based on the normal incidence impedance of a single resonator. Using the Helmholtz integral formulation, Horoshenkov and Sakagami (2001) studied sound reflection and absorption by a finite, poroelastic plate in an infinite rigid baffle. The plate is supported by a shallow cavity which is filled with sound absorption materials. The cavity is infinite in the direction parallel to the panel. The analytical method employed in this study is rigorous, but the configuration differs from that of the present study where the cavity is enclosed. A recent study conducted by Mechel (2001) dealt with both locally reactive and bulk reactive poroelastic panels supported by an enclosed cavity.

Compared with all the related studies mentioned above, the theoretical model and the solution method described in the present paper have the following distinct features: (a) the membrane and the cavity are finite but not necessarily compact; (b) there is no simplification assumption made regarding the radiation pressure on either side of the membrane; (c) the pressure on the panel surface facing the cavity is expressed in terms of the duct acoustics modes instead of the room acoustics modes. Multiple resonant peaks are found and, when an appropriate combination of cavity and membrane properties are chosen, the breakdown of transmission loss between these peaks is moderate, and a level of 10 dB is sustained for a continuous frequency range wider than an octave band. This represents an attractive performance considering the fact that an expansion chamber of the same cavity size yields a maximum transmission loss of 5 dB when it experiences a quarter wavelength resonance. In what follows, Sec. II outlines the analytical formulation for the membrane response to a grazing incident wave. Section III gives the spectra of transmission loss as well as details of all modal interactions for a typical example which can be easily materialized. Section IV focuses on the wave-reflecting behavior of the coupled system at the low frequency region, where an approximate and closed-form solution is possible. Qualitative analysis is extended to moderate frequencies where the transmission loss attains peak values. The main physical ingredients are summarized in the conclusion section.

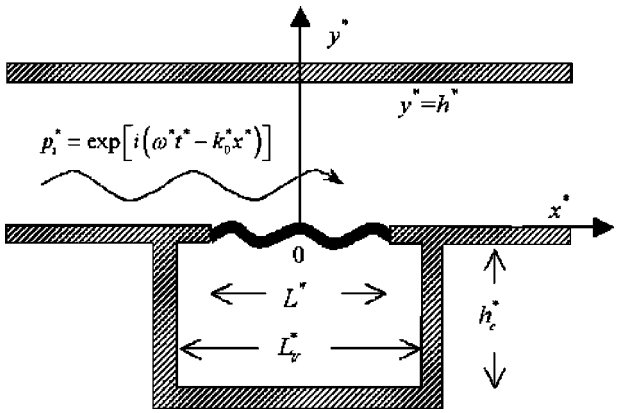


FIG. 1. Configuration of the theoretical model.

II. METHOD OF SOLUTION

The configuration is shown in Fig. 1. It has a two-dimensional duct, or channel, of height h^*, lined in part by a membrane of length L^* on the lower wall. The asterisks denote dimensional variables while the corresponding dimensionless ones are introduced shortly without asterisks. The membrane is simply supported at the two edges, $|x^*| = L^*/2$, and is enclosed by a rigid-walled cavity of depth h_c^* and length L_v^*. L_v^* is held equal to L^* in the numerical example to be given later, but a general formulation for $L_v^* > L^*$ is sought for the sake of parametric studies on the shape of the cavity. The question now is how the membrane responds to an incident wave of unit amplitude, $p^{*\prime} = \exp[i(\omega^* t^* - k_0^* x^*)]$, where $\omega^* = 2\pi f^*$ and $k_0^* = \omega^*/c_0^*$ are, respectively, the angular frequency and the wave number based on the speed of sound in free space, c_0^*. The flexible

wall has a mass per unit length m^*, and a tensile force T^* is applied. The governing equation for the membrane vibration is

$$m^*\frac{\partial^2 \eta^*}{\partial (t^*)^2} - T^*\frac{\partial^2 \eta^*}{\partial (x^*)^2} + (p_+^* - p_-^*) = 0, \tag{1}$$

where η^* is the membrane displacement, and $(p_+^* - p_-^*)$ is the fluid loading on the membrane located at $y^* = 0$. Positive loading means a net force acting downwards, and positive modal impedance, which is introduced shortly, refers to positive loading induced by an upward membrane vibration velocity. Structural damping in a thin membrane like this is normally small and is excluded at present, so is bending stiffness. Dividing Eq. (1) by $\rho_0^*(c_0^*)^2$, where ρ_0^* is the fluid density, yields

$$\frac{m^*}{\rho_0^* h^*}\frac{\partial^2(\eta^*/h^*)}{\partial(t^* c_0^*/h^*)^2} - \frac{T^*}{h^*\rho_0^*(c_0^*)^2}\frac{\partial^2(\eta^*/h^*)}{\partial(x^*/h^*)^2} + \frac{p_+^* - p_-^*}{\rho_0^*(c_0^*)^2} = 0. \tag{2}$$

All variables are now nondimensionalized as follows by three basic quantities, ρ_0^*, h^*, and c_0^*:

$$x = \frac{x^*}{h^*}, \quad y = \frac{y^*}{h^*}, \quad t = \frac{c_0^* t^*}{h^*},$$

$$h_c = \frac{h_c^*}{h^*}, \quad f = \frac{f^* h^*}{c_0^*}, \quad \omega = \frac{\omega^* h^*}{c_0^*}, \tag{3}$$

$$k_0 = 2\pi f, \quad m = \frac{m^*}{\rho_0^* h^*}, \quad T = \frac{T^*}{h^*\rho_0^*(c_0^*)^2}, \quad p = \frac{p^*}{\rho_0^*(c_0^*)^2}.$$

m is the mass ratio and T is the dimensionless tensile force. In a real, three-dimensional laboratory model, T is calculated by dividing the total force applied on the membrane by $\rho_0^*(c_0^*)^2 A^*$, where A^* is the cross section of a rectangular duct. Note that $\omega = k_0$ according to the above normalization scheme. The first cut-on frequency of the duct is $f = 0.5$. Equation (2) becomes

$$m\frac{\partial^2 \eta}{\partial t^2} - T\frac{\partial^2 \eta}{\partial x^2} + (p_+ - p_-) = 0. \tag{4}$$

Note that, in reality, the two-dimensional model can only be realized properly when the duct has a finite width in the third direction, say $z \in [0, h_z]$, and the width h_z cannot be too wide as to cause higher order modes in that direction. Meanwhile, the membrane should be allowed to move freely at the edges of $z = 0+$, h_z- with a tiny gap with the duct walls. The solution for a simplified model excluding the cavity effect was given in Huang (1999). The same procedure is followed here. The main steps are described very briefly and issues unique to the current model are discussed in more details.

For harmonic vibrations, and introducing membrane vibration velocity $V = \partial\eta/\partial t = i\omega\eta$, Eq. (4) becomes

$$mi\omega V - (T/i\omega)\partial^2 V/\partial x^2 + (p_+ - p_-) = 0. \tag{5}$$

Equation (5) can be solved via the standard Galerkin procedure, in which V is expanded as a series of *in vacuo* modes with modal amplitude V_j,

$$V_j(t) = 2\int_0^1 V(x,t)\sin(j\pi\xi)\,d\xi, \tag{6}$$

where

$$\xi = x^*/L^* + 1/2, \tag{7}$$

is the local axial coordinate. Equation (5) then becomes

$$\mathcal{L}_j V_j + 2\int_0^1 (p_+ - p_-)\sin(j\pi\xi)\,d\xi = 0,$$
$$\mathcal{L}_j = \left[mi\omega + \frac{T}{i\omega}\left(\frac{j\pi}{L}\right)^2 \right], \quad j = 1,2,3,\ldots. \tag{8}$$

A. Radiation and cavity reflection impedances

Traditionally, the fluid loading inside the cavity, p_-, is found by cavity modes while the upper surface loading, p_+, is specified as part of the external excitation force. This approach is invalid for the current problem since the coupling of air and membrane motion is very strong. For this reason, it is more convenient to take the following approach. The fluid loading, $(p_+ - p_-)$, is divided into three parts. Part one is the upper surface pressure due to the specified incident wave without considering the membrane response,

$$p_i(x,t) = \exp[i(\omega t - k_0 x)]. \tag{9}$$

Part two is the sound radiated by the membrane vibration excluding the reflections by the vertical walls of the cavity located at $|x| = L_v/2$. The radiation pressure on the lower side, denoted as $p_{-\mathrm{rad}}$, is formulated in a way similar to that in the main duct, $p_{+\mathrm{rad}}$. In the special case where $h_c = 1$, it is shown that $p_{-\mathrm{rad}} = -p_{+\mathrm{rad}}$, and the pressure difference is simply twice the radiation pressure on the upper side. The third part is due to the reflection from the two vertical cavity walls, denoted as $p_{-\mathrm{ref}}$. Hence,

$$(p_+ - p_-) = p_i + (p_{+\mathrm{rad}} - p_{-\mathrm{rad}}) - p_{-\mathrm{ref}}. \tag{10}$$

The radiation pressure in the main duct is well known (Doak, 1973), and is written here in the dimensional form first. The pressure consists of contributions from all duct acoustics modes of index n,

$$p_{+\mathrm{rad}}^*(x^*,y^*,t^*) = \frac{\rho_0^*}{2h^*}\sum_{n=0}^{\infty} c_n^*\psi_n(y^*) \times \int_{-L^*/2}^{+L^*/2} \psi_n^*(y^{*\prime}) V^*(x^{*\prime},y^{*\prime},t^*) \times [H(x^*-x^{*\prime})e^{-ik_n^*(x^*-x^{*\prime})} + H(x^{*\prime}-x^*)e^{+ik_n^*(x^*-x^{*\prime})}]\,dx^{*\prime}, \tag{11}$$

where H is the Heaviside function, c_n^*, k_n^*, and ψ_n^* are, respectively, the modal phase speed, the modal wave number, and the modal velocity potential,

$$c_n^*=\frac{ic_0^*}{\sqrt{(n\pi c_0^*/\omega^* h^*)^2-1}},\quad k_n^*=\frac{\omega^*}{c_n^*},$$
$$\psi_n^*(y^*)=\sqrt{2-\delta_{0n}}\cos(n\pi y^*/h^*), \tag{12}$$

and δ_{0n} is the Kronecker delta. Equation (11) is valid for any time history and location of the source V^*. From this point onwards, the source is specified as a harmonic membrane vibration at $y^{*\prime}=0$, $x^{*\prime}\in[-L^*/2,L^*/2]$. All variables are represented by their amplitudes, such as $p^*_{+\mathrm{rad}}(x^*,y^*)$ and $V^*(x^{*\prime})$. There are two important length scales, one is h^* and another is L^*. The former concerns the duct acoustics modes while the latter concerns the membrane dynamics. The local dimensionless coordinate ξ introduced in Eq. (7) is used together with x, and ξ' with x'. The dimensionless radiation pressure can be established from Eq. (11),

$$p_{+\mathrm{rad}}(x,y)=\frac{L}{2}\sum_{n=0}^{\infty}c_n\psi_n(y)\int_0^1\psi_n(y')V(x')\times[H(x-x')e^{-ik_n(x-x')}+H(x'-x)e^{+ik_n(x-x')}]d\xi'. \tag{13}$$

The dimensional Eq. (11) allows easy translation of Eq. (13) into the lower duct where the relevant duct acoustics scale is h_c^* and the membrane scale remains as L^*. The radiation pressure inside the cavity for the general case of $h_c^*\neq h^*$ is

$$p_{-\mathrm{rad}}(x_c,x_c)=\frac{L_c}{2}\sum_{n=0}^{\infty}c_{nc}\psi_n(y_c)\int_0^1\psi_n(y_c')[-V(x_c')]\times[H(x_c-x_c')e^{-ik_{nc}(x_c-x_c')}+H(x_c'-x_c)e^{+ik_{nc}(x_c-x_c')}]d\xi', \tag{14}$$

where the quantities peculiar to the cavity are distinguished by an additional subscript c as defined below:

$$x_c=\frac{x^*}{h_c^*},\quad y_c=\frac{y^*}{h_c^*},\quad L_c=\frac{L^*}{h_c^*},\quad L_v=\frac{L_v^*}{h_c^*},$$
$$\omega_c=\frac{\omega^* h_c^*}{c_0^*}=\omega h_c,\quad c_{nc}=\frac{i}{\sqrt{(n\pi/\omega_c)^2-1}},\quad k_{nc}=\frac{\omega_c}{c_{nc}}. \tag{15}$$

The reflection in the cavity is calculated in a form similar to the radiation pressure (14) except that the Heaviside functions are replaced by unknown coefficients A and B. The reflection for sound radiated by an elemental source of length $d\xi'$ is

$$dp_{-\mathrm{ref}}=\frac{L_c}{2}\sum_{n=0}^{\infty}c_{nc}\psi_n(y)\psi_n(y')\times[Ae^{-ik_{nc}(x_c-x_c')}+Be^{+ik_{nc}(x_c-x_c')}][-V(x_c')]d\xi',$$

where A and B are determined by the rigid wall boundary conditions,

$$\left.\frac{\partial(p_{-\mathrm{rad}}+p_{-\mathrm{ref}})}{\partial x_c}\right|_{|x_c|=L_v/2}=0,$$

to be satisfied for each duct acoustics mode n. Hence,

$$(1+B)e^{ik_{nc}(-L_v/2-x_c')}-Ae^{-ik_{nc}(-L_v/2-x_c')}=0,$$
$$-(1+A)e^{-ik_{nc}(L_v/2-x_c')}+Be^{ik_{nc}(L_v/2-x_c')}=0.$$

Solving these two equations yields

$$A=\frac{e^{ik_{nc}(L_v-2x_c')}+1}{e^{ik_{nc}(2L_v)}-1},\quad B=\frac{e^{ik_n(L_v+2x_c')}+1}{e^{ik_n(2L_v)}-1},$$

hence the reflection pressure

$$p_{-\mathrm{ref}}(x_c,y_c)=\frac{L_c}{2}\sum_{n=0}^{\infty}c_{nc}\psi_n(y_c)\int_0^1\psi_n(y_c')[-V(x_c')]\times\frac{2}{e^{ik_{nc}(2L_v)}-1}[\cos k_{nc}(x_c-x_c')+e^{ik_{nc}L_v}\cos k_{nc}(x_c+x_c')]d\xi'. \tag{16}$$

In order to solve the membrane dynamics, Eq. (8), the fluid loading (p_+-p_-) has to be related to the modal vibration velocity amplitudes V_j by the following modal decomposition:

$$(p_+-p_-)=\sum_{j=1}^{\infty}V_j p_j^1(x),$$

where $p_j^1(x)$ is the fluid loading caused by the modal vibration of unit amplitude, $V=\sin(j\pi\xi)$, which is to be substituted into Eqs. (13) and (16). The resulting loading contains not only the source mode of index j, but also of all other modes. Modal impedance is now defined as

$$Z_{jl}=\int_0^1 2\sin(l\pi\xi)p_j^1(x,0)d\xi,$$

where subscript j refers to the source vibration mode while l the resulting pressure coefficient. Part of Z_{jl} derives from $p_{+\mathrm{rad}}$ which was given in Huang (1999), and is now given the new symbol of Z_{+jl}. The contribution from the lower side of the membrane is denoted as Z_{-jl}, while that caused by the vertical wall reflections as Z_{rjl}. The total modal impedance Z_{jl} is calculated as follows:

$$Z_{jl}=Z_{+jl}+Z_{-jl}+Z_{rjl}=\int_0^1 2\sin(l\pi\xi)(p_{+\mathrm{rad}}-p_{-\mathrm{rad}}-p_{-\mathrm{ref}})_j^1\,d\xi. \tag{17}$$

Z_{+jl} given in Huang (1999) can be recovered here from Z_{-jl} by specifying $h_c=1$. The results for Z_{-jl} and Z_{rjl} are

$$Z_{-jl}=L_c\sum_{n=0}^{\infty}c_{nc}(2-\delta_{0n})\mathcal{I}_{2c}(n,j,l),$$
$$Z_{rjl}=L_c\sum_{n=0}^{\infty}c_{nc}(2-\delta_{0n})\mathcal{I}_{2r}(n,j,l), \tag{18}$$

where $\mathcal{I}_{2c}$ and $\mathcal{I}_{2r}$ are the results of double-integration over the membrane surface,

$$\mathcal{I}_{2c}(n,j,l)=\frac{l\pi j\pi(\cos j\pi-e^{-ik_{nc}L_c})(\cos j\pi+\cos l\pi)}{[(j\pi)^2-(k_{nc}L_c)^2][(l\pi)^2-(k_{nc}L_c)^2]}+\frac{ik_{nc}L_c\delta_{jl}}{(j\pi)^2-(k_{nc}L_c)^2}, \tag{19}$$

$$\mathcal{I}_{2r}(n,j,l)=\frac{2j\pi l\pi(1-\cos j\pi\cos k_{nc}L_c)(\cos j\pi-e^{ik_{nc}L_v})(\cos j\pi+\cos l\pi)}{[(j\pi)^2-(k_{nc}L_c)^2][(l\pi)^2-(k_{nc}L_c)^2][e^{-ik_{nc}(2L_v)}-1]}. \tag{20}$$

Two observations are made. First, the results for $\mathcal{I}_{2c}(n,j,l)$ and $\mathcal{I}_{2r}(n,j,l)$ vanish when j is even and l is odd, or when l is even and j is odd. This property decouples the even and odd modes of the membrane, and all subsequent analytical processing is focused on the case of $\cos j\pi=\cos l\pi$. The second observation is that, when $k_{nc}L_c$ coincides with $j\pi$ or $l\pi$, $\mathcal{I}_{2c}(n,j,l)$ becomes 0/0-type, and the finite results are similar to those for Z_{+jl} given in Huang (1999). The only possible mode which may cause this is the plane wave mode $n=0$, unless the cavity is very deep and the incident wave has higher order modes in the cavity. When $L_v=L$, and $k_{0c}L_v\to\ell\pi$, where ℓ is also an integer, the cavity reflection term $\mathcal{I}_{2r}$ may diverge,

$$\lim_{k_{nc}L_v\to\ell\pi}\mathcal{I}_{2r}(0,j,l)=\begin{cases}\infty, & \text{when } (-1)^{\ell}=-(-1)^{j}=-(-1)^{l};\\ -i/2, & \text{when } \ell=j=l;\\ 0, & \text{otherwise.}\end{cases} \tag{21}$$

This implies that, for a duct length equal to an odd multiple of half wavelength, it's impossible to excite the even modes on the membrane.

B. Membrane response and transmission loss

Having found the modal radiation and cavity reflection impedances, the dynamics Eq. (8) can be cast as a truncated set of linear equations for the modal vibration amplitudes, V_j, $j=1,2,3,...,N$,

$$\begin{bmatrix}Z_{11}+\mathcal{L}_1 & Z_{12} & \cdots & Z_{1N}\\ Z_{21} & Z_{22}+\mathcal{L}_2 & \cdots & Z_{2N}\\ \cdots & \cdots & \cdots & \cdots\\ Z_{N1} & Z_{N2} & \cdots & Z_{NN}+\mathcal{L}_N\end{bmatrix}\begin{bmatrix}V_1\\V_2\\V_3\\\vdots\\V_N\end{bmatrix}+\begin{bmatrix}I_1\\I_2\\I_3\\\vdots\\I_N\end{bmatrix}=0, \tag{22}$$

where I_j is the modal coefficient of the incident wave,

$$I_j=\int_0^1 p_i 2\sin(j\pi\xi)d\xi=2j\pi e^{ik_0L/2}\left[\frac{1-e^{i(-k_0L+j\pi)}}{(j\pi)^2-(k_0L)^2}\right]. \tag{23}$$

Note that $N=25$ is normally enough as further increase in N does not make any noticeable difference.

The final transmitted wave is found by adding the incident wave, p_i, to the far-field radiation wave, $p_{+\mathrm{rad}}$, which can be found from Eq. (13) taking only the plane wave mode $n=0$ for $x>L/2$:

$$p_t=p_{+\mathrm{rad}}|_{n=0,x\to+\infty}+p_i. \tag{24}$$

For an incident wave of unit amplitude, $|p_i|=1$, the transmission loss is calculated as

$$\mathrm{TL}=-20\log_{10}|p_t|. \tag{25}$$

Denoting the reflected sound as $p_re^{ik_0x}$, the coefficients of sound power reflected and absorbed, if there is structural damping, are calculated as follows:

$$\beta=|p_r|^2,\quad \alpha=1-|p_r|^2-|p_t|^2. \tag{26}$$

When two identical membrane-cavity systems are used, one at each side of a two-dimensional channel, the two membranes can be shown to respond in an identical manner. The pressure loading on the upper side of the lower membrane, $p_{+\mathrm{rad}}$, must account for that caused by the opposite membrane, which is the same as the reciprocal pressure on $y=1$ caused by the lower membrane. The latter can be calculated by setting $y=1$ in $\psi_n(y)$ in Eq. (13). The performance of a two-membrane configuration can also be calculated by regarding the channel height as $h^*/2$, but caution must be taken to avoid mistakes when converting the results back to the dimensionless form based on the full duct height h^*.

The complex amplitude of the reflected sound, p_r, is the sum of contributions made by all individual membrane vibration modes, which is found as follows with the help of Eq. (13):

$$p_r=\frac{p_{+\mathrm{rad}}|_{n=0,x\to-\infty}}{e^{ik_0x}}=\frac{1}{2}\int_{-L/2}^{+L/2}V(x')e^{ik_0x'}\,dx'=\sum_{j=1}^{\infty}V_jR_j, \tag{27}$$

where R_j is the complex amplitude of the reflected sound by the induced vibration of the jth mode with unit amplitude,

$$R_j=\frac{1}{2}\int_{-L/2}^{+L/2}\sin(j\pi\xi')e^{-ik_0x'}\,dx'=I_jL/4, \tag{28}$$

and the last identity is found by comparing the integration with that in Eq. (23). As each complex number can be represented by a vector in the real-imaginary space, the contribution of each single mode, denoted as $\mathbf{V}_j\mathbf{R}_j$, towards the total reflected sound, $\mathbf{p}_r$, can be found by projecting the former to the latter as illustrated in Fig. 2. The result is called the modal reflection contribution, γ_j, which is defined below together with the interference index, $\Gamma_{1,2}$, between the contributions from the two modes,

$$\gamma_j=\mathrm{Re}[V_jR_j\hat{p}_r/|p_r|],\quad \Gamma_{1,2}=\cos(\theta), \tag{29}$$

where $\hat{p}_r$ is the conjugate of p_r, and θ is shown in Fig. 2. Note that γ_j is a real scalar, and is a fraction of $|V_jR_j|$. Note also that the magnitudes of both V_j and R_j can be greater than unity, but that of the reflected wave, $p_r=\Sigma_jV_jR_j$, is

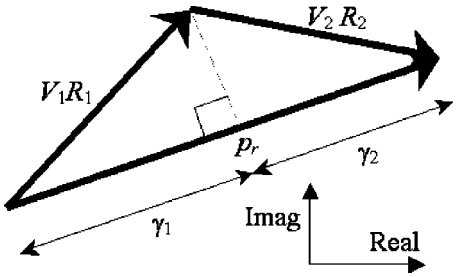


FIG. 2. Projection of reflection caused by each individual mode onto the sum vector $\mathbf{p}_r$.

less than unity since the reflected sound power derives from the incident wave which has a specified magnitude of unity.

III. SPECTRA OF MODAL INTERACTIONS

The results of the overall spectra are presented before the physics of modal response for specific frequencies are discussed. The following set of parameters are used as the default values for a device with two identical membranes, one on each side,

$$m=1, \quad T=0.475, \quad h_c=1, \quad L=L_v=5. \tag{30}$$

The mass ratio is considered to be in the practical range as illustrated by the example of 0.077 mm thick aluminum foil used in a duct of height 170 mm, $m=2700\times0.077/(1.225\times170)=1$. The dimensionless tensile force $T=0.475$ translates into a dimensional force of

$$F^*=T\rho_0^*(c_0^*h^*)^2=0.475\times1.225\times(340\times0.17)^2=1944N$$

for a square duct and the corresponding tensile stress is 148.5 MPa, which is close to but still within the yielding strength of the material. The choice of $h^*=170$ mm as the example is based on the convenience of the frequency unit: $c^*/h^*=2$ kHz, which falls right on the center of octave bands in practical use. The first cut-on frequency for a rigid duct of height 170 mm is thus 1 kHz, and the dimensionless lower stopband to be shown below, $f=0.054$, corresponds to $f^*=108$ Hz. The choice of the default set of parameters in Eq. (30) derives from a parametric study of which a brief summary is given in the conclusions section.

A. Stopband and spectral peaks

The main objective of the study is to see whether a tensioned membrane can achieve a broad stopband, preferably in the low-frequency range. The stopband here is defined as the frequency range, $f\in[f_1,f_2]$, in which TL $\geqslant$10 dB. The cost function for the optimization is set as the ratio of the band limits, f_2/f_1, namely the logarithmic bandwidth. The reason why 10 dB is chosen is the following. When the membranes are removed, the cavities, one on each side, form an expansion chamber of area ratio $(1+2h_c)=3$. The best performance of sound reflection by this reference expansion chamber is about TL=5 dB. It is felt that TL=10 dB would represent a substantial improvement. It is possible that the performance of cavities filled with sound absorption materials might be even better than 10 dB at moderate to high frequencies, but the emphasis of the current study is on the control of low-frequency noise. Most of the

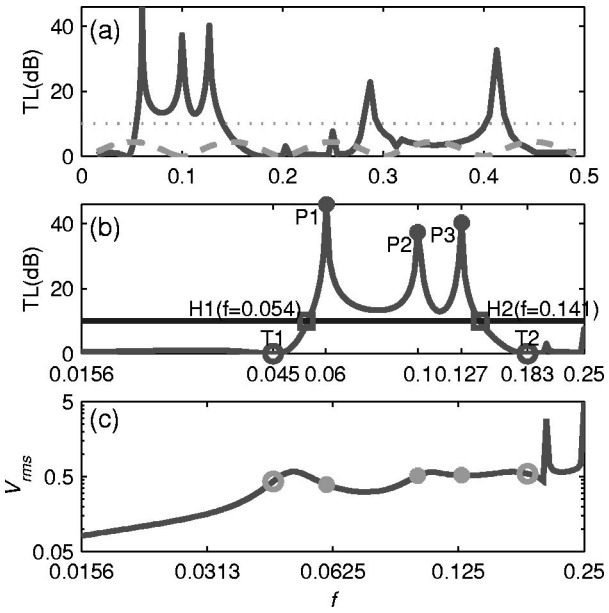


FIG. 3. Performance of two opposite membranes under tension $T=0.475$. The dashed curve in (a) is for a simple expansion chamber with an area ratio of $(1+2h_c)=3$. The peaks are illustrated by ● and the troughs by ○.

following analyses show a spectrum over four octaves. The frequency of four octaves down from $f=0.5$ is $0.5/2^4=0.0313$, which is 62.5 Hz for the above example of the 170 mm duct.

Figure 3 shows the results for the default configuration specified in (30). The scale of frequency in Fig. 3(a) is linear, and it covers the whole plane wave frequency for the rigid channel, $f\in(0,0.5)$. The horizontal dashed line is the threshold level 10 dB. The dashed curve in Fig. 3(a) is drawn for the plane-wave theory of the expansion chamber of an area ratio of 3. A wide stopband is found in the low frequency range below $f=0.15$, which is the range of focus for the present study. Figures 3(b) and (c) use a logarithmic scale for the four octave frequency band from $f=0.0156$ to 0.25. The ranges of frequencies shown in these two subfigures are identical, but in Fig. 3(b) the frequencies of troughs and peaks are labeled for easy reference. Figure 3(b) shows that the stopband begins from $f_1=0.054$ (marked as point H1) and ends at $f_2=0.141$ (marked as point H2), the ratio being $f_2/f_1=2.6$. Three peaks are marked in both Fig. 3(b) and Fig. 3(c), by asterisks and labeled as P1, P2, and P3. Two troughs are identified as open circles, and labeled as T1 and T2, respectively. Figure 3(c) shows a gradually increasing rms level of the membrane response calculated by $V_{\rm rms}=(\Sigma_j|V_j|^2/2)^{1/2}$. The vertical coordinate is logarithmic to show more detail features. The important observation here is that the peaks and troughs do not correspond to the level of high or low membrane response, as shown by the lack of correlation between the marked peaks and troughs with the actual variation pattern of $V_{\rm rms}$. In other words, the effectiveness of the membrane to reflect sound does not solely depend on the amplitude of the induced membrane vibration. As is shown later, it mainly depends on the acoustic interference of sound radiated by different parts of the membrane. Since the membrane vibration is decomposed into axial modes of *in vacuo* membrane vibration, $\sin(j\pi\xi)$, it is con-

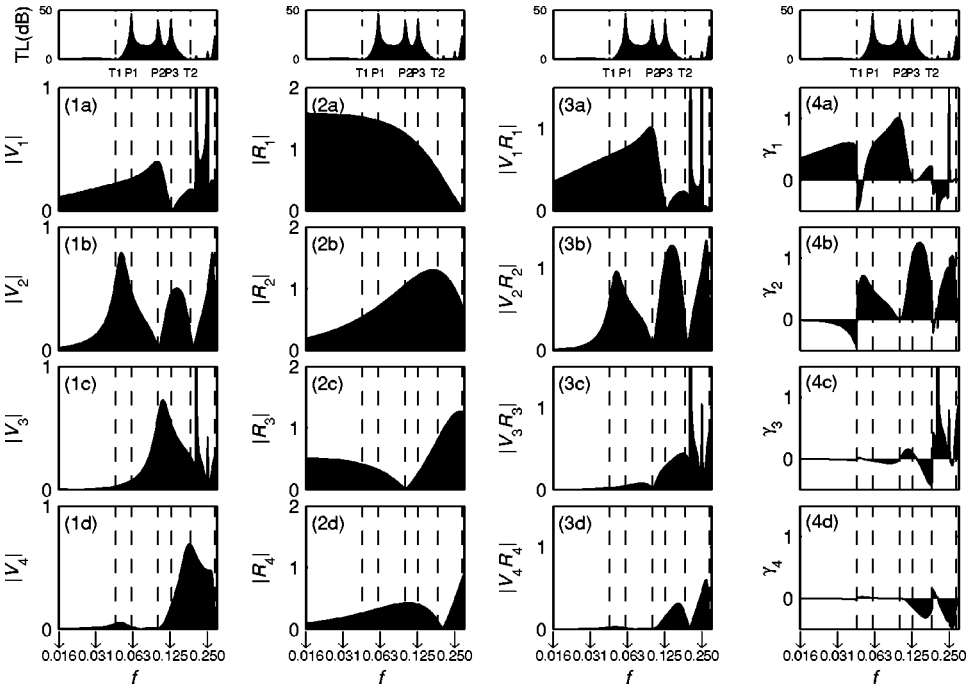


FIG. 4. Modal reflections. The first column on the left is for the modal amplitude, $|V_j|$, the second column is for the modal reflection coefficient, $|R_j|$, the third column is for the single mode reflection $|V_jR_j|$, and the last column on the right is for the modal contribution γ_j defined in Fig. 2. The top row is the TL spectrum for the purpose of identifying important frequencies.

venient to investigate the interference pattern in terms of these modes.

B. Modal reflections

The modal response for the default membranes is shown in Fig. 4. The vibration amplitude of each individual mode, $|V_j|$, is shown in the subfigures on the left column. A TL spectrum is attached on top of each column to facilitate the correlation study between the peak and trough frequencies analyzed in Fig. 3, and the modal response. The second column is the amplitude of the modal sound reflection coefficient, $|R_j|$, which is also related to the modal loading of the incident wave I_j by $R_j = I_jL/4$, cf. Eqs. (28) and (23). The third column is the amplitude of the single mode reflection, $|V_jR_j|$. The fourth column shows the modal contribution towards the total sound reflection, γ_j, defined in Eq. (29).

The membrane resonance points are analyzed first. As shown in Fig. 4(1a), there are two sharp peaks in the first mode amplitude, $|V_1|$, one next to the frequency of T2, and one right on the border frequency of $f=0.25$. There is also a 3rd mode peak, $|V_3|$, shown in Fig. 4(1c). These peaks are also seen in the response curve of Fig. 3(c). It might be tempting to relate these frequencies with conditions where the membrane length is equal to the multiples of the incident wavelength or the quarter wavelength, but such simple analysis does not hold since these frequencies change with tension. The two frequency points certainly represent membrane resonance (to be discussed later), but such resonance does not produce much sound reflection since TL is not high at these frequencies. The reason why TL is low can be found from the figures in the fourth column. Taking $f=0.2$ as an example, there is a sharp negative peak in Fig. 4(4a), which is clipped in the figure, and a sharp positive peak in Fig. 4(4c). Both modal ineffectiveness and intermodal cancellation are manifestations of destructive acoustic interference of sound radiated by different portions of the membrane. Details of the modal response for the first two modes and the specific frequency points where TL is high are analyzed in the next section.

The degree of intermodal coupling can be seen in the left column of Fig. 4. For example, in the frequency range below about 0.05, $|V_3|$ shown in 4(1c) is much smaller than $|V_1|$ shown in Fig. 4(1a). Therefore, a single mode analysis can be conducted for the first mode below $f=0.05$. Beyond this point, the coupling effect is strong and a detailed analysis becomes difficult and less rewarding. For the coupling between the second and the fourth modes shown in Figs. 4(1b) and (1d), single mode analysis can be conducted for the second mode below $f=0.1$. Further analysis is concentrated in these low-frequency ranges where the cross-modal coupling is weak.

IV. LOW FREQUENCY AND PEAK POINTS

Within the frequency range where the first two modes are decoupled from the higher order modes, it is possible to find a closed-form solution and reveal the functional relationship for important features of the coupled system. For this purpose, the core expressions of modal impedances shown in Eqs. (19) and (20) are analyzed here for both odd-odd and even-even combinations of modal indices, j, l. Introducing the following membrane and cavity parameters:

$$\theta_n = k_{nc}L_c, \quad \theta \equiv \theta_0 = k_{0c}L_c = k_0L, \quad \Theta = k_{0c}L_v, \qquad (31)$$

and recalling from Eq. (18),

$$Z_{-jl} = L_c\mathcal{I}_{2c}(0,j,l) + 2\sum_{n=1}^{\infty} c_{nc}\mathcal{I}_{2c}(n,j,l),$$

$$\mathcal{I}_{2c}(n,j,l) = \frac{2j\pi l\pi[1-e^{ij\pi - i\theta_n}]}{[(j\pi)^2-\theta_n^2][(l\pi)^2-\theta_n^2]} + \frac{i\theta_n\delta_{jl}}{(j\pi)^2-\theta_n^2},$$

where c_{nc}, $k_{nc}L_c$ for the higher order duct modes are purely imaginary and can be approximated as follows:

$$c_{nc}=\frac{i}{\sqrt{(n/2f)^2-1}}\approx\frac{i2f}{n},$$

$$\theta_n=k_{nc}L_c\approx\frac{2\pi f}{i2f/n}L_c=-in\pi L_c.$$

Since $|c_{nc}|$ is very small and $|\theta_n|$ very large, the denominator $[(j\pi)^2-(\theta_n)^2]\approx[(L_cn\pi)^2+(j\pi)^2]$ diverges as $(L_cn)^2$ as n increases. The two terms in $\mathcal{I}_{2c}$ $(n>0,j,l)$ diverge as $(L_cn)^{-4}$ and $(L_cn)^{-1}$, respectively. Numerical analysis shows that, for the typical parameters set in Eq. (30), the higher order modes $\mathcal{I}_{2c}$ $(n>0,j,l)$ contribute little to the radiation impedance Z_{-jl}. It is therefore only necessary to concentrate on the plane wave mode, $n=0$. The same conclusion holds for the reflection impedance Z_{rjl}.

A. Plane wave approximation

The plane wave approximation, which neglects all higher order duct acoustics modes, gives

$$Z_{-jl}\approx L_c\left[\frac{2jl\pi(1-e^{il\pi-i\theta})}{[(j\pi)^2-\theta^2][(l\pi)^2-\theta^2]}+\frac{i\theta\delta_{jl}}{(j\pi)^2-\theta^2}\right],$$

$$Z_{rjl}\approx L_c\frac{8jl\pi(1-\cos j\pi\cos\theta)(1-e^{i\Theta+ij\pi})}{[(j\pi)^2-\theta^2][(l\pi)^2-\theta^2][e^{i2\Theta}-1]}.$$

For $L=L_v=5$, and $f=0.0156$, 0.0313, 0.0625, 0.125, $\theta=\Theta$ is, respectively, 0.491, 0.982, 1.964, and 3.927. In the frequency range where the first two modes are quite decoupled from the higher order modes, as discussed before this section, the solutions to the dynamics Eq. (22) are simply

$$V_1\approx\frac{-I_1}{(Z_{11}+\mathcal{L}_1)},\quad V_2\approx\frac{-I_2}{(Z_{22}+\mathcal{L}_2)}.$$

For a configuration in which there are two identical membranes and accounting only the duct acoustics mode of $n=0$, the total modal impedance is $Z=(2Z_{+\mathrm{rad}}+Z_{-\mathrm{rad}})+Z_{-\mathrm{ref}}$. The plane wave approximation of Z_{11} is

$$Z_{11}=(L_2+2L)\left[\frac{2\pi^2(1+e^{-i\theta})}{[\pi^2-\theta^2]^2}+\frac{i\theta}{\pi^2-\theta^2}\right]-2L_c\frac{2\pi^2(1+\cos\theta)}{[\pi^2-\theta^2]^2[1-e^{i\Theta}]},$$

where the last group is from $Z_{-\mathrm{ref}}$, which can be further divided into resistance and reactance term by

$$\frac{1}{1-e^{i\Theta}}=\frac{1-\cos\Theta+i\sin\Theta}{2-2\cos\Theta}=\frac{1}{2}+\frac{i}{2}\cot\frac{\Theta}{2}.$$

It can be shown that the radiation resistance in the cavity is cancelled out by the real part of the cavity reflection impedance, so that

$$Z_{11}=\overbrace{\left[(2L)\frac{2\pi^2(1+\cos\theta)}{[\pi^2-\theta^2]^2}\right]}^{\text{Radiation resistance}}-\overbrace{i(L_c+2L)\left[\frac{2\pi^2\sin\theta}{[\pi^2-\theta^2]^2}-\frac{\theta}{\pi^2-\theta^2}\right]}^{\text{Radiation reactance (stiffness)}}-\overbrace{iL_c\left[\frac{2\pi^2(1+\cos\theta)}{[\pi^2-\theta^2]^2\tan(\Theta/2)}\right]}^{\text{Cavity reflection reactance (stiffness)}}. \tag{32}$$

Similarly,

$$Z_{22}=\overbrace{\left[(2L)\frac{8\pi^2(1-\cos\theta)}{[(2\pi)^2-\theta^2]^2}\right]}^{\text{Radiation resistance}}+\overbrace{i(L_c+2L)\left[\frac{8\pi^2\sin\theta}{[(2\pi)^2-\theta^2]^2}+\frac{\theta}{(2\pi)^2-(\theta)^2}\right]}^{\text{Radiation reactance (mass)}}+\overbrace{iL_c\left[\frac{8\pi^2(1-\cos\theta)\tan(\Theta/2)}{[(2\pi)^2-\theta^2]^2}\right]}^{\text{Cavity reflection reactance (mass)}}. \tag{33}$$

Notice that, when $\Theta<\pi$, the imaginary part of Z_{11} presents stiffness, but that of Z_{22} a mass. The magnitude of the reactance is small for Z_{22} when $\theta\to0$, but that of Z_{11} diverges like any spring stiffness, and the divergent term is from the cavity reflection. When $\Theta>\pi$, the reverse is true. When Θ passes π, $\tan(\Theta/2)\to\infty$, the cavity has zero stiffness to the first mode but infinite stiffness for the second mode. From this point onwards, Θ is assumed to be the same as θ.

The accurate values of the two impedances are shown in Fig. 5 with sources of contributions marked along the curves.

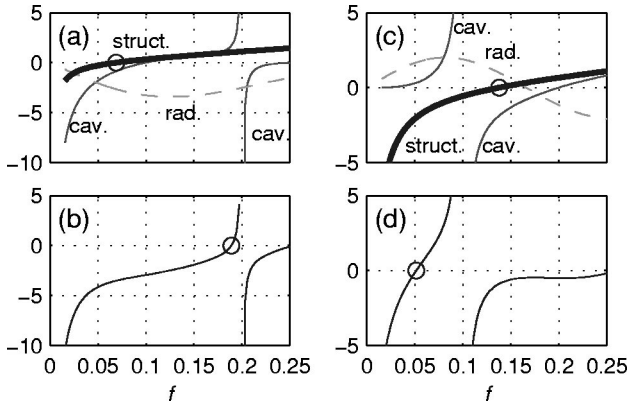


FIG. 5. Reactance of the first two modes and resonant frequencies. (a) shows the components of the first mode reactance. The components are from radiation (marked rad), cavity reflection (marked cav), and structural properties (struct). The open circle marks the frequency of the first *in vacuo* eigen frequency. (b) shows the total reactance for which the zero reactance is also marked by an open circle. (c) shows the second mode components where the second *in vacuo* eigen frequency is marked by an open circle. (d) shows the total reactance of the second mode, where the open circle marks the frequency where the total reactance vanishes.

The two subfigures on the left-hand side are for the first mode reactance including the structural contribution, $\mathrm{Im}(Z_{11}+\mathcal{L}_1)$, while those on the right-hand side for the second mode, $\mathrm{Im}(Z_{22}+\mathcal{L}_2)$. The dashed lines in Fig. 5(a) and Fig. 5(c) are for the radiation components, which are not remarkable except that the second mode radiation reactance vanishes at $f=0.166$. The thin solid lines in these two subfigures are for the cavity reflection reactance. The reactance diverges at the cavity modes where the cavity length is a multiple of the half wavelength, $\theta=j\pi$, $f=0.1,0.2$, $j=2,1$. The two thick solid lines are the structural dynamics contribution which is the combination of inertia and stiffness terms. The open circles indicate the *in vacuo* membrane modes at frequencies $f=(j\pi/L)\sqrt{T/m}=0.069$, 0.138 for $j=1,2$, respectively. Figures 5(b) and (d) show the total reactance. When the cross-modal coupling is ignored, the coupled system resonates where the total reactance vanishes. Note that the frequency range shown in this figure is larger than the interaction-free range discussed at the end of Sec. III.

Figures 5(a) and (b) show that the first mode is rather stiff for very low frequencies mainly due to the cavity reflection effect. But the stiffness domination continues beyond the first *in vacuo* mode frequency due to the radiation stiffness. For a piston or a compact surface to radiate sound into an open space with or without baffle, the radiation reactance is normally positive, i.e., there is virtual mass from the surrounding air. But the situation for sound radiation in a duct is different. The opposite wall causes a mirror effect, and part of the radiation resistance is transformed into negative radiation mass or positive radiation stiffness (Huang, 2000). In such case, it makes more sense to regard this part of reactance as negative mass since its frequency dependence is inertialike [see the initial segment of the dashed line in Fig. 5(a)]. This mass can cancel out the structural mass, the former being larger in amplitude in the current example. This observation leads us to the conclusion that the structural mass is not important at all for the first mode response at low frequencies. The very high stiffness in the low-frequency range also indicates that the first mode cannot be relied upon for sufficient membrane reflection of the grazing incident sound. The total reactance is shown in Fig. 5(b). The first membrane mode now occurs at $f=0.189$, which is beyond the second *in vacuo* mode frequency. Another zero point can be found in Fig. 5(b) beyond $f=0.25$. The two resonance points found in Fig. 5(b) roughly correspond to the two peaks of $|V_1|$ shown on the right-hand edge of Fig. 4(1a). The qualifying word roughly is used here to indicate the lack of cross-modal coupling between the first and third mode. Again, if there was no cross-modal interactions, the first cavity mode at $f=0.2$ should have caused the membrane to have zero first mode response. The fact that this is not the case is shown in Fig. 4(1a).

Since $|V_4|$ is very small for $f<0.1$, the cross-modal coupling between the second and the fourth mode occurs strongly only beyond $f=0.1$. It is sufficient to analyze the reactance of the second mode alone in Figs. 5(c) and (d). Here, the cavity mode at $f=0.1$ indeed eliminates the second mode resonance, as shown in Fig. 4(1b). For $f<0.1$, both radiation and cavity reflection reactance are masslike, leaving the structural stiffness to dominate the coupled membrane response. The frequency of the second *in vacuo* mode, $f=0.138$, is shifted to $f=0.0515$ due to fluid added mass. This frequency corresponds to the peak of second mode response shown in Fig. 4(1b). This peak is higher than the first mode response in this region, shown in Fig. 4(1a), but its effectiveness in reflecting sound is not as good, cf. Figs. 4(2a) and (2b). The contribution towards sound reflection match in magnitude only beyond the second mode peak, see Figs. 4(3a) and (3b), where satisfactory overall performance of the membrane can be expected. It is therefore very important to determine analytically the second mode resonance.

B. Second mode resonance

For the second mode, $j=2$, θ in $[(j\pi)^2-\theta^2]$ can be ignored for a rough estimate without loss of qualitative information. The impedance Z_{22} given in Eq. (33), and the incident wave coefficients, I_2 given in Eq. (23), are approximated as follows:

$$4\pi^2 Z_{22}\approx(4L)(1-\cos\theta)+i(L_c+2L)[2\sin\theta+\theta]+iL_c 2(1-\cos\theta)\tan(\Theta/2),$$

$$I_2\approx\frac{1}{\pi}e^{i\theta/2}(1-e^{-i\theta}).$$

The second mode response is found to be

$$\begin{aligned} V_2&\approx\frac{-I_2}{Z_{22}+\mathcal{L}_2}=\frac{-i4\pi L^{-1}}{8\sin(\theta/2)+iX},\\ X&=(h_c^{-1}+2)[4\cos(\theta/2)+\theta/\sin(\theta/2)]\\ &\quad+4h_c^{-1}\sin(\theta/2)\tan(\Theta/2)\\ &\quad+(2\pi/L)^2[m\theta-T4\pi^2/\theta]/\sin(\theta/2). \end{aligned}\tag{34}$$

A weak resonance occurs when the amplitude of the denominator approaches a minimum, and this minimum can be shown to be approximately where the reactance term X in the denominator vanishes. By assuming a vanishing θ, and that Θ is also rather small even if it's larger than θ, the approximated resonance condition $X=0$ can be simplified as

$$[6(h_c^{-1}+2)8(\pi/L)^2(m-T4\pi^2/\theta^2)]\approx 0,$$

so that

$$f_{2\,\mathrm{res}}\approx\frac{gv}{2\pi L}=\frac{1}{L}\sqrt{\frac{T}{3(2+h_c^{-1})(L/2\pi)^2+m}}.\tag{35}$$

For the example in which $T=0.475$, $m=1$ and $L_c=L=5$, $f_{2\,\mathrm{res}}$ is found by this formula to be $f_{2\,\mathrm{res}}\approx 0.0533$, while the accurate root for $X=0$ is $f_{2\,\mathrm{res}}=0.0504$. Note that $\theta=0.5$ is not so small, but somehow the errors cancel themselves out coincidentally. Note also that X here only accounts for the plane wave contribution, which differs slightly from the full reactance shown in Fig. 5(d). Although the optimal value of T depends on L and m, it can be observed that the membrane length is the most influential parameter for the second mode resonance frequency.

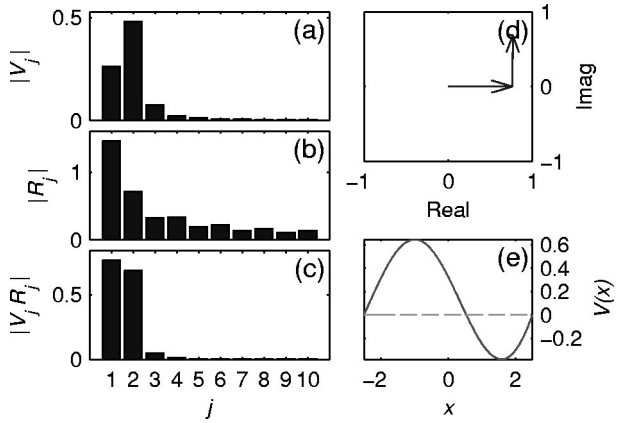


FIG. 6. Performance of the first peak at $f=0.06$, P1. (a) shows the velocity amplitude of each mode, $|V_j|$. (b) is the amplitude of the modal reflection coefficient, $|R_j|$. (c) is the sound reflection by each mode alone, $|V_jR_j|$. (d) shows how the complex modal contributions add up. (e) is the real (solid line) and imaginary (dashed line) parts of the membrane vibration velocity distribution, $V(x)$.

C. Peak performance points

The interference patterns of the first three peaks shown in Fig. 3(a), P1, P2, and P3, are analyzed as follows. Figure 6 is for P1 at $f=0.06$. The amplitudes of V_j, R_j, and V_jR_j are shown in Figs. 6(a), (b), and (c), respectively. The first mode proves more difficult to excite than the second, but the first mode is more effective in reflecting sound, $|R_1|>|R_2|$. As a result, the amplitudes of V_1R_1 and V_2R_2 are about equal. As shown in Fig. 6(d), the contributions from the two modes are nearly out-of-phase, hence no interference. The resultant membrane vibrations are, however, almost in-phase over the entire length as there is almost zero imaginary part in $V(x)$ shown in Fig. 6(e). The vibration is approximately described by $\sin(\pi\xi)+1.84\sin(2\pi\xi)$, which appears to be a distorted second mode with the left half-length extended and the right half-length reduced.

The modal contributions and their interference for P2 at $f=0.1$ are shown in Fig. 7. Here, the second mode proves almost impossible to excite since there is a first cavity mode, as explained by the divergence of the cavity reflection impedance in Eq. (21). The result is an in-phase combination of the first and the third mode in the form of $\sin(\pi\xi)$ $-1.50\sin(3\pi\xi)$. At this frequency, the third mode is not effective in reflecting sound. So that the only contribution towards reflection comes from the first mode. The reason why the first and third modes are excited so much at the cavity mode is that the cavity stiffness vanishes at this frequency, see the term $\tan^{-1}(\Theta/2)$ in Eq. (32), where $\Theta=\pi$ for this cavity mode.

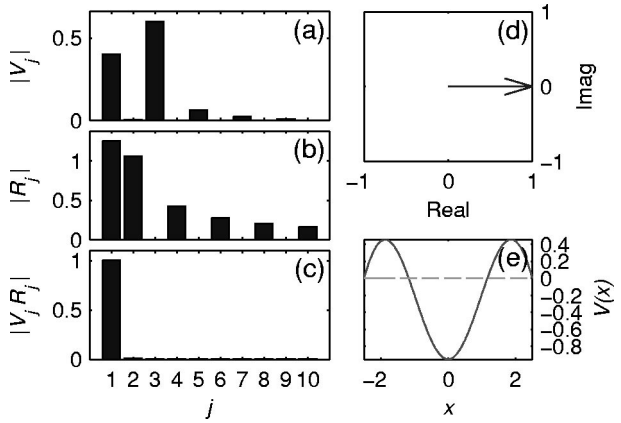


FIG. 7. The response of the membrane at the second peak of $f=0.102$, P2, cf. caption of Fig. 6.

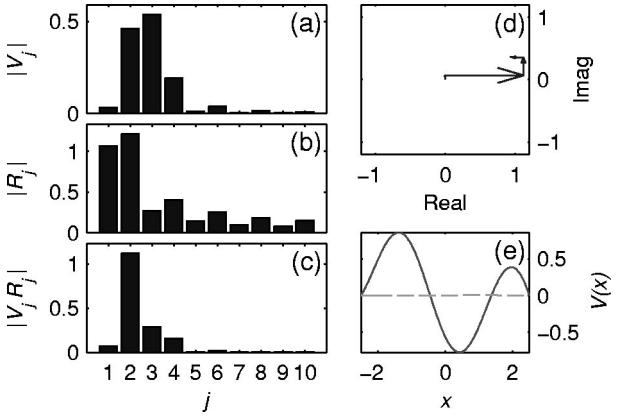


FIG. 8. The response at the third peak at $f=0.125$, P3, cf. caption of Fig. 6.

The third peak at $f=0.127$, P3, is analyzed in Fig. 8. In this case, the first mode is almost absent. The presence of the second and third modes is almost equally strong. The two modes are roughly in phase and the vibration velocity can be approximated by $\sin(2\pi\xi)+1.17\sin(3\pi\xi)$. Sound reflection mainly derives from the second mode, which is only slightly cancelled out by that from the fourth mode. The strong action by the second mode is caused by the increase in both the strength of the incident wave, I_2, and the modal reflectivity coefficient R_2, the two being related by $R_2=I_2L/4$. The excitation of the second mode vanishes when the incident wave coincides with the wavelength of the first mode, or the first cavity mode discussed above, which is $f=0.1$ in this case. As f moves away from this value, I_2 increases, as shown in Fig. 4(2b), and the stiffness term in Z_{22} reduces rapidly, as shown in Fig. 5(d).

V. CONCLUSIONS

Three figures are central to the understanding of the modal response and interaction pattern. They are Figs. 3, 4, and 5. The following conclusions are drawn for the membrane performance in the order of increasing frequency.

(a) When the frequency is extremely low, say, $f=0.0165$, the cavity is compact. The cavity is very stiff to the first mode and imposes high mass load to the second. In other words, the air in the cavity is barely compressible to the volume-displacing mode. The vibration of the second mode is not affected by the incompressibility, but it suffers from two adverse factors: (1) it's acoustically ineffective in reflecting sound and it's not excited, and (2) it's structurally hard due to the curvature of the deformed membrane. The limited reflection by the second mode only serves to cancel out that from the first mode until beyond the frequency of its own resonance. It can be said that the response of the membrane is controlled by the cavity for the first mode, and by the structural properties for the second mode.

(b) As frequency gradually increases towards $f=0.05$, the response of the first mode increases mainly due to the rapid reduction of the cavity stiffness. Since the structure properties do not have much influence in the particular configuration studied in this article, it can be read from Fig. 5(b) that the total stiffness reaches a rather low level around $f=0.05$, or when the ratio of the membrane length to wavelength is $L^*f^*/c^*=fL=0.25$, i.e., a quarter wavelength. Further reduction in system stiffness is gradual as f goes beyond 0.05. The increase in the action of the second mode is even more rapid when f approaches 0.05 for two reasons. One is that the structure is less stiff, which is natural to all modes, and the other is that the excitation, I_2, increases rapidly. The fluid exerts only virtual mass on the second mode. As a result, the structural stiffness is balanced by the virtual mass. The fluid and structure are coupled on equal footing to bring out a second mode resonance. As the frequency increases, the interference pattern between the first mode reflection and the second mode interference has changed from one of cancellation at DC frequency to one of out-of-phase relationship. The first peak in transmission loss, P1, appears at a frequency slightly beyond the second mode resonance. The main feature during the frequency running up to P1 is that the cavity has released its grip on the first mode, and the fluid couples closely with the second structural mode.

(c) Between the first peak point P1 and the second peak point P2, the pattern of reflection wave noninterference between the first two modes is rather stable. The contribution from the first mode, γ_1, grows with frequency due to the continuous decrease in cavity effect, which seems to have overridden the adverse effect of the decreasing receptivity to incident waves. Without a detailed analysis, it is difficult to determine the effect of the cross-modal coupling between the first and third modes within this frequency range. The dominant feature within the range, however, is the declining contribution of the second mode contribution, γ_2, until it vanishes at P2 at $f=0.1$, where the second mode is totally prohibited by the first cavity mode. The trade-off between the declining γ_2 and the increasing γ_1 resulted in a dip which in this example is still above 10 dB. If γ_2 were to decline a little faster than it does, the dip could go below 10 dB, leaving the TL pattern as one of normal Helmholtz resonator in which the sharp resonant peak is isolated. In order to halt the pace of this decline, the increase in the total reactance $\mathrm{Im}(Z_{22}+\mathcal{L}_2)$, shown as the first branch curve in Fig. 5(d), should be as gradual as possible. Since Z_{22} is by and large a function of geometry, the structural property in $\mathcal{L}_2$ can be adjusted in order to prevent the dip to go too low.

(d) The events between the second and third peaks begin to involve the third and fourth modes. However, their contribution, γ_3 and γ_4, remain to be small. The second peak P2 features the solo action of the first mode, as explained earlier. The third peak is contributed to mainly by γ_2. The reason why γ_2 is gaining ground towards this frequency is that the second mode continues to gain receptivity to the incident sound, and it moves away from the banned region of the first cavity mode. Part of the reason why γ_1 disappears totally is that the modal response V_1 decreases in amplitude due to the poor modal receptivity, and partly due to the cross-modal interaction between the first and third modes, which is too involved to discuss in detail. There is also a dip between the two peaks, which is again determined by the trade-off between the declining γ_1 and increasing γ_2. The trade-off in this region is exactly opposite to what happens between P1 and P2, where γ_1 increases but γ_2 decreases. Overall, it is found that the actions of the first two modes are by and large supplementary to each other, a key element in sustaining a high TL over a broad frequency band.

The analyses so far have been conducted for one particular set of parameters listed in Eq. (30) for the purpose of in-depth understanding of the fluid-membrane coupling mechanism. A detailed parametric study has to be conducted before an actual design can be made. Without giving details of such a study, the following discussions are reported. First, the membrane tension plays a pivotal role in tuning the relative response of various axial modes and the positions of the peaks in the transmission loss spectrum as shown in Fig. 3. The choice of $T=0.475$ is made in such a way that the spectral troughs between P1 and P2 is roughly level with that between P2 and P3. The result is a balanced spectrum with a wide stopband. When the tension vanishes, it can be demonstrated that the response of the system is essentially a simple expansion chamber. When tension approaches infinity, the membranes do not vibrate at all as they become hard walls. An optimal tension exists and the value of such tension depends on the choice of the objective function in the optimization process. In the particular example given with a square duct of height 17 cm, the tension corresponding to $T=0.475$ is 1944N. The dimensional force obviously increases with h^{*2} when other dimensionless parameters are fixed, and the dimensional frequency for the stopband decreases in proportion to c_0^*/h^*. The ratio of the membrane length to the duct height, $L=L^*/h^*$, is also very important. When L is too short, one can expect a very ineffective second mode in receiving and reflecting long waves, see Eqs. (23) and (28). When L is too long, however, the method of traditional duct lining with an acoustically transparent membrane cover may become more attractive. Again, some optimal length exists for a specific design task.

ACKNOWLEDGMENT

The author thanks the Hong Kong Polytechnic University for its support through the research project G-YC93.

Cheng, L. (**1994**). ''Fluid-structural coupling of a plate-ended cylindrical shell: vibration and internal sound field,'' J. Sound Vib. **174**, 641–654.

Doak, P. E. (**1973**). ''Excitation, transmission and radiation of sound from source distributions in hard-walled ducts of finite length (I): the effects of duct cross-section geometry and source distribution space-time pattern,'' J. Sound Vib. **31**, 1–72.

Dowell, E. H., Gorman, G. F., and Smith, D. A. (**1977**). ''Acoustoelasticity: general theory, Acoustic natural modes and forced response to sinusoidal excitation, including comparisons with experiment,'' J. Sound Vib. **52**, 519–542.

Dowell, E. H., and Voss, H. M. (**1963**). ''The effect of a cavity on panel vibrations,'' AIAA J. **1**, 476–477.

Frommhold, W., Fuchs, H. V., and Sheng, S. (**1994**). ''Acoustic performance of membrane absorbers,'' J. Sound Vib. **170**, 621–636.
Fuchs, H. V. (**2001a**). ''From advanced acoustic research to novel silencing procedures and innovative sound treatments,'' Acustica **87**, 407–413.
Fuchs, H. V. (**2001b**). ''Alternative fiberless absorbers—new tools and materials for noise control and acoustic comfort,'' Acustica **87**, 414–422.
Guy, R. W. (**1979**). ''The response of a cavity backed panel to external airborne excitation: a general analysis,'' J. Acoust. Soc. Am. **65**, 719–731.
Horoshenkov, K. V., and Sakagami, K. (**2001**). ''A method to calculate the acoustic response of a thin, baffled, simply supported poroelastic plate,'' J. Acoust. Soc. Am. **110**, 904–917.
Huang, L. (**1999**). ''A theoretical study of duct noise control by flexible panels,'' J. Acoust. Soc. Am. **106**, 1801–1809.
Huang, L. (**2000**). ''A theory of reactive control of low-frequency duct noise,'' J. Sound Vib. **238**, 575–594.
Kang, J., and Fuchs, H. V. (**1999**). ''Predicting the absorption of open weave textiles and micro-perforated membranes backed by an air space,'' J. Sound Vib. **220**, 905–920.
Lyon, R. H. (**1963**). ''Noise reduction of rectangular enclosures with one flexible walls,'' J. Acoust. Soc. Am. **35**, 1791–1796.
Mechel, F. P. (**2001**). ''Panel absorber,'' J. Sound Vib. **248**, 43–70.
Pan, J., and Bies, D. A. (**1990**). ''The effect of fluid-structural coupling on sound waves in an enclosure—theoretical part,'' J. Acoust. Soc. Am. **87**, 691–707.
Pretlove, A. J. (**1965**). ''Free vibrations of a rectangular panel backed by a closed rectangular cavity,'' J. Sound Vib. **2**, 197–209.
Sum, K. S., and Pan, J. (**1998**). ''A study of the medium frequency response of sound field in a panel-cavity system,'' J. Acoust. Soc. Am. **103**, 1510–1519.

Experimental studies of a drumlike silencer

Y. S. Choy and Lixi Huang[a)]
Department of Mechanical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong

(Received 26 January 2002; revised 15 July 2002; accepted 15 July 2002)

The theoretical finding of the broadband performance of a reactive silencer is validated experimentally. The silencer consists of two highly stretched membranes lining part of the duct and backed by two long and shallow cavities. The test rig was built with a small square duct of 5 cm in dimension, and each cavity is 5 cm deep and 25 cm long. Two types of metal foils, stainless steel and copper, were used, and the lowest membrane-to-air mass ratio was 1.3. A transmission loss in excess of 10 dB was achieved over more than one octave band. For one configuration close to the optimal parameters, the predicted ratio of the frequency band limits is 2.47, while the experiment gave 2.35. Three spectral peaks were found in the stopband, as predicted, but the peaks were broader than prediction, indicating the presence of significant sound energy dissipation mechanisms. Comparison with theoretical simulation shows that the cavity damping dominates over membrane friction. Tests using heavier membranes and membrane with different levels of tension also agree with predictions. Issues of practical implementation of the concept as a flow-through silencer are also addressed. [DOI: 10.1121/1.1508779]

PACS numbers: 43.50.Gf, 43.20.Tb, 43.20.Ks [MRS]

I. INTRODUCTION

In a recent theoretical study, the second author explored the use of tensioned membranes as a sound reflection device with zero pressure loss penalty in a flow duct (Huang, 1999). In that study, the effect of air external to the membrane, which forms part of the duct wall, was excluded. In order to explore the use of the concept in a real silencer, a cavity has to be added to prevent the noise break-out. The formulation has since been extended to include the important cavity effect. Preliminary tests were carried out by the first author to validate the theoretical framework (Choy, 2001). With further refinement on the prediction, it now becomes clear that far better results are possible with membranes stretched far beyond the elastic limit of the plastic material used in those initial tests. The present paper reports the full comparison between the latest prediction and the tests using metal foils as membranes. The performance of these metal foils is comparable with that in many related studies, which are very briefly reviewed below in the wider context of low-frequency noise control.

Low-frequency sound radiation often carries most of the sound power in many engineering applications, such as the ventilation fans and the vibrating machinery. In general, the thickness of sound absorption material required in noise control engineering is tied to the wavelength. Effective absorption of low-frequency noise is impractical in most circumstances. Before the discussion can be extended to reactive and active noise control methods, it is necessary to specify the circumstances of the noise problem. In the area of room acoustics, in which sound strikes the walls with normal or random incidence, a few effective treatments have been developed, such as Helmholtz resonator and membrane, or plate, absorbers. Helmholtz resonators have been used to enhance low-frequency noise absorption since, perhaps, very long time ago in human history (Brüel, 1951). Recently, the principle of Helmholtz resonator is again becoming essential as the microperforated (submillimeter holes) noise absorption techniques (Maa, 1975, 1998) are being rapidly developed to offer fiberless and relatively broadband solutions (Zha *et al.*, 2002) for broadcasting studios, music and concert halls. Theoretical studies in this area have also increased in their sophistication, e.g., Brown, 1964; Ford, 1969; Sakagami *et al.*, 1996; Kang and Fuchs, 1999; and Horoshenkov and Sakagami, 2001. In its simplest setup, a panel absorber is a mass-spring system in which the mass derives from the panel, and the stiffness is mainly determined by the air cavity although the panel may also have a significant contribution. A shallow cavity gives high stiffness, and this is one of the main reasons why low-frequency, broadband noise is difficult to deal with. In this regard, the concept of magnetic field neutralizing the cavity air stiffness introduced by Huang (2000) could be a potential source of advancement in this field.

In the area of duct noise control, absorption of low-frequency noise by duct lining is more difficult than in rooms due to the grazing incidence in duct. Effective noise control is achieved with intrusive devices, such as a splitter silencer, or normal vehicle silencers in which the air flow passage is tortuous. But there are many important utilities, such as power plant chimneys and central ventilation systems, where the pressure loss caused by such a design means a substantial running cost. One alternative to achieving low-frequency noise abatement in a duct is the active noise control technique. Despite the rapid development in this field in the last few decades, successful applications of this old concept are still few. Apart from many electronic intricacies, the ineffectiveness of the secondary sound source at low frequencies may also be a factor.

The group led by Fuchs (2001) have been in the forefront of taming the low-frequency noise without imposing

[a)]Author to whom correspondence should be addressed. Electronic mail: mmlhuang@polyu.edu.hk

much pressure loss penalty. Among many patents they have developed, two are particularly relevant to the present studies. One is called the membrane absorber box, as described in detail by Ackermann, Fuchs, and Rambausek (1988), and the other is called the 'cleanable reactive silencer' [see Zha *et al.* (2002) for a brief scientific discussion]. The latter essentially looks like an expansion chamber except that the smaller tube runs all the way through the chamber with a small perforated segment immediately downstream of the first junction. The device is reactive, similar to the drumlike silencer described in the present study, but the insertion loss spectrum is typical of a resonator, which is what it essentially is. The biggest advantage here is that there is no pressure loss with this design, and the disadvantage is that many such silencers have to be installed in a row to achieve a broadband performance.

The device of membrane absorber box is a dissipative one with a broadband performance in the low to medium frequency range. It is made by a host of combined Helmholtz and panel resonators in a honeycomb structure. A typical resonator unit consists of a hollow rectangular cavity covered by a perforated membrane. Another impervious thin membrane is placed in front of the perforated membrane to form a narrow channel between the two sheets. The system of perforated panel with cavity resonates at a frequency about twice the resonant frequency of the Helmholtz system formed by the cavity and the aperture. The two resonances are only weakly coupled. Without the impervious membrane cover, the effect of the combined resonator would be rather poor as the spectrum of each resonator is peaky. But the narrow channel between the two sheets helps to increase the system resistance and the dip in the sound absorption spectrum is reduced. The details of the unit impedance are described in Frommhold *et al.* (1994). A splitter silencer built on this concept was successfully used in the exhaust stack of a papermill (Ackerman and Fuchs, 1989). The duct was 1 m wide and the silencer was 1 m long. Each splitter is 100 mm thick and the distance between adjacent splitters is 150 mm. An insertion loss of 10 dB or above was achieved over one octave band starting from 175 Hz. The device is mainly absorptive although some reactive mechanism might also be at work. This contrasts with the entirely reactive mechanism of the drumlike silencer being explored in the current study. This difference aside, two parallels can still be drawn. One is that the two devices share exactly the same goal of low to medium frequency duct noise abatement, and the other is that both achieve a stopband wider than one octave.

The current paper compares the experimental data of the drumlike silencer, henceforth drum silencer, with the predicted performance. The theory necessary for prediction is outlined very briefly in the next section, followed by the description of the experimental rig. The optimal result obtained in the tests is then analyzed in terms of both transmission loss and energy loss mechanisms. The effects of membrane mass and tension are discussed towards the end.

II. THEORETICAL OUTLINE

As shown in Fig. 1, the theoretical model resembles a standard, two-dimensional expansion chamber with a main

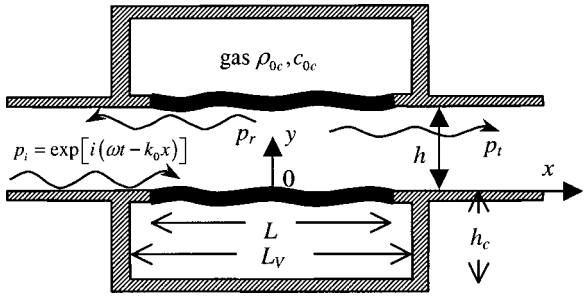


FIG. 1. Theoretical model of sound incident on two identical stretched membranes each backed by a cavity filled with a fluid which may differ from air.

channel height of h, and two identical cavities of length L_v and depth h_c. Two membranes are used to cover the two cavities, and their length can be less than that of the cavities, $L \leqslant L_v$. With the membrane cover, the gas medium inside the cavities can differ from air, and the gas density and speed of sound are denoted as ρ_{0c}, c_{0c}, respectively. The membranes have a mass per unit length m and a tensile force of T is applied. The basic method of transmission loss prediction follows that of an earlier study (Huang, 1999) in which the effect of the cavity air was ignored. In addition, the fluid inside the cavity in the present model is allowed to differ from that of air in the main duct. The need of doing so arises from the experimental finding (see Sec. IV) that there is a substantial amount of sound energy absorption inside the cavity.

When a harmonic sound of time dependence $e^{i\omega t}$ is incident on the membrane, it responds by vibrating at a velocity V which gives rise to a pressure difference Δp, for which the positive value denotes the downward action, hence fluid loading. This loading is borne by the membrane in addition to the incident wave perturbation, $p_i = e^{i(\omega t - k_0 x)}$, where $k_0 = \omega/c_0$ is the wave number and c_0 is the speed of sound in air. The dynamics of the fluid-structural coupling is described by the following equation for the displacement calculated as $\eta = V/(i\omega)$,

$$m\partial^2\eta/\partial t^2 - T(1+i\sigma)\partial^2\eta/\partial x^2 + \Delta p + p_i = 0, \tag{1}$$

where σ is a dimensionless loss factor of the membrane. The choice of this particular form of friction force is based on the following considerations. The usual dash-pot model would result in an inertia term like $m(1-i\sigma)\partial^2\eta/\partial t^2$. This implies that the tensioned membrane experiences friction even when it undergoes a uniform heaving motion, which we argue is not proper. Having said that, the physics of the real damping mechanisms in a thin tensioned membrane is rather complicated and is beyond the scope of current studies. Equation (1) is an *ad hoc* attempt to relate the energy loss with the shearing motion of membrane, and as such, σ is not necessarily identical to the loss factor normally quoted for materials, but the latter can be used as a reference. In fact, there are other acoustic loss mechanisms involved in the experimental rig, and more discussions on the modeling of the lumped energy loss are given later. The main objective of the current model is to provide some measure of accounting for the total vibration energy loss, so that the behavior of the

major part of sound power, which is shown to be reflected, may be better predicted. Comparison with the dash-pot model with the same σ shows little quantitative difference except that the dash-pot model gives very low damping at low frequencies while the current model gives more uniform loss throughout the spectrum.

The Galerkin procedure can be followed to solve Eq. (1) in terms of the modal expansion for a simply supported membrane,

$$V=\sum_{j=1}^{\infty} V_j \sin(j\pi\xi), \quad V_j=2\int_0^1 V\sin(j\pi\xi)d\xi, \quad \xi=x/L+1/2. \tag{2}$$

The main task in doing so is to relate the fluid loading Δp with the vibration velocity V through a modal impedance matrix $\{Z_{jl}\}$ as shown below,

$$\Delta p=\sum_{l=1}^{\infty}\left(\sum_{j=1}^{\infty} V_j Z_{jl}\right)\sin(l\pi\xi), \tag{3}$$

where the single mode impedance Z_{jl} is defined as the lth fluid loading coefficient caused by the vibration of the jth *in vacuo* mode of unit amplitude, $\sin(j\pi\xi)$. Substitution of Eq. (3) into Eq. (1) then forms a set of linear equations for the vibration velocity coefficients V_j, $j=1,2,3,...$, which may be truncated to an order like $j=N=25$, and solved by standard matrix inversion techniques like Gaussian elimination.

The task of finding the modal impedance Z_{jl} is one in which, unlike the coupled membrane dynamics, the membrane vibration is specified, and the loading is customarily called the radiation pressure. The radiation pressure on the membrane surface facing the main duct, denoted as $p_{+\mathrm{rad}}$, can be found by the summation over all duct acoustics modes, ψ_n, as shown below (see, for example, Doak, 1973):

$$p_{+\mathrm{rad}}(x,y,t)=\frac{\rho_0}{2h}\sum_{n=0}^{\infty} c_n\psi_n(y/h)\int_{-L/2}^{+L/2}\psi_n(y'/h)\times V(x',y',t)[H(x-x')e^{-ik_n(x-x')}+H(x'-x)e^{+ik_n(x-x')}]dx',$$
$$\psi_n(y/h)=\sqrt{2-\delta_{0n}}\cos(n\pi y/h), \quad k_n=\frac{\omega}{c_n}, \quad c_n=\frac{ic_0}{\sqrt{(n\pi/k_0h)^2-1}}, \tag{4}$$

where ρ_0 is the air density, c_n and k_n are, respectively, the phase speed and wave number of the nth duct mode, H is the Heaviside function, δ_{0n} is the Kronecker delta. The integration is carried out over the source surface: $x'\in[-L/2,L/2]$, $y'=0$. This result can be easily adapted for finding the radiation pressure in the lower (cavity) channel, denoted as $p_{-\mathrm{rad}}$, when ρ_0 and c_0 are replaced by the corresponding medium properties inside the cavity, ρ_{0c} and c_{0c}, respectively, the vibration velocity V replaced by $-V$, and h replaced by h_c. Sound reflection by the cavity walls at $x=\pm L_v/2$, denoted as $p_{-\mathrm{ref}}$, is also expressed as a summation of duct acoustics modes ψ_n like in Eq. (4) except that the two Heaviside functions are replaced by two constants, A and B,

$$p_{-\mathrm{ref}}(x,y,t)=\frac{\rho_{0c}}{2h_c}\sum_{n=0}^{\infty} c_{nc}\psi_n(y/h_c)\int_{-L/2}^{+L/2}\psi_n(y'/h_c)\times[-V(x',y',t)]\times[Ae^{-ik_{nc}(x-x')}+Be^{+ik_{nc}(x-x')}]dx',$$
$$c_{nc}=\frac{ic_{0c}}{\sqrt{(n\pi/k_{0c}h_c)^2-1}}, \quad k_{nc}=\frac{\omega}{c_{nc}}, \tag{5}$$

which are found by the rigid wall boundary conditions at $x=\pm L_v/2$,

$$\left.\frac{\partial(p_{-\mathrm{rad}}+p_{-\mathrm{ref}})}{\partial x}\right|_{x=\pm L_v/2}=0\rightarrow A=\frac{e^{ik_{nc}(L_v-2x')}+1}{e^{ik_{nc}(2L_v)}-1},$$
$$B=\frac{e^{ik_{nc}(L_v+2x')}+1}{e^{ik_{nc}(2L_v)}-1}. \tag{6}$$

Both ρ_{0c} and c_{0c} can be complex quantities if sound absorption material is used and modelled as an equivalent fluid, in which case both are functions of frequency. In the next section, mechanisms of cavity energy dissipation is modelled by such an approach using the properties of glass-fiber of various porosity.

Once the vibration velocity is found via modal coefficients V_j through solving the coupled dynamics equation, Eq. (1), the reflected wave, denoted as p_r, can be found by evaluating the radiated wave into the far left from Eq. (4), while the transmitted wave, p_t, is found by the superposition of the incident wave, p_i, with the radiation into the far-right. The coefficients of energy flux reflection, β, wave absorption, α, and the transmission loss, TL can be evaluated as follows:

$$p_r=p_{+\mathrm{rad}}|_{x\rightarrow-\infty}, \quad p_t=p_{+\mathrm{rad}}|_{x\rightarrow+\infty}+p_i,$$
$$\beta=\left|\frac{p_r}{p_i}\right|^2, \quad \alpha=1-\beta-\left|\frac{p_t}{p_i}\right|^2, \quad \mathrm{TL}=20\log_{10}\left|\frac{p_i}{p_t}\right|. \tag{7}$$

The theory can be checked by the calculation of sound reflection by a simple expansion chamber without membranes by allowing the membrane mass and tension to vanish. This is described in the next section, and the result is compared with experimental data for the expansion chamber, a step which also helps to check the experimental method.

III. EXPERIMENTAL SETUP

As shown in Fig. 2, the transmission loss across the drum silencer is measured by the four-microphone, two-load method similar to the one described by Munjal and Doige (1990), and the resolution of the standing wave pattern follows the principle described by Chung and Blaser (1980). Briefly, two microphones are required to resolve the two travelling waves in a standing wave pattern. The theoretical definition of transmission loss involves a downstream condition which is anechoic. This is not normally possible without using excessively long duct lining for low frequencies.

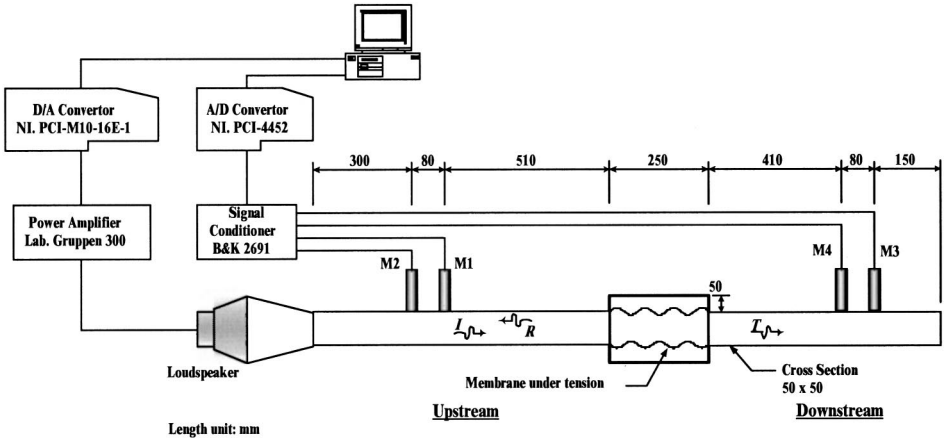


FIG. 2. The four-microphone, two-load measurement system. Both data acquisition and source sound generation are controlled a Labview code in a computer equipped with A/D and D/A cards.

Therefore the wave pattern downstream of the muffler remains a standing wave. A total of four microphones are needed for the simultaneous measurements of the two standing wave patterns, one upstream and one downstream. Even this is insufficient to define the transmission loss since the downstream segment still has reflection wave, say $R_{d1}e^{i(\omega t+kx)}$. It is possible, and indeed necessary, to change the configuration at the downstream without affecting the muffler itself, so that a different standing wave pattern is created and measured giving another residual reflection wave $R_{d2}e^{i(\omega t+kx)}$. The linear combination of the two sets of data forms another set which is acoustically correct for the muffler. The combination can be done in such a way that $R_{d1}+\varphi R_{d2}=0$, i.e., the downstream becomes anechoic. φ is a complex constant which is determined experimentally by this anechoic requirement. The derived data set is non-trivial only when the two actual experiments are linearly independent of each other, i.e., when they are conducted with different downstream impedance conditions. The magnitudes of the incident and transmitted waves in the derived data set can be determined by the same combination constant φ and the transmission loss is found.

The incident noise is simulated by a loudspeaker connected to the duct through a contraction cone. A digital-to-analogue conversion card (PCI-M10-16E-1 from National Instruments) is used to generate the signal which is amplified by one of B&K's amplifiers (Lab Gruppen 300). Two pairs of 1/2 inch, phase-matched microphones (B&K type 4187) are used together with conditioning amplifier (B&K's Nexus 2691). The separation distance for each pair is 8 cm. The large separation is chosen to reduce experimental error at low frequencies up to 1500 Hz. Signals from the four microphones, labeled M1 through M4, are digitized by the National Instruments card PCI-4452 at the sufficient sampling rate of 16 KHz. Both A/D and D/A processes are controlled by National Instruments' Labview program and the test is run by a loop of discrete frequencies from 20 Hz to 1.5 kHz with a 10 Hz increment. The discrete frequency approach is preferred since it offers opportunities to adjust the loudspeaker volume at nodal frequencies where the output can be very low. Since the frequency of the incident wave is accurately controlled by the Labview program, the length of each sample can be tailor-made to be an integer number of cycles, eliminating any possible spectral leakage through the use of digital windowing technique. The program can also be run with random noise measurement, but it is felt that the time required to scan through the whole frequency spectrum is acceptable thanks to the automatic looping function of the Labview code.

The exact locations of the microphones are shown in Fig. 2 with the length unit of mm. Before the two-load tests are conducted, microphones M2, M3, and M4 are calibrated against M1 by the swapping procedure despite the fact that all microphones are very well phase matched. The resultant variations of the amplitude ratio and phase angle difference with respect to frequency are fit to a polynomial curve and saved in files to be read by the Labview code which runs the two-load tests. The first test is conducted with the far end of the downstream section filled with some sound absorption materials providing a partially anechoic termination. The second test is conducted without any sound absorption material. In both tests, the far downstream end is terminated by a rigid plate so that the signal will not be contaminated by the background noise. The rigid end and partially anechoic end conditions provide two linearly independent tests as required by the two-load method.

The duct used in the tests has a square cross section of dimension $h=5$ cm. The two cavities also have the nominal cross section of 5 cm by 5 cm and the length is 25 cm. These dimensions can be adjusted by adding or retrieving some solid blocks in the rig. The first cut-on frequency in the rigid duct is 3430 Hz. According to the prediction, the lower limit of the stopband frequency may be down to a dimensionless frequency of $fh/c=0.05$ or $f=343$ Hz. Such a frequency is not that low in the usual sense, but the general agreement reached between the experimental data and the theoretical prediction gives reasonable confidence that substantial noise reflection at frequencies below 100 Hz can be achieved for a duct with a height of 20 cm or so. The choice of the 5 cm duct in this experiment is partly based on the convenience of an existing rig. In contrast, the choice of the relatively large separation distances between the two junctions of the chamber housing the silencer and the nearest microphones is forced upon by the inconvenience of accommodating the external device required to apply and change the membrane tension. Account is taken of the acoustic damping by the duct walls according to the well-known formula of complex wave numbers (see Pierce 1991, Eqs. 10-5.10 and 10-5.8):

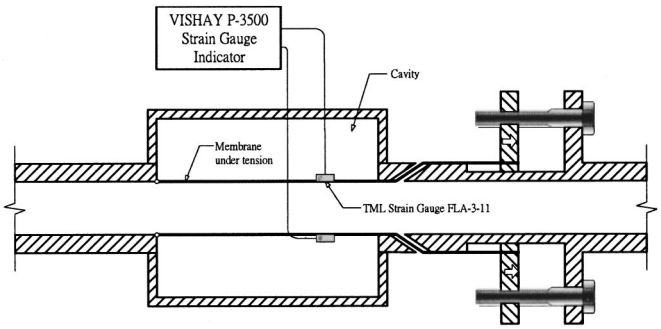


FIG. 3. The mechanism to apply and measure the membrane tension.

$$k=\frac{\omega}{c_0}+(1+i)\alpha_{\text{walls}},$$
$$\alpha_{\text{walls}}=2^{-3/2}\sqrt{\frac{\omega\mu}{\rho_0 c_0^2}}\left[1+\frac{\gamma-1}{\text{Pr}^{1/2}}\right]\frac{L_p}{A}, \tag{8}$$

where μ is the air viscosity, γ is the ratio of specific heats, Pr the Prandtl number, L_p the perimeter, and A the cross sectional area. The energy loss has a square-root frequency dependency. Take 500 Hz as an example, friction for the present duct over 1 m gives around 5% loss in sound energy flux. Corrections for travelling wave amplitudes are made by finding out the true incident and reflection waves at the left-hand side junction of the chamber from the sound intensities measured between M1 and M2. The same applies to the downstream. But the effect of friction inside the chamber is regarded as part of the damping mechanism of the silencer. No attempt is made to separate friction from other mechanisms such as the inevitable vibration of the so-called rigid walls of the cavity.

The membranes are clamped and stretched by a mechanism shown in Fig. 3. The upstream end of the membrane is simply fixed by wrapping it around a cylinder of 6 mm in diameter and pressed against the junction walls. The downstream end of the membrane runs through a 0.05 mm slit, which is exaggerated in the figure, to be fixed by two plates outside the duct. The axial positions of these clamp plates, hence the strain in the membrane, are easily adjusted by the screws shown in the figure. The tensile force applied is measured by a strain gauge glued to the surface of the membrane in the test section. The dimension of the strain gauge sensor (TML FLA-3-11) is 3 mm and its attachment is believed to cause no noticeable influence on the dynamics of the stretched membrane. However, since the membrane used is thinner than the epoxy used to attach the strain gauge, significant measurement uncertainties may occur. To eradicate this problem, a calibration procedure is undertaken by attaching dead weights to the membrane through some connection joints in the upright position.

In the direction perpendicular to the figure, say, z, the two edges of the membranes (of width 52 mm) are inserted into a very thin gap between the two constituent plates of the cavity walls. The gap is less than 0.5 mm and has a depth around 1–2 mm in the z direction. This is a deliberate attempt to simulate the two-dimensional theoretical model shown in Fig. 1. Tests were also conducted using a membrane narrower than 50 mm leaving a small gap between the membrane edges and the end walls, but results were quite similar. It is anticipated that, in real applications, the edges of tensioned membranes should be entirely glued to the end walls, for which a three-dimensional theoretical model should be developed.

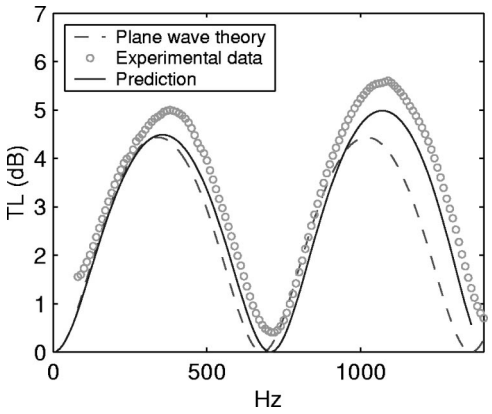


FIG. 4. Comparison of the measured data with predictions by the plane wave approximation and the modal expansion calculation with zero membrane mass and tension.

When the membranes are removed and the slit allowing the membranes to get out of the duct is covered, the rig becomes a simple expansion chamber with an area expansion ratio of 3. The transmission loss is measured and compared with two theoretical predictions in Fig. 4. One is the plane-wave theory (dashed line), and the other is the present theory using 25 modes of the virtual membranes (solid line). It is found that the comparison between the experimental data with the present theory is better than that with the plane-wave approximation. The latter ignores the acoustic scattering at the junctions leading to two characteristic differences with reality. In reality, the frequency range of the lobes of the transmission loss is larger than $c_0/2L$, and the peak transmission loss is higher than that determined by the plane-wave theory. These are all captured correctly by the present theory. Having said that, the measured transmission loss is typically 0.5 dB higher than the theoretical prediction. This is probably caused by some energy dissipation mechanisms which are necessarily present in the rig but excluded in the theory. Apart from the damping by the sharp edges, such as the 6 mm cylinder used to clamp the membranes, one possible mechanism is the vibration of many constituent parts of the rig which is required to adjust the cavity size during the experimental study. When the same measurement system and procedure were used earlier in a single-piece circular expansion chamber, the discrepancy between the experimental data and the prediction using a finite element calculation was reduced to 0.2 dB at the nodal frequencies of the transmission loss lobes. Despite the finite discrepancy between the prediction and the experimental data, the test of the expansion chamber serves as a mini validation for both experimental setup and the theory.

IV. BASIC ANALYSIS

According to the theoretical prediction, the membrane mass has adverse effect on the silencer performance. Two

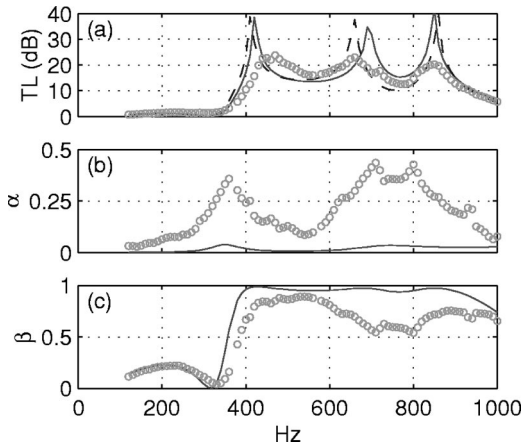


FIG. 5. Comparison of experimental data with theoretical predictions for (a) transmission loss TL, (b) absorption coefficient α, and (c) reflection coefficient β. The experimental data (open circles) was obtained for the thin stainless steel under 190N tension. The predictions were for the same membranes with a loss factor of $\sigma=0.5\%$ under tension 190N (solid lines) and 181N [dashed line in (a)].

types of metal foils were available at the time of experiment: stainless steel and copper. The stainless steel foils were 0.01 mm and 0.02 mm thick. The membrane is weighed, and the mass ratio for the thinner one is calculated as follows:

$$m_1=\frac{m}{\rho_0 h}=\frac{7800\times 0.01}{1.225\times 50}=1.3. \tag{9}$$

This is sufficiently low although even lower values can be achieved by using aluminum foil in a larger duct. This section focuses on the best result using this foil, leaving the results for heavier membranes and the effect of tensile force to the next section.

Here, the stopband is defined as the frequency range, $f\in[f_1,f_2]$, in which the transmission loss is everywhere equal to or higher than 10 dB, a level which is much higher than the peak performance of around 5 dB for an empty expansion chamber with an area ratio of 3. Based on the mass ratio of $m_1=1.3$, and assuming a membrane loss factor of $\sigma=0.5\%$ for stainless steel, an optimization calculation can be conducted to see what tension T would yield the highest bandwidth represented by f_2/f_1. The result is shown below in dimensionless form first,

$$\begin{aligned} &f_1 h/c_0=0.0548, \quad f_2 h/c_0=0.1378,\\ &f_2/f_1=2.516, \quad T_{\mathrm{opt}}/(\rho_0 c_0^2 h^2)=0.502. \end{aligned} \tag{10}$$

The predicted dimensional optimal tension, T_{opt} and band frequencies for the particular rig are given below,

$$f_1=377\ \mathrm{Hz}, \quad f_2=947\ \mathrm{Hz},$$

$$T_{\mathrm{opt}}=0.502\times 1.225\times(343\times 0.05)^2=181\mathrm{N}.$$

This prediction is shown in Fig. 5(a) as the dashed line. Three peaks are found and the dip between the second and third peaks is right on the critical level of 10 dB. Study of the variation around this tension shows that the dip goes below 10 dB for a slightly lower tension, indicating that the tension does play a very subtle role in setting the frequency intervals between the spectral peaks. The dip remains above 10 dB for higher tension but the stopband of TL$>$10 dB is slightly narrower. In the experiments, tensions were applied with an increment of 10N, and the stopband of 190N was found to be wider than that of 180N. So, the experimental data of 190N is chosen for analysis. In Fig. 5, the experimental data is shown in small open circles and the prediction for 190N is shown as solid curves.

As shown in Fig. 5(a), the predicted stopband is $f\in[385,953]$ Hz, while the experimental result is $f\in[398,935]$ Hz. The ratio of frequency limits, f_2/f_1, are 2.47 and 2.35, respectively. Figure 5(b) compares the sound absorption coefficient, α, and Fig. 5(c) shows the reflection coefficient of sound energy flux, β. The following observations are made.

(1) In agreement with the prediction, there are three distinct spectral peaks found within the experimental stopband. This means that the basic theoretical model is correct for the experimental rig.
(2) The predicted TL curve has sharp peaks while that from the experiment is smoother. It means that there are significant energy dissipation mechanisms in the rig which are excluded from the theory. The experimental value of α ranges from 10% to 45% within the stopband, which is roughly 10 times higher than the predicted value with $\sigma=0.5\%$.
(3) Also in agreement with theory, the experimental data shows that the mechanism of sound reflection dominates over sound absorption in the whole experimental stopband. As shown in Fig. 5(c), the prediction agrees with the experimental data almost perfectly below 300 Hz, which constituents the first small lobe.

In order to make a proper account of the sound energy damping in the rig, four possible mechanisms are suggested,

(1) membrane damping over the bulk vibrating surface;
(2) membrane damping at the edges;
(3) cavity damping; and
(4) sound induced vibration of some walls during the low frequency test.

Since the membrane is so thin and the dynamics of air and structure are fully coupled, it is impossible to separate mechanisms (1) from (3). The identifications of mechanisms of (2) and (4) are also difficult. Two models are proposed below, one for mechanisms (1) and (2), and another for (3) and (4).

It might be assumed that mechanism (2) dominates over mechanism (1). The losses concentrated at the leading and trailing edges of the membrane should be described by delta functions, $\delta(|x\pm L/2|)$, but the fact that there are two other edges in the membrane length direction gives some hope that a model of distributed energy loss might at least capture some essential characteristics of mechanisms (1) and (2). This is done by assigning an equivalent material loss factor σ which should be much higher than 0.5%, the typical value listed in material handbooks for metals. This model is called the σ model.

Cavity damping, mechanism (3), is again hard to model

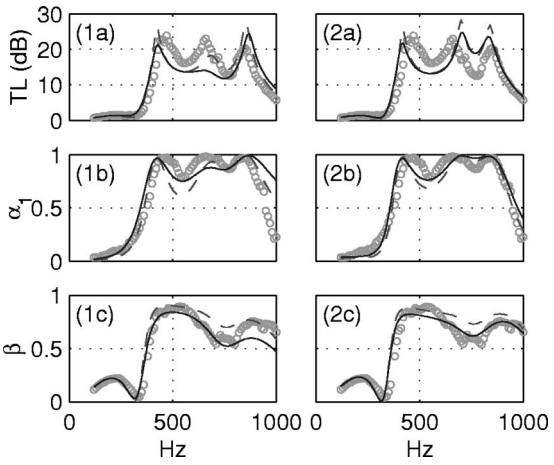


FIG. 6. Modeling of damping mechanisms by two methods: equivalent membrane loss factor (1a), (1b), and (1c), and equivalent sound absorption material in the cavities (2a), (2b), and (2c). The experimental data are shown in open circles, and two levels of damping mechanisms are calculated in each model. In the left column, σ is 5% for the dashed lines, and 10% for the solid lines. In the right column, R_{sam} is 40 kg/m^3 s for the dashed lines and 80 kg/m^3 s for the solid lines.

exactly. Usually, a loss coefficient is assigned to each cavity mode. This procedure becomes awkward in the present model using the expansion of membrane vibration in terms of the *in vacuo* modes of the simply supported membrane. It is assumed that the combined effects of mechanisms (3) and (4) may be modeled by treating the empty cavity as one filled with a very fluffy sound absorption material. This is done by assigning a small value of flow resistivity R_{sam} and the model is labeled as the sam model. Once R_{sam} is given, the complex density and speed of sound in the cavity, ρ_{0c}, c_{0c}, can be specified as functions of frequency and flow resistivity (or porosity) using empirical properties of common materials like fiberglass available in Mechel and Vér (1992).

Going back to Figs. 5(b) and (c), it is observed that high sound absorption rate occurs when the reflection coefficient is low. A more useful comparative study can be made by the relative sound absorption coefficient

$$\alpha_1 = \alpha/(1-\beta), \tag{11}$$

which is a more sensible measure of how much noise that actually enters the silencer is absorbed. Using this concept, the effects of σ and R_{sam} are studied separately, and the results are shown in Fig. 6. The subfigures on the left-hand-side columns are for the σ model, while those on the right-hand side are for the sam model. For each model, two levels of losses are tested. The dashed lines are for the lower levels and the solid lines higher values. The two levels are chosen in such a way that the experimental data for 190N (open circles) generally fall within the range predicted by the two levels with a heuristic curve-fitting. As shown in Fig. 6(1b), the comparison between the experimental data and the result of the level 1 loss factor σ=5% shows a large deficit of relative damping (α_1) around 540 Hz. This discrepancy is reduced by the level 2 model with $\sigma=10\%$. But, as shown in Fig. 6(1a), the transmission loss of the two predicted falls below the experimental value due to the inability to simultaneously predict the correct reflection coefficient in this region, as shown in Fig. 6(1c). The modelling of the spectral dip around 775 Hz and the third peak around 855 Hz is more successful, but that for the second peak around 670 Hz is not. The transmission loss for $\sigma=5\%$ retains a distinct peak, but it overpredicts the reflection. The model of $\sigma=10\%$, on the other hand, simulates β correctly but the TL peak disappears due to the lack of sufficient membrane response. However, assigning a lower value of σ for the first mode does not yield a better overall agreement with the experimental data. These tests, together with the fact that the frequencies of TL peaks are not predicted accurately, seems to suggest that some three-dimensional effects inherent in the experiment is responsible for the discrepancy between the data and the two-dimensional theoretical model. These include the possible corrugation of the membrane when it's slightly twisted during the application of tensile force, nonuniformity of the tensile force across the width in the z direction, and the possible effect of the gap between the membrane edges and the duct walls along the axial direction, etc. There are also other factors in the rig which are normally excluded in any theoretical model. For example, the effect of the attachment of the strain gauge is believed to be minor but, admittedly, it is hard to assess this accurately.

The sam-model based on the fiberglass properties is tested for two levels of flow resistivity, $R_{\text{sam}}=40$, 80 kg/m^3 s, which are about three orders of magnitude lower than that used in usual duct linings. The results are shown in the right-hand-side column of Fig. 6. Figure 6(2b) shows that the level 1 model underpredicts the loss around the first spectral dip at 540 Hz but, contrary to the σ models, sam models of both levels overpredict the loss around the second spectral peak at 685 Hz. The match between the sam model and the experimental data is very good beyond the third spectral peak for both α_1 and β. Figure 6(2a) shows that three spectral peaks are simulated with satisfactory sharpness although, like the essentially lossless model shown in Fig. 5(a), the interval between the second and third peaks is narrower than the experimental result. The overall conclusion is that the sam model simulates better than the σ model, and an intermediate level of $R_{\text{sam}}=60$ kg/m^3 s will be used with the basic material loss factor of σ=0.5% in all subsequent comparative studies between theory and experiment.

V. EFFECT OF MEMBRANE PROPERTIES

The variation of transmission loss spectrum with respect to the membrane tension is shown in Fig. 7 for the thin stainless steel foil ($m_1=1.3$). Four spectra for 110, 180, 190, and 200N are chosen for display. The experimental results (open circles) are compared with the theoretical modeling (solid lines) with $R_{\text{sam}}=60$ kg/m^3 s, and $\sigma=0.5\%$. Figure 7(a) is for the tension of 110N. This is a typical situation where there are only two peaks in the stopband since the third one is too far away from the second. The theoretical modelling is seen to be rather successful except that the heights of all the peaks are over-estimated. Figure 7(b) is for 180N, which is shown earlier to be close to the theoretical optimal value of 181N. Although the second peak is rather close to the third, the breakdown between the two peaks is still too serious to qualify for a broadband performance un-

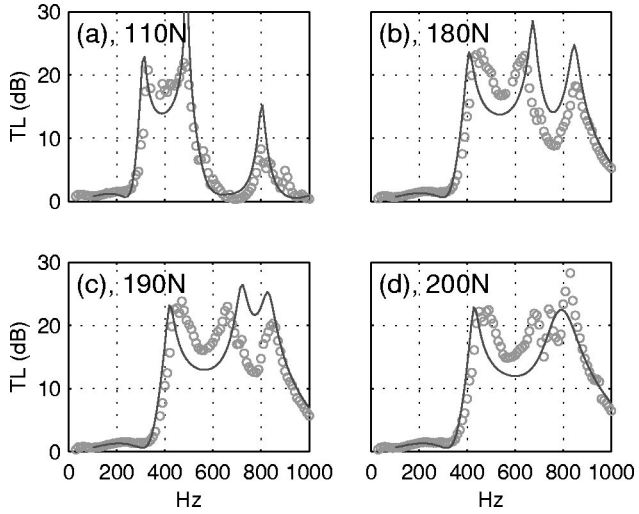


FIG. 7. Variation of TL spectrum with respect to tension for the thin stainless steel foil of $m_1=1.3$. The values of tension are labeled in subfigures. Open circles are the experimental data, and the solid lines are predictions incorporating a cavity damping model with $R_{\text{sam}}=60\ \text{kg/m}^3\ \text{s}$ and $\sigma=0.5\%$.

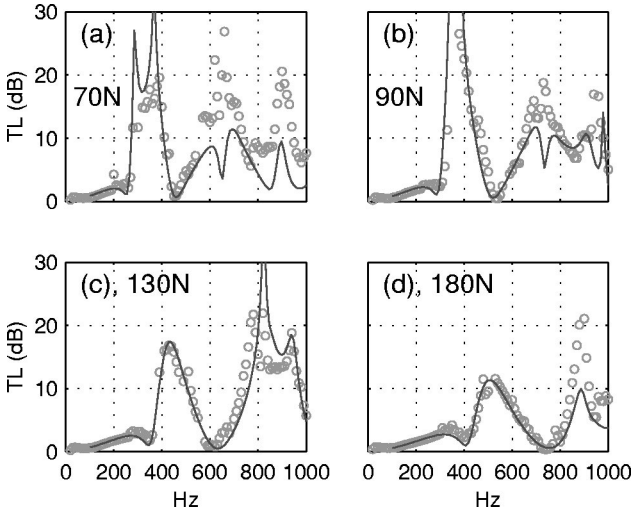


FIG. 8. The performance of a drum silencer made of copper foils of 0.03 mm thick and 198 mm long, with a 65 mm deep cavity. The open circles are the experimental data, and the solid curves are predictions using a cavity damping model with $R_{\text{sam}}=60\ \text{kg/m}^3\ \text{s}$ and $\sigma=0.5\%$.

der the strict criteria of TL$\geqslant$10 dB at all discrete frequencies. If one third octave band averaging is applied, as is done in most practical applications, this breakdown can be overlooked. The point to be emphasized here is that the spectral variation with respect to T is gradual in every sense, and the sudden change of the bandwidth with T is purely caused by the rigid definition of the stopband. The modelling in this case can also be regarded as being successful given the fact that the peak frequencies cannot really be determined accurately with a two-dimensional theory. Figure 7(c) is for 190N and is analyzed in detail earlier. Comparing with the experimental spectrum in Fig. 7(b), the interval between the second and the third peaks of the experimental data is reduced, leading to a more balanced picture with a wider stopband. This change of pattern is captured rather accurately despite the two-dimensional approximation. Figure 7(d) is for 200N in which the second and third peaks from the experiment seem to merge together. The merge is more complete in the theoretical prediction in this case, consistent with the overprediction of the second peak frequency in all previous cases.

Experiments using copper foil of thickness 0.03 mm were also conducted with less impressive results. This outcome is expected, but the results are also analyzed here to confirm the theoretical prediction of adverse effects of membrane mass. For this foil, the cavity was changed to a shorter and deeper one following the theoretical search for a reasonable performance for the heavy foil. The parameters are $h_c=65$ mm, $L_v=L=198$ mm, $m_1=4.43$. The results are compared with theoretical modeling in Fig. 8. Figure 8(a) is for the tensile force of 70N, and the data seems to be rather scattered relative to the theoretical modeling. The same phenomena are observed for tests with lower tension values such as 30N and 50N. This was found to be caused by the difficulty in setting up a uniform stress for the foil at such low tension. The agreement between theory and experiment becomes much improved for tests with higher tensile forces of 90N, 130N and 180N in Figs. 8(b), (c), and (d), respectively. Comparing the results from the lighter foil of stainless steel in Fig. 7, it is found that the heavier copper foil yields stopband in lower frequencies but the bandwidth is also narrower.

Since noise abatement at extremely low frequencies is difficult, it might be reasonable to adopt a performance criterion which varies somewhat with frequency. In that case, heavy metal foils with a medium stress level will be the most suitable choice. To demonstrate the effect of mass on spectrum, the typical result of the thin stainless steel foil shown in Fig. 7(c) is plotted together with that of the thick copper foil shown in Fig. 8(a) for comparison. These are presented as Figs. 9(a) and (c), respectively. A thicker stainless steel

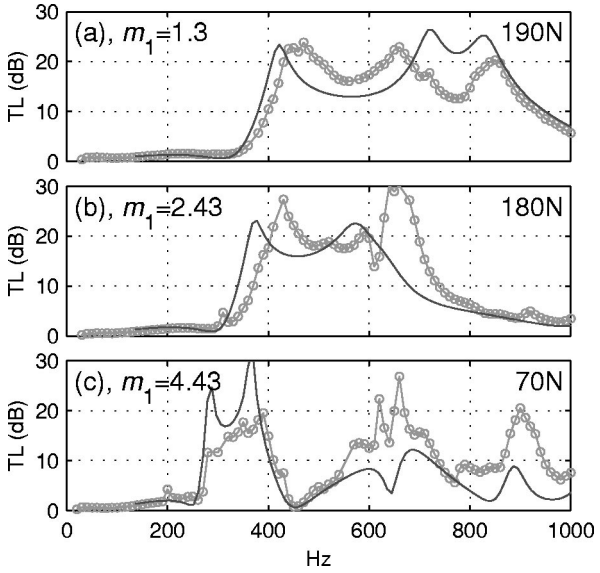


FIG. 9. Comparison of stopbands for three different values of membrane mass which is marked in each subfigure. (a) is the thin stainless steel with $L_v=L=25$ cm, $h_c=5$ cm. (b) is for a thicker stainless steel foil with a cavity length $L_v=30$ cm which is longer than the membrane length $L=25$ cm, while $h_c=5$ cm. (c) is copper foil with $L_v=L=198$ mm, $h_c=65$ mm.

foil was also tested. Its mass ratio of 2.43 provide the intermediate value, so the result is shown in Fig. 9(b). The general trend of spectral shifting towards lower frequencies is made clear. Notice that, for the thicker stainless steel foil shown in Fig. 9(b), the stopband is underpredicted, contrasting with all other comparisons.

VI. DISCUSSIONS

Several issues related to the actual implementation of the concept of drum silencer are discussed here based on what is learned from the two-dimensional modelling. The first issue is the significance of possible molecular relaxation of the stretched membrane after a period of use. For this purpose, the last test with the thicker stainless steel, shown in Fig. 9(b), was left in the laboratory for more than a week, and the reading for the tensile force was reduced from 180N to 178N, which is insignificant considering the accuracy of the strain gauge itself. The transmission loss was measured again and there was hardly any noticeable change. Strictly speaking, this test does not really simulate the actual situation where the silencer might be exposed to turbulent flow. For the moment, it can be said that the change of transmission loss spectrum with possible relaxation in tensile force is at least very gradual. The other related issue is acoustic fatigue. The stress level for this particular rig is around 380 MPa, which falls within the range of fatigue stress of carbon steels. In order to make sure that no fatigue occurs, good quality material should be chosen. But in actual application with larger duct, thicker membranes can be used, and the stress level comes down even with the same tensile force, so the choice of material should be plenty.

The second issue is the possibility of flow-induced vibration, and secondary sound radiation caused by this vibration. This issue should be divided into two parts. The first is the possibility of sound radiation by forced membrane vibrations due to turbulent flow eddies. The second is the possibility of self-induced vibration which would lead to a possible breakdown of the membrane, a problem much more serious than the noise itself. For the first part, reference can be made to the work of Ffowcs, Williams, and Hill (1987), who treated the problem of scattering of frozen, convective wave pattern within the boundary layer flow by abrupt changes in structural properties, which in the drum silencer is from the rigid wall to the tensioned membrane. Considering the small length scale of the turbulent eddies, see Howe and Baumann (1992), such transfer of energy from the near field to the far field is most likely to concentrate in the region of high frequencies. This possible problem of high frequency radiation is much less important than the low-frequency noise being addressed in the present study. For the second part, extensive studies have been carried out by the second author on this topic (Huang, 1998, 2001), and it can be concluded that it is unlikely that the membrane would experience any form of flutter (transfer of mean flow energy to growing structural vibration) when the flow speed is well below the *in-vacuo* tensile wave speed in the membrane. The latter is very high for the drum silencer studied here. Take the thin stainless membrane as an example,

$$c_T = \sqrt{\frac{T}{mh}} = \sqrt{\frac{190}{7800 \times (0.01 \times 10^{-3}) \times 0.05}} = 221 \text{ m/s}.$$

This speed is unlikely to be approached in ordinary ventilation systems. If the flow speed does approach such value, the acoustic impedance of the system would have been changed drastically. The tensile force of 190N may no longer be the optimal tension, and the whole device has to be redesigned. The problem of self-induced vibration should be more seriously considered if the drum silencer is to be used in a high-speed wind-tunnel.

The third issue is the effect of three-dimensional features in any realistic design of drum silencer. Most likely the membranes will have to be fixed on all the edges. If no tension is applied in the third dimension *z*, most conclusions reached in the two dimensional model might be preserved although the membrane motion would become considerably more complicated. Three-dimensional modelling will have to be developed to predict and design such realistic drum silencers. Further investigation is also needed to better identify and model the damping mechanisms.

VII. CONCLUSIONS

The general conclusion of this study is that the theoretical predictions of the broadband performance of the so-called drum silencer are validated by experimental data despite two major limitations of the theory. One is the two-dimensional nature of the theoretical model, and the other is the exclusion of the damping mechanisms which are difficult to describe quantitatively. The following are specific conclusions derived from various tests conducted.

(1) The frequency range. Due to the small duct size, the dimensional frequencies do not fall in the low frequency range, say below 200 Hz, in engineers' common sense, but the dimensionless frequency range does agree with the prediction. This gives the confidence that low-frequency broadband noise can be tackled by the new concept of drum silencer. The main feature of the broadband performance is that three peaks appear with suitable intervals such that the breakdown between adjacent peaks remain moderate.

(2) Mechanisms of energy dissipation. In the specific design of the validation experiments, there exists a significant source of energy dissipation at frequencies when reflection is less effective. The exact nature of such damping mechanisms has not been identified convincingly, but it appears from the limited attempts of modelling that the cavity damping is dominant. The other major source of damping may arise from the clamped membrane edges, and the two nominally free edges in the third dimension. When membrane damping is introduced in a distributed manner, the amount of damping required to match the measured overall sound absorption coefficient suppresses the second spectral peak which is definitely present in the experiment. Simulation of cavity damping by fluffy sound absorption material proves to be capable of preserving all three peaks.

(3) Dominance of reflection. Sound reflection dominates in every part of the stopband despite the fact that the dimensional frequency encountered in the small duct rig is relatively high. It is speculated that the significance of the unaccounted contribution from damping will be reduced if a larger duct is used.
(4) Effect of membrane mass. A light membrane is shown, both theoretically and experimentally, to yield better performance than its heavier counterparts, membrane of zero mass being the theoretical optimum. However, according to the present theory, an aluminum foil of thickness 0.05 mm or more can still provide a stopband wider than one octave in a duct of 20 cm in cross sectional dimension. This excludes the potential technical problems that might be encountered by using ultrathin membranes.
(5) Optimal membrane tension. The optimal tension is identified theoretically and proved rather accurately in the tests. The problem of possible material relaxation of stretched membrane was only investigated briefly by measuring the same rig again after one week. No detectable change in tension nor performance was found in this test. However, this does not exclude the possibility of membrane relaxation in all practical situations. At this junction, it is stressed that the actual performance of the drum silencer does not change with tension in any abrupt manner although the level of tension does play a delicate role of regulating the intervals between spectral peaks.
(6) Multiple drum silencer. Since the stopband frequencies can be manipulated by many variables such as mass and tension, it is theoretically possible to come up with a very broadband design with the sequential combination of drum silencers of different membrane properties. The most profitable design would be the combination of many narrower stopbands in the very low frequency range while that in the medium to high frequency range is better taken care of by normal duct lining covered by light membranes under low tension.
(7) Self-induced vibration and secondary radiation of noise. With a brief reference to the results of related studies, it was shown that a drum silencer is unlikely to experience self-induced vibration because the optimal tensile stress required should be normally very high. Of course, the level of stress should still be within the elastic limits of normal materials like steel or aluminum. The *in vacuo* flexural wave speed far exceeds the usual flow speed in a normal ventilation tunnel, hence no flutter.

ACKNOWLEDGMENT

The first author thanks the Hong Kong Polytechnic University for the Ph.D. studentship.

Ackermann, U., and Fuchs, H. V. (**1989**). ''Noise reduction in an exhaust stack of a papermill,'' Noise Control Eng. J. **33**, 57–60.
Brown, S. (**1964**). ''Acoustic design of broadcasting studios,'' J. Sound Vib. **1**, 239–257.
Brüel, P. V. (**1951**). *Sound Insulation and Room Acoustics* (Chapman and Hall, London).
Choy, Y. S. (**2001**). ''Experimental study of absorption and reflection of grazing sound by flexible panels,'' Proceedings of the 8th International Congress on Sound Vibration, Hong Kong, China, pp. 1003–1010.
Chung, J. Y., and Blaser, D. A. (**1980**). ''Transfer function method of measuring in-duct acoustic properties. I. Theory,'' J. Acoust. Soc. Am. **68**, 907–913.
Doak, P. E. (**1973**). ''Excitation, transmission and radiation of sound from source distributions in hard-walled ducts of finite length (I): the effects of duct cross-section geometry and source distribution space-time pattern,'' J. Sound Vib. **31**, 1–72.
Ford, R. D., and McCormick, M. A. (**1969**). ''Panel sound absorbers,'' J. Sound Vib. **10**, 411–423.
Frommhold, W., Fuchs, H. V., and Sheng, S. (**1994**). ''Acoustic performance of membrane absorbers,'' J. Sound Vib. **170**, 621–636.
Fuchs, H. V. (**2001**). ''Alternative fiberless absorbers—new tools and materials for noise control and acoustic comfort,'' Acustica **87**, 414–422.
Horoshenkov, K. V., and Sakagami, K. (**2001**). ''A method to calculate the acoustic response of a thin, baffled, simply supported poroelastic plate,'' J. Acoust. Soc. Am. **110**, 904–917.
Ffowcs Williams, J. E., and Hill, D. C. (**1987**). ''On the scattering of evanescent waves into sound,'' J. Fluid Mech. **184**, 101–121.
Howe, M. S., and Baumann, H. D. (**1992**). ''Noise of gas flows,'' in *Noise and Vibration Control Engineering: Principles and Applications*, edited by L. L. Beranek and I. L. Vér (Wiley, New York), Chap. 14.
Huang, L. (**1998**). ''Reversal of the Bernoulli effect and channel flutter,'' J. Fluids Struct. **12**, 131–151.
Huang, L. (**1999**). ''A theoretical study of duct noise control by flexible panels,'' J. Acoust. Soc. Am. **106**, 1801–1809.
Huang, L. (**2000**). ''A theory of reactive control of low-frequency duct noise,'' J. Sound Vib. **238**, 575–594.
Huang, L. (**2001**). ''Viscous flutter of a finite elastic membrane in Poiseuille flow,'' J. Fluids Struct. **15**, 1061–1088.
Kang, J., and Fuchs, H. V. (**1999**). ''Predicting the absorption of open weave textiles and micro-perforated membranes backed by an air space,'' J. Sound Vib. **220**, 905–920.
Maa, D.-Y. (**1975**). ''Theory and design of microperforated panel sound absorbing constructions,'' Sci. Sin. **18**, 55–71.
Maa, D.-Y. (**1998**). ''Potential of microperforated panel absorber,'' J. Acoust. Soc. Am. **104**, 2861–2866.
Mechel, F. P., and Vér, I. L. (**1992**). ''Sound absorbing materials and sound absorber,'' in *Noise and Vibration Control Engineering: Principles and Applications*, edited by L. L. Beranek and I. L. Vér (Wiley, New York), Chap. 8.
Munjal, M. L., and Doige, A. G. (**1990**). ''Theory of a two source-location method for direct experimental evaluation of the four-pole parameters of an aeroacoustic element,'' J. Sound Vib. **141**, 323–333.
Pierce, A. D. (**1991**). *Acoustics, An Introduction to its Physical Principles and Applications* (Acoustical Society of America, New York).
Sakagami, K., Kiyama, M., Morimoto, M., and Takahashi, D. (**1996**). ''Sound absorption of a cavity-backed membrane: a step towards design method for membrane-type absorber,'' Appl. Acoust. **49**, 237–247.
Zha, X., Fuchs, H. V., and Drotleff, H. (**2002**). ''Improving the acoustic working conditions for musicians in small spaces,'' Appl. Acoust. **63**, 203–221.

A jet engine noise measurement and prediction tool

Abdelkader Frendi[a)]
Mechanical and Aerospace Engineering Department, University of Alabama in Huntsville, Huntsville, Alabama 35899

Wade D. Dorland and Thein Maung
AI Signal Research, Inc., Huntsville, Alabama 35805

Tom Nesman and Ten-See Wang
NASA Marshall Space Flight Center, Huntsville, Alabama 35812

(Received 8 October 2001; revised 14 June 2002; accepted 22 July 2002)

In this paper, the authors describe an innovative jet engine noise measurement and prediction tool. The tool measures sound-pressure levels and frequency spectra in the far field. In addition, the tool provides predicted results while the measurements are being made. The predictions are based on an existing computational fluid dynamics database coupled to an empirical acoustic radiation model based on the far-field approximation to the Lighthill acoustic analogy. Preliminary tests of this acoustic measurement and prediction tool produced very encouraging results. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1506690]

PACS numbers: 43.58.Kr, 43.58.Ta [SLE]

I. INTRODUCTION

In the last 20 years, tremendous progress has been made in computational fluid dynamics (CFD). This progress could not have been achieved without the quantum leap in computer power. In addition, computational algorithms have become very efficient and robust. However, with all these advancements, CFD has remained a tool used by engineers and scientists mainly to understand experimental observations. In recent years, however, there have been few cases where CFD was used extensively in the research and development phase.[1] This gave CFD a vote of confidence from the engineering community, which in itself is a great achievement. However, using CFD for real-time predictions has not been attempted because of its slow turnaround time. One way to overcome this delay is to create a large CFD database with various parameters.

In this paper, we present a detailed description of a measurement and prediction tool that can be used during a real-time experiment. This tool uses a CFD database, made up of results from various Reynolds-averaged-Navier–Stokes computations, a semiempirical far-field acoustic model, and the experimental results being measured. Results, both measured and predicted, are displayed on the same screen in real time. The semiempirical far-field acoustic model[2] is derived from the Lighthill acoustic analogy.[3] This equation gives the sound pressure at a far-field point as radiated by a localized unsteady source or turbulent flow. Ribner[4] used the Lighthill acoustic analogy to arrive at a relation between the mean-square pressure radiated by a jet plume into the far field as the integral over the plume volume of quadrupole correlations. Ribner[5,6] also showed the importance of convection and refraction on the far-field jet noise. Several other significant contributions to jet noise theory have been made by various researchers and are summarized in the review papers by Ffowcs Williams,[7,8] Goldstein,[9] and Tam.[10]

Mani *et al.*[11] used Ribner[4] and Ffowcs Williams[12] jet noise models in their extensive studies on high-velocity jet noise. Good agreement was obtained between the model and the experimental results for various nozzle geometries and flow regimes. More recently, Khavaran *et al.*[13] used a $(k-\varepsilon)$ turbulence model to compute the jet mixing noise from a supersonic axisymmetric converging–diverging nozzle. They used Mani *et al.*'s[11] approach to arrive at their results. Comparisons with experiments showed good agreement. Other computational studies using the $(k-\varepsilon)$ turbulence model were carried out (for example, Refs. 14 and 15) with varying degrees of success.

Recently, Tam *et al.*[16] proposed a self-contained, semi-empirical jet noise theory for the prediction of fine-scale turbulence noise from high-speed jets. This theory also uses jet flow turbulence information supplied by a $(k-\varepsilon)$ turbulence model. In addition to the empirical constants found in the turbulence model, Tam *et al.*'s jet noise model introduces three additional empirical constants that need to be fixed using experimental data. The resulting noise predictions exhibited excellent agreement with measurements. This theoretical approach uses linearized Euler equations to determine the far-field acoustic pressure, which is a departure from the widely used Lighthill acoustic analogy. However, Tam *et al.*'s model cannot be used for real-time predictions because of its computational cost and time requirements.

In the remainder of the paper, the authors describe the semiempirical acoustic model used, CFD database results, experimental measurements, and, finally, the acoustic model validation. We close the paper with a description of the integrated measurement and prediction tool and related conclusions.

[a)]Author to whom correspondence should be addressed. Electronic mail: frendi@mae.uah.edu

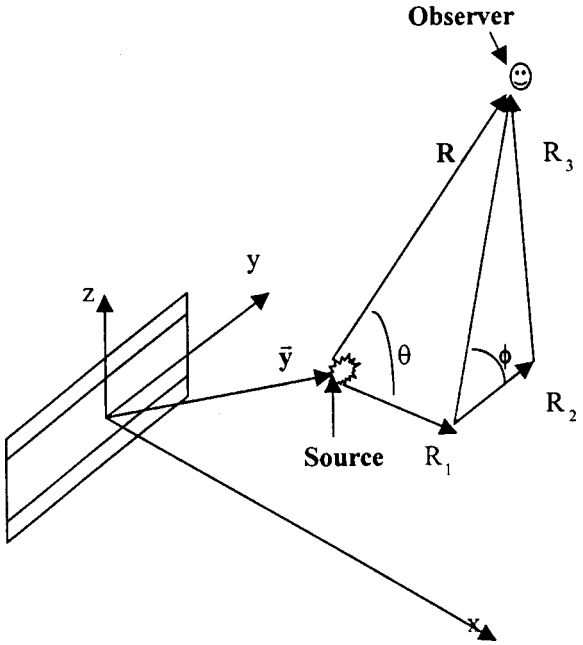


FIG. 1. Reference coordinate system.

II. THE SEMIEMPIRICAL MODEL

To formulate the mathematical model, we begin with the far-field approximation of the Lighthill acoustic analogy,[3] which can be written as

$$p(\mathbf{R},t)=\frac{R_iR_j}{4\pi C_\infty^2R^3}\int_V\left[\frac{\partial^2T_{ij}}{\partial t^2}\right]d^3\mathbf{y}, \tag{1}$$

where T_{ij} is the Lighthill stress tensor given by

$$T_{ij}=\rho v_iv_j+\tau_{ij}+(\pi-C_\infty^2\rho)\,\delta_{ij}\,. \tag{2}$$

In Eq. (2), the dominant term is the unsteady momentum flux ρv_iv_j, τ_{ij} is the viscous compressive stress tensor, π the local pressure, ρ the density, C_∞ the ambient, adiabatic speed of sound, v_i the velocity, and δ_{ij} is the Kronecker symbol ($=0$ if $i\neq j$, $=1$ if $i=j$). The notation "[ ]" in Eq. (1) denotes evaluation at the retarded time, $\tau=t-\mathbf{R}/C_\infty$. On retaining the dominant term, i.e., ρv_iv_j, the mean-square pressure $\overline{p^2}$ radiated at an observer point $(\mathbf{R},\theta,\phi)$ in polar coordinates (see Fig. 1) can be written as

$$\overline{p^2}(\mathbf{R},\theta,\phi)=\frac{R_iR_jR_kR_l}{16\pi^2C_\infty^4R^6}\int_V\int_V\overline{\frac{\partial^2(\rho v_iv_j)\partial^2(\rho v_k'v_l')}{\partial t'^2\partial t''^2}}\,d^3\mathbf{y}'\,d^3\mathbf{y}''. \tag{3}$$

The first term under the overbar is evaluated at (y',t') while the second term is evaluated at (y'',t''). The emission times (t',t'') are retarded with respect to the reception time t. The quadrupole correlation shown with an overbar can be expressed as a function of the midpoint $\mathbf{y}$ and the separation in space and time

$$\mathbf{y}=\tfrac{1}{2}(\mathbf{y}'+\mathbf{y}''),\quad \mathbf{r}=\mathbf{y}'-\mathbf{y}'',\quad \tau=t'-t''. \tag{4}$$

If the observer distance R is large compared to the flow dimensions, then

$$C_\infty\tau\cong\frac{\mathbf{r}\cdot\mathbf{R}}{R}. \tag{5}$$

A convenient transformation of Eq. (3) is (see Ribner[4])

$$\overline{p^2}(\mathbf{R},\theta,\phi)=\frac{R_iR_jR_kR_l}{16\pi^2C_\infty^4R^6}\int_V\frac{\partial^4}{\partial\tau^4}\overline{(\rho v_iv_j)(\rho' v_k'v_l')}\,d^3\mathbf{r}. \tag{6}$$

When the turbulence is assumed quasi-incompressible, Eq. (6) becomes

$$\overline{p^2}(\mathbf{R},\theta,\phi)=\frac{\rho^2R_iR_jR_kR_l}{16\pi^2C_\infty^4R^6}\int_V\frac{\partial^4}{\partial\tau^4}\overline{v_iv_jv_k'v_l'}\,d^3\mathbf{r}, \tag{7}$$

where $S_{ijkl}=\overline{v_iv_jv_k'v_l'}$ is the fourth-order velocity correlation tensor. Ribner[4] expressed S_{ijkl} in terms of linear combinations of second-order correlations as follows:

$$S_{ijkl}=S_{ik}S_{jl}+S_{il}S_{jk}+S_{ij}S_{kl}\,, \tag{8}$$

where $S_{ij}=\overline{v_iv_j'}$. Ribner[4] has shown that for an axisymmetric source field and for a locally isotropic turbulent field, only few of the 81 possible combinations of indices (i,j,k,l) have a significant contribution. The various source terms are then evaluated using the isotropic turbulence model of Batchelor[17] and by assuming that the second-order correlations are separable functions of $\mathbf{r}$ and τ (see Ribner[4])

$$S_{ij}(\mathbf{r},\tau)=\Lambda_{ij}(\mathbf{r})g(\tau), \tag{9}$$

with

$$\Lambda_{ij}(r)=k[(f+\tfrac{1}{2}rf')\,\delta_{ij}-\tfrac{1}{2}f'r_ir_j/r], \tag{10}$$

where k is the local turbulent kinetic energy and f some universal function of r given by

$$f(r)=e^{-\pi^2r^2/L^2},\quad \text{with}\ \ r^2=x^2+y^2+z^2. \tag{11}$$

In the above definition of f, L is a longitudinal macroscale of the turbulence. In Eq. (9), $g(\tau)$ is the time delay factor assumed by Ribner[4] to be of the form

$$g(\tau)=e^{-(\tau/\tau_0)^2}, \tag{12}$$

where τ_0 represents the characteristic time delay which determines the minimum significant correlation in a moving reference frame. The length scale L in Eq. (11) is related to τ_0 by

$$L=c_lk^{1/2}\tau_0\,, \tag{13}$$

where c_l is an empirical constant. Using Eqs. (8)–(13) in Eq. (7) and taking the Fourier transform of the resulting equation with respect to τ, we arrive at the spectrum

$$I(\Omega)=c_l^3\rho^2k^{7/2}(\Omega\tau_0)^4e^{-(1/8)(\Omega\tau_0)^2}, \tag{14}$$

where Ω is the source frequency, which is related to the observed frequency by (see Mani *et al.*[11])

$$\Omega=2\pi f\sqrt{(1-M_c\cos\theta)^2+(c_dk^{1/2}/C_\infty)^2}. \tag{15}$$

In Eq. (15), M_c is the convection Mach number given by

$$M_c=0.5M+c_cM_j\,, \tag{16}$$

where M is the local Mach number, M_j is the nozzle exit Mach number, and (c_c,c_d) are empirical constants. Light-

hill's acoustic analogy approach, which is based on the classical wave equation, doesn't include the effects of the surrounding mean flow. Several studies were carried out to quantify the mean flow effects on sound radiation.[18,19] It was found that the mean flow results in not only the refraction of the radiated sound, but also convection amplification due to fluid motion. The Mani *et al.*[11] formulation is used to link the turbulent properties of a jet plume to the radiated acoustic field as follows:

$$\overline{p^2}(R,\theta,\Omega)=\int_V \Psi(a_{xx}+4a_{xy}+2a_{yy}+2a_{yz})d^3r, \quad (17)$$

with Ψ given by

$$\Psi=\frac{I(\Omega)}{16\pi^2R^2C_\infty^4}\left(\frac{\rho_\infty}{\rho}\right)^2\left(\frac{C_\infty}{C}\right)^2(1-M\cos\theta)^{-2} \times(1-M_c\cos\theta)^{-1}, \quad (18)$$

where ρ_∞ is the ambient density of the fluid, and C is the local speed of sound. The directivity factors $(a_{xx},a_{xy},a_{yy},a_{yz})$ have different forms depending on the location of the source, the velocity, and temperature profiles in the vicinity of the source. These factors depend explicitly on a shielding function g^2 that has the form

$$g^2(r)=\frac{(1-M\cos\theta)^2(C_\infty/C)^2-\cos^2\theta}{(1-M_c\cos\theta)^2}. \quad (19)$$

Depending on the location and the shape of the profiles (i.e., velocity and temperature), the function g^2 may have both positive and negative regions in space. When a negative region exists, fluid shielding of the source is possible. The directivity factors are given by

$$a_{xx}=\frac{\cos^4\theta}{(1-M_c\cos\theta)^4}\beta_{xx}, \quad (20a)$$

$$a_{xy}=\frac{g_0^2\cos^2\theta}{2(1-M_c\cos\theta)^2}\beta_{xy}, \quad (20b)$$

$$a_{yy}=\tfrac{3}{8}g_0^4\beta_{yy}, \quad (20c)$$

$$a_{yz}=\tfrac{1}{8}g_0^4\beta_{yz}, \quad (20d)$$

where g_0^2 is the value of $g^2(r)$ at the source radius $r=r_0$. The shielding coefficients $(\beta_{xx},\beta_{xy},\beta_{yy},\beta_{yz})$ depend on the location of the source and are set equal to 1 in the present study (i.e., no shielding of the source).

In Eq. (12), a characteristic time delay, τ_0, was introduced. The most widely used value for this time scale is

$$\tau_{0,\varepsilon}=c_t^{\varepsilon}\frac{k}{\varepsilon}, \quad (21)$$

where ε is the dissipation of the turbulent kinetic energy. In this paper, one more time scale is considered

$$\tau_{0,\mathrm{Pr}}=c_t^{\mathrm{Pr}}\frac{k}{\mathrm{Pr}}. \quad (22)$$

The production of the turbulent kinetic energy, Pr, in Eq. (22) is written as

$$\mathrm{Pr}=\frac{\mu_t}{\rho}\left[2(u_x^2+v_y^2+w_z^2)+(u_y+v_x)^2+(v_z+w_y)^2 +(w_x+u_z)^2-\frac{2}{3}(u_x+v_y+w_z)^2\right], \quad (23)$$

where μ_t is the turbulent viscosity, and (u,v,w) are the plume velocity components in the (x,y,z) directions, respectively. In the extended version of the $(k-\varepsilon)$ turbulence model[20] used in the current CFD computation, the source term in the dissipation rate equation is given by

$$S_\varepsilon=\rho[C_1\,\mathrm{Pr}(\varepsilon/k)-C_2\varepsilon(\varepsilon/k)+C_3\,\mathrm{Pr}(\mathrm{Pr}/k)], \quad (24)$$

which shows the presence of two time scales; k/ε and k/Pr. Therefore, an acoustic model that accounts for radiation from both time scales is needed. This model determines the far-field mean-square-pressure as the average of the mean-square-pressure given by the individual time scales following

$$\overline{p^2}=\tfrac{1}{2}[\overline{p_\varepsilon^2}+\overline{p_{\mathrm{Pr}}^2}], \quad (25)$$

where $\overline{p_\varepsilon^2}$ and $\overline{p_{\mathrm{Pr}}^2}$ are associated with $\tau_{0,\varepsilon}$ and $\tau_{0,\mathrm{Pr}}$ of Eqs. (21) and (22), respectively. The model defined by Eq. (25) is referred to as the ''dual time scale'' model.

III. THE CFD DATABASE

In order to use the far-field acoustic radiation model derived above, one needs the local steady-state engine plume data $(\rho,C,M,k,\varepsilon,\mathrm{Pr},U)$ and the plume grid. If the test conditions are known accurately, only one dataset is required. However, if more test conditions are planned then one needs to have a CFD solution for each set of conditions. In the example presented here, the Reynolds-averaged-Navier–Stokes (RANS) equations are integrated over a domain 2.5 times the length of the Venture Star space vehicle of 25 m (this vehicle is also known as the X-33). The CFD code used, known as FDNS,[21] uses a second-order finite difference scheme to discretize the diffusion fluxes and source terms of the governing equations. The convective terms are discretized using a second-order total-variation-diminishing difference scheme. The volume grid is composed of 2.2 million grid points as shown in Fig. 2. Details about the grid refinement techniques and the treatment of the various boundary conditions and source terms can be found in Ref. 22. An extended $(k-\varepsilon)$ turbulence model describes the turbulence. Figure 3 shows the temperature distribution in the plume region.

IV. EXPERIMENTAL MEASUREMENTS

NASA conducted ten single-engine tests of the linear aerospike engine, XRS-2200 at the Stennis Space Center in late 1999 and early 2000. Figure 4 is a photograph of this liquid-fueled rocket engine and the initial part of the exhaust plume. The engine was suspended above a J-deflector where the vertical distance (along the centerline of the engine) from the engine exit ramp plane to the deflector was 15.244 m. For CFD models, relevant dimensions of the test configuration include the nozzle ramp (2.82×2.24 m at one end and 1.07 ×2.24 m at the other end; the ramp is 1.27 m long and

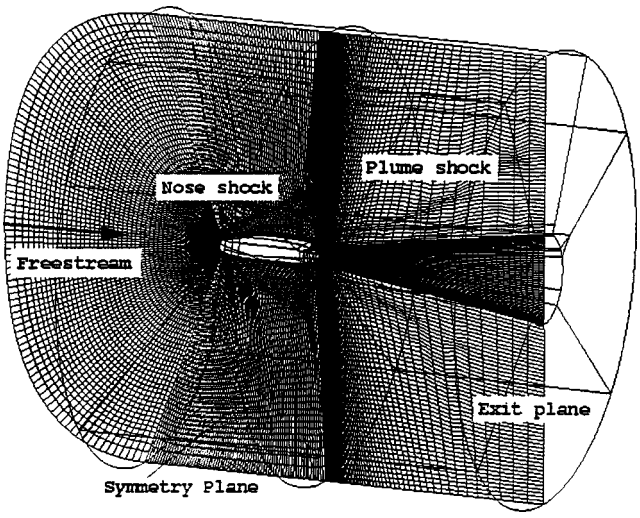


FIG. 2. Computational grid for the X-33 vehicle.

possesses an area ratio of 58). The jet Mach number was 3.042. The test setup included microphones that were arranged at ground level at various distances and angles from the test stand, as shown in Fig. 5. The ''Ms'' and ''Ns'' of Fig. 5 designate the locations of the microphones and amplifiers. The microphones were vented, polarized capacitor type, 1.27 cm in diameter (B&K 4190 and 4191). The microphone signals were preamplified (B&K 2669) and passed through shielded cabling to amplifiers (Nexus 2690), then through coaxial cables to an analog tape recorder. A Teac XR-510 tape recorder captured the signal with operational conditions of 5 volts full scale at 76-cm/s tape speed in the FM mode. Prior to each XLS-2200 test, an end-to-end calibration procedure was used. The crew placed a calibrator (B&K 4231) over each microphone that produced an acoustic output of 94 dB and 1000 Hz: the microphone output signal for this tone was recorded. The reference pressure used for the decibel level is 20 mPa.

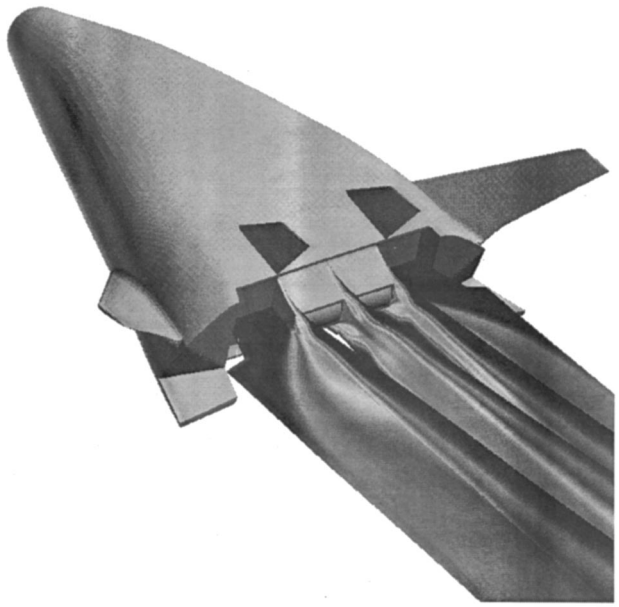


FIG. 3. Temperature distribution in the plume region near the vehicle obtained by CFD.

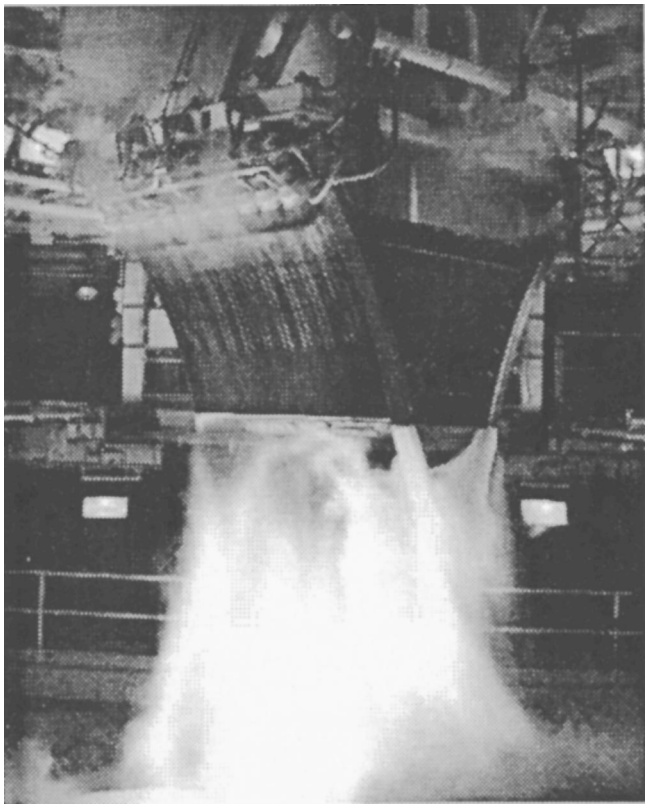


FIG. 4. Picture of the hotfire test showing the water vapor enhanced engine plume.

The microphone frequency range was 5 Hz to 20 kHz, and full-scale sound levels varied from 140 to 160 dB depending on location. For each microphone, the background noise was measured and removed from the test data (at these high noise levels, the background noise was negligible). These data records were processed after each test and provide a database for comparison with linear aerospike acoustic environment predictions. The database includes spectral shapes, overall sound-pressure-level distributions, and decay rates.

V. ACOUSTIC MODEL VALIDATION

Initially, the empirical constants of the model $(c_c, c_d, c_l, c_t^{\mathrm{Pr}}, c_t^{\varepsilon})$ are selected by using a measured spectrum in 1/3-octave frequency bands at a fixed observer location. As depicted in Fig. 5, several sound measurements are available at a radius of 90 m from the XRS-2200 test stand and at various angular locations with respect to the deflected plume. Figure 6 shows a comparison between the measured and calculated spectrum at the 90-deg location. This spectrum is obtained after a few trial and error runs with different empirical constants. In addition to using the experimental spectrum, the overall sound-pressure level (OASPL) at the same location is used in choosing the constants. The final set of selected empirical constants is $c_c = 0.27$, $c_d = 0.55$, $c_t^{\mathrm{Pr}} = 0.04$, $c_t^{\varepsilon} = 0.045$, and $c_l = 1.13$. Figure 7 shows the calculated and measured overall sound-pressure level as a function

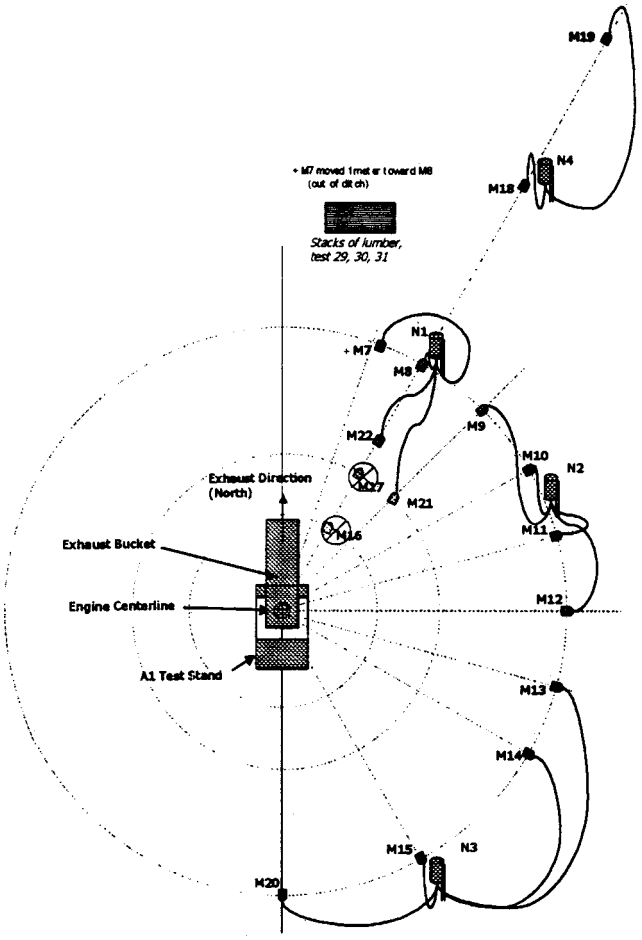


FIG. 5. Midfield measurement locations around the test stand.

of the angular location. Good agreement is obtained at most locations except near 180°, where a wall separates the microphone from the engine. This is not taken into account by the far-field acoustic model.

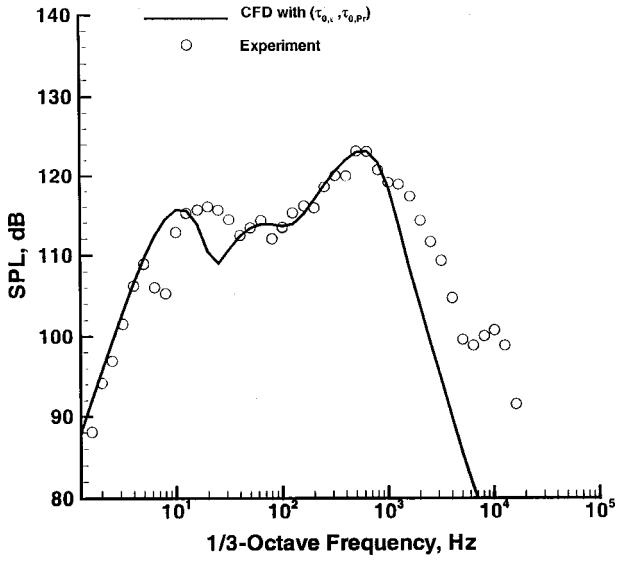


FIG. 6. Comparison of the measured and predicted spectrum at a midfield location 100 meters away from the test stand and at 90° with the deflected plume.

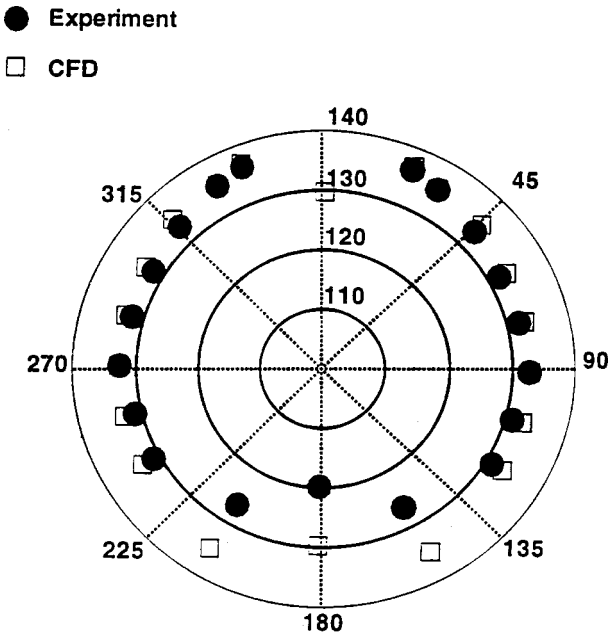


FIG. 7. Measured and predicted overall sound-pressure level at various angular locations in the midfield.

VI. THE MEASUREMENT AND PREDICTION TOOL

The innovative aspect of the proposed measurement tool resides in its capability to make predictions while taking experimental measurements and display both the predicted and measured results in real time (i.e., during a test). The acoustical engineer, using the tool, can adjust two of the empirical constants (c_c, c_t) on the display screen and enter the location of the microphone so that the semiempirical model can provide predictions at that particular location. The two empirical constants given above are chosen because the model is more sensitive to these constants than to the other ones. Figure 8 shows the major functional components of the system. In order to perform its task correctly, the tool needs a CFD database (results from prior CFD computations for the jet engine being tested and at the same conditions), which will be used in the integration of the far-field acoustic model, and the experimental data being collected in real time.

These components can be used as stand-alone or integrated. In a stand-alone application, the semiempirical model will use previously optimized empirical coefficients to make acoustic radiation predictions for any combination of new engine test conditions. A CFD plume solution that matches the new test condition is necessary to complete the model. Figure 9 shows four screen displays obtained by adjusting the above-mentioned empirical constants. Each display shows the experimental data together with the predicted spectrum.

VII. CONCLUSIONS

A new, innovative measurement and prediction tool has been designed and tested. The purpose of the tool is to measure the far-field noise radiated by an engine plume. In addition, the tool is equipped with a prediction package based on a semiempirical far-field acoustic radiation model and an existing CFD database of engine plume. Test results show

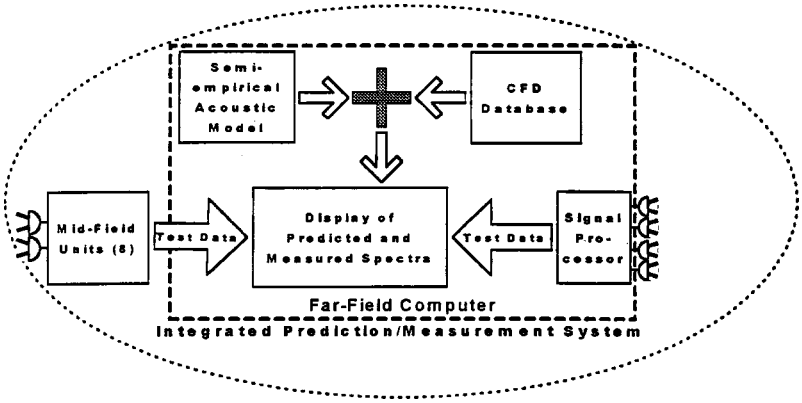


FIG. 8. The various functional components of the proposed measurement and prediction tool.

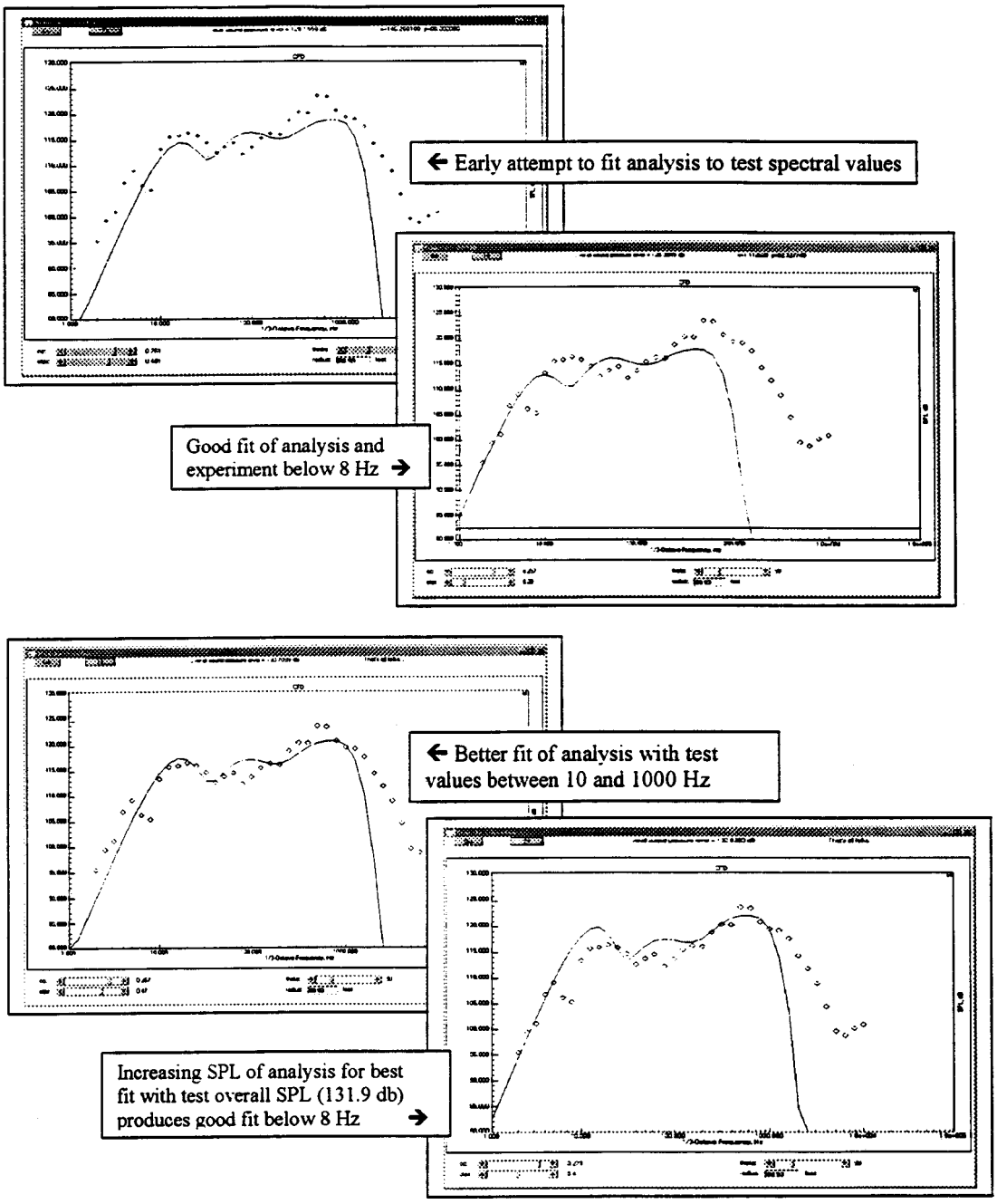


FIG. 9. Various screen displays of the measured and predicted results using proposed tool.

that the tool is able to display both the predicted and measured results in real time. The semiempirical model has been tested and shown to give good agreement with far-field measurement.

ACKNOWLEDGMENTS

A. Frendi, Wade D. Dorland, and Thein Maung would like to acknowledge the support of NASA Marshall Space Flight Center under an SBIR Phase I Grant No. ASRI-NAS8-01009-01 with Tom Nesman as the technical monitor.

[1] A. Frendi, ''On the CFD Support for the Hyper-X Aerodynamic Database,'' AIAA paper 99-0885 (1999).
[2] A. Frendi, T. E. Nesman, and T.-S. Wang, ''Computational and Experimental Study of the Linear Aerospike Engine Noise,'' AIAA J. **39**(8), 1485–1492 (2001).
[3] M. J. Lighthill, ''On Sound Generated Aerodynamically. I. General Theory,'' Proc. R. Soc. London, Ser. A **211**, 564–587 (1952).
[4] H. S. Ribner, ''Quadrupole Correlations Governing the Pattern of Jet Noise,'' J. Fluid Mech. **38**(1), 1–24 (1969).
[5] H. S. Ribner, ''On the Strength Distribution of Noise Sources Along a Jet,'' J. Acoust. Soc. Am. **31**, 245–246 (1958).
[6] H. S. Ribner, ''On Spectra and Directivity of Jet Noise,'' J. Acoust. Soc. Am. **35**, 614–616 (1963).
[7] J. E. Ffowcs Williams, ''Hydrodynamics Noise,'' Annu. Rev. Fluid Mech. **1**, 197–222 (1969).
[8] J. E. Ffowcs Williams, ''Aeroacoustics,'' Annu. Rev. Fluid Mech. **9**, 447–468 (1977).
[9] M. E. Goldstein, ''Aeroacoustics of Turbulent Shear Flows,'' Annu. Rev. Fluid Mech. **16**, 263–285 (1984).
[10] C. K. W. Tam, ''Supersonic Jet Noise,'' Annu. Rev. Fluid Mech. **27**, 17–43 (1995).
[11] R. Mani, E. J. Stringas, J. C. F. Wang, T. F. Balsa, P. R. Gliebe, and R. A. Kantola, ''High Velocity Jet Noise Source Location and Reduction,'' Task 2, FAA-RD-76-79-II (1977).
[12] J. E Ffowcs Williams, ''The Noise from Turbulence Convected at High Speed,'' Philos. Trans. R. Soc. London, Ser. A **255**, 469–503 (1963).
[13] A. Khavaran, E. A. Krejsa, and C. M. Kim, ''Computation of Supersonic Jet Mixing Noise from an Axisymmetric CD Nozzle Using $(k-\varepsilon)$ Turbulence Model,'' AIAA Paper 92-0500, Jan. 1992.
[14] C. Baily, W. Bechara, P. Lafon, and S. Candel, ''Jet Noise Predictions Using a $(k-\varepsilon)$ Turbulence Model,'' AIAA Paper 93-4412, Oct. 1993.
[15] C. Baily, P. Lafon, and S. Candel, ''Prediction of Supersonic Jet Noise from a Statistical Acoustic Model and a Compressible Turbulence Closure,'' J. Sound Vib. **194**(2), 219–242 (1996).
[16] C. K. W. Tam and L. Auriault, ''Jet Mixing Noise from Fine-Scale Turbulence,'' AIAA J. **37**(2), 145–153 (1999).
[17] G. K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge University Press, Cambridge, 1953).
[18] T. F. Balsa, ''The Acoustic Field of Sources in Shear Flow with Application to Jet Noise: Convective Amplification,'' J. Fluid Mech. **79**(1), 33 (1977).
[19] M. E. Goldstein, ''High Frequency Sound Emission from Moving Point Multipole Sources Embedded in Arbitrary Transversely Sheared Mean Flows,'' J. Sound Vib. **80**(4), 499 (1982).
[20] Y.-S. Chen and S. W. Kim, ''Computation of Turbulent Flows Using an Extended $k-\varepsilon$ Turbulence Closure Model,'' NASA CR-179204, Oct. 1987.
[21] Y.-S. Chen, ''FDNS—A General Purpose CFD Code, Version 4.0,'' Engineering Sciences, Inc., ESI-TR-97-01, Huntsville, AL, May 1997.
[22] T.-S. Wang, ''Analysis of Linear Aerospike Plume-Induced X-33 Base-Heating Environment,'' J. Spacecr. Rockets **36**(6), 777–783 (1999).

Sweep-spread carrier for underwater communication over acoustic channels with strong multipath propagation

Konstantin G. Kebkal[a)]
State Oceanarium of Ukraine, Epronovskaya 9, 335024 Sevastopol, Ukraine

Rudolf Bannasch
Technical University Berlin, Ackerstr. 71-76 (ACK1), 13355 Berlin, Germany

(Received 24 August 2000; revised 21 April 2001; accepted 9 July 2002)

Multipath propagation causes the transmitted signal to take numerous, time-varying different-length paths to the receiver. Exploitation of conventional frequency-constant carrier signals for communication over underwater acoustic channels typically shows that intricate mutual time variations of multipath arrivals create an extremely hard problem for equalizers to compensate for nonstable intersymbol interactions. Communication over such channels can be, however, substantially improved by using a new method based on the implementation of a sweep-spread carrier (S2C). Such a carrier consists of a succession of sweeps, which cause permanent rapid fluctuation of signal frequency. Apart from some additional useful effects such as providing multiuser access and reducing influence of narrow-band noise, the method provides significant processing improvement enabling clear separation of multipath arrivals by converting their time delays into their frequency reallocations—the steeper the sweeps, the better the multipath resolution that can be achieved. When converting the signal into constant intermediate frequencies, the best suitable multipath arrival can be separated not as traditionally in time domains, applying complex equalizers, but in frequency domains by means of usual band-pass filters. High transmission stability of this approach was confirmed in both computer simulations as well as in validation experiments carried out since summer 1999. © *2002 Acoustical Society of America.*
[DOI: 10.1121/1.1504855]

PACS numbers: 43.60.Dh, 43.30.Cq [JCB]

I. INTRODUCTION

A major problem for achieving reliable and high rate underwater acoustic (UWA) communication is the severe time-variable intersymbol interference (ISI) strongly introduced in horizontal shallow water channels and, especially strong, between moving objects. In shallow water channels of 1–2 km range, no universal equation determining multipath features is available.[1] Reflections from channel boundaries and diverse objects dominate the multipath structure. The structure is geometry specific and not predictable for every communication event. Time variant multipath propagation is recognized as a basic channel characteristic, and single-path temporal fluctuations are considered as primary performance problems.

The various approaches developed for communication in shallow water can be divided into two main groups: phase-coherent and noncoherent methods. The last group of methods accomplishes robust communication link but provides rather low data rates. Contrary—phase coherent—methods allow much faster link using, however, typically highly complex processing procedures and expensive, noncompact array constructions. Designed for high data rate communication in the severe ISI environment, systems based on the phase-coherent methods employ either some form of array processing, or equalization methods, or their combinations. Three main approaches have recently been developed for implementation in phase-coherent receivers. They are introduced in (1) processing methods based on joint synchronization and equalization for combating the effects of phase variations and severe ISI,[2,3] (2) array processing methods[4] focused on the use of arrays to excite only a single path of propagation, ensuring that only a single deterministic path makes a major contribution to received signal, and (3) methods of adaptive beamforming for multipath rejection at the receiver end providing an adaptive algorithm for nulls steering in the direction of a surface reflected wave.[5,6]

The first approach has been developed at Northeastern University and WHOI.[2,3,7] It contains the strategy of the optimal phase synchronization and channel equalization jointly. The phase coherent communication system has been shown achieving high data rates by successfully tracking and updating time-varying phase fluctuations in signals dominated by multipath components.[2] However, serious application limitations emerge for this kind of coherent communication if impulse response functions within an equalized packet become significantly decorrelated.[8] In practical applications, source-receiver movements can considerably induce such coherence diminution. Most suitable in dynamic channels ''fast'' equalizers are typically applied, being realized on the base of recursive least squares (RLS) algorithm. However, even the fastest among others, an adaptive RLS equalizer encounters convergence problems when a channel impulse response changes significantly in less than 10 impulse response durations (symbol durations), which is often approached in ocean

[a)]Current address: Technical University Berlin, Ackerstr. 71-76 (ACK1), 13355 Berlin, Germany. Electronic mail: kebkal@bionik.tu-berlin.de

channels.[1,9] Time-variant multipath structures also create difficult synchronization problems, typically encountered by communication in nearly all channels of interest.[8–10]

The second approach uses differently coherent detection and relies on array processing to eliminate multipath arrivals.[4] It focuses on the use of transmitter arrays to excite only a single path of propagation. Equalization is not used, since the shallow water channel is considered exhibiting too rapid Rayleigh fading for an adaptive equalizer to follow. In the receiving end instead, an array is employed to compensate for random propagation irregularities. Generally, it was found that the technique becomes remarkably less effective at longer ranges.

The third approach, developed at the University of Newcastle, uses an adaptive beamforming at the receiver end and carries out the isolation of an undesired propagation path based on angular separation of multipath arrivals.[11] The approach exploits a least-mean-squares-type algorithm to adaptively steer nulls in the direction of reflected wave.[5,6] Similarly as in the last case, it was found that the beamforming system reduces its performance with growing communication distance as a consequence of decreasing the angular separation rate for all the propagation paths in the channel.[12] To compensate for this effect, the use of an equalizer was proposed to complement the performance of the beamformer.[5] Such combination implies again, however, the likely equalizer convergence problems often encountered in dynamic fast-varying channels.

The limitations of current approaches motivate the search for another strategy by which the time-varying ISI can be effectively treated.

As noticed by Catipovic,[1] the multipath propagation itself does not introduce the fundamental performance limitation for acoustic telemetry system; rather, single-path temporal fluctuations are primary performance problems. This concern stays perfectly true so far as a time-based filtering approach, like equalization based on a tapped delay line filter, takes place. Temporal phase/frequency impairments, obtained by the signal during its propagation over different paths, are coupled at the receiver so that no accurate time-domain filtering can be accomplished. The longer the time spread the stronger the dynamic influence, the less accuracy that can be achieved.

In this paper, an alternative approach is suggested, where the separation of every single path occurs not in time but in frequency domains by means of usual band-pass filters. In order to facilitate it, a revision of carrier signals (Sec. II) must initially take place.

A significant advantage of the method consists in the fact that an accurate multipath separation can be easily achieved without application of complex processing structures like adaptive equalizers. Instead, a simple preprocessing unit will be implemented into the receiver section capable of transforming a complex time-varying multipath mixture to a combination of frequency-spaced multipath arrivals making use of a new sort of carrier signal applied. An important aspect is the effect that every multipath signal is separated together with only its individual time-varying (phase) impairments, obtained with the signal during its propagation over a certain path. That is, signal impairments, specific for a given propagation path, are completely separated from signal deformations obtained by the signal during its propagation over every different path. Hence, each single time-varying multipath arrival can be processed individually and, thus, in the best form for further evaluation of its information symbols. If necessary, all single-path deformations can be tracked during the processing by means of conventional and rather simple components like phase tracking loops.

In the paper, we introduce an innovative sweep-spread carrier communication (S2C-communication) method capable of significantly improving the signal-to-multipath and signal-to-noise ratio. The method considers sweeps not as information but as carrier signals. Using the property of acoustic signals to propagate under water with relatively low speed, the method allows us to accurately separate multipath arrivals by converting their time delays into their frequency reallocations provided that sufficiently steep sweeps are applied.

The operating environment of the proposed system would be a broad range of shallow water channels both with strongly expressed thermally layered water and with well-mixed isothermal water. It can be applied to any acoustic channel where the effect of multipath propagation has a significant influence. The respective system can be made compact and easy to deploy since only one transmitter is needed; real time analysis can be implemented using relatively inexpensive processing components.

It should be noted that this study was designed to investigate only the feasibility of the new method, therefore, no large data rates were applied, and also no tests for long distance link were carried out. Since a pool or small shallow water lake represents the ''worst case'' in terms of reflections strength and their spread delay, those conditions were primarily chosen for our experiments. The effect of distance enlargement was initiated by varying the transmission power, channel variations were achieved via vertical and horizontal moving of the transmitter hydrophone, which, together with the existence of strongly introduced multipath propagation, allowed for the creation of experimental conditions very close to those usually experienced in practical applications.

In Sec. II, the sweep-spread carrier is defined. Section III gives an analysis of the behavior of signals with such a carrier in a multipath channel. The system design is given in Sec. IV. The off-line processed experimental results are presented in Sec. V, followed by discussion and conclusions.

II. THE SWEEP-SPREAD CARRIER

The current approach is associated with spread-spectrum communication methods.

Let us concentrate on a phase shift keying (PSK) system. The peculiarity of the new method is that the phase-encoded signal is modulated not only to a commonly used frequency-constant or stepped varying carrier, but onto a carrier that experiences permanent frequency variations. We shall refer to it as the sweep-spread carrier (S2-carrier), which consists of a succession of sweeps spreading the signal spectrum over a broad bandwidth.

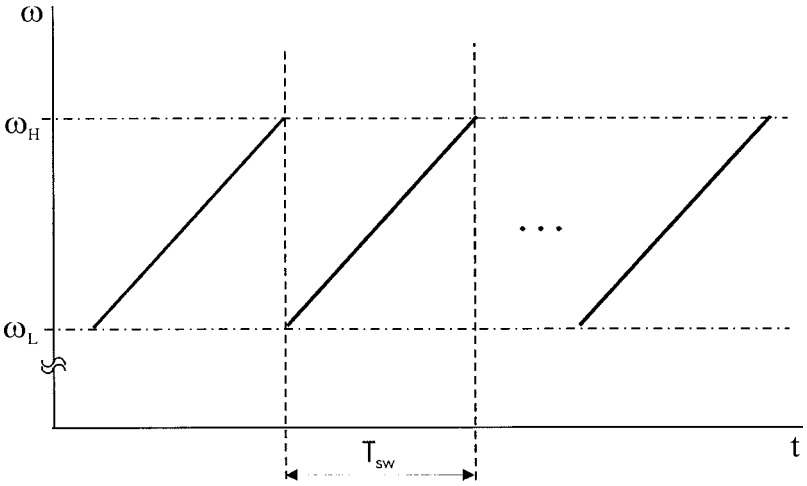


FIG. 1. S2-carrier consisting of a succession of sweeps.

What is the advantage of such a solution for underwater acoustic communication? Usually, when sending a signal of constant frequency through a channel with strong time varying and extended reverberation, the highly unstable multipath arrivals will interfere with each other so that the phase as well as the amplitude of the resulting signal will permanently fluctuate. Under those conditions, it becomes extremely complicated to work with phase- or amplitude-encoded signals. However, by applying signals with a frequency gradient (sweeping) carrier, the problem can be solved. If the frequency gradient is rather large, then each multipath signal, arriving with its individual time delay to receiver, will have its own instant frequency, which significantly differs from instant frequencies of all other multipath arrivals. Thus, for example, an instant frequency of a reflected sweep signal can lag essentially behind the frequency of the direct sweep signal, and, in this way, the multipath-induced masking effect can be effectively avoided. The larger the sweeping rate, the better multipath resolution can be achieved.

Before considering the behavior of the S2-carrier in a numerical model of an UWA channel, it should be defined in mathematical terms.

Let the S2-carrier consist of a succession of sweeps with frequency variation from ω_L to ω_H within a time interval T_{sw}, and let all the sweeps be uniformly produced in a linear manner with rapid frequency variation following each other successively without any gap between them (Fig. 1). Then such an S2-carrier can be written as

$$c(t)=A_c\ \exp\left[j\left(\omega_L\left(t-\left\lfloor\frac{t}{T_{\mathrm{sw}}}\right\rfloor T_{\mathrm{sw}}\right)+m\left(t-\left\lfloor\frac{t}{T_{\mathrm{sw}}}\right\rfloor T_{\mathrm{sw}}\right)^2\right)\right], \tag{1}$$

where A_c is the amplitude, $m=(\omega_H-\omega_L)/2T_{\mathrm{sw}}$ is a coefficient denoting the frequency variation rate, ω_L and ω_H denote respectively the lowest and highest angular frequencies, T_{sw} is the sweep duration, and the term $[t/T_{\mathrm{sw}}]$ denotes the operand for truncating the value to the nearest least integer. Per definition,

$$\left(t-\left\lfloor\frac{t}{T_{\mathrm{sw}}}\right\rfloor T_{\mathrm{sw}}\right)=\left\{\frac{t}{T_{\mathrm{sw}}}\right\}T_{\mathrm{sw}}, \tag{2}$$

which can be interpreted in Eq. (1) as an actual cycle time with the cycle duration T_{sw}.

The peculiarity of the carrier expressed with Eq. (1) consists in the presence of a quadratic member in the parentheses. (Here we use linearly modulated sweeps with constant m, but in general, any sonogram form can be defined by changing m with time.)

III. BEHAVIOR SIGNALS WITH THE S2-CARRIER IN A MULTIPATH CHANNEL

Let $s(t)$ symbols be phase encoded data. The symbols are modulated onto the S2-carrier, that is, $x(t)=s(t)\cdot c(t)$, what produces a sweep-spread signal (S2-signal). The signal is transmitted over the UWA channel, whose overall model is shown in Fig. 2. The part of the model which represents the water medium consists of a number of delay elements with τ_i, denoting time intervals between two successive multipath arrivals, and a number of multiplication elements with coefficients V_i taking into account possible attenuations on interfering multipath signals.

The term $h_s(t)$ represents the impulse response of the transmitting filter and $h_r(t)$ that of the receiving filter for intermediate frequencies. A real, practical model contains typically not constant but time-varying delay elements as well as time-varying coefficients V_i. However, to simplify the analysis of the received signal, we neglect these variations on the current step and derive an expression for a signal received after propagation over this channel.

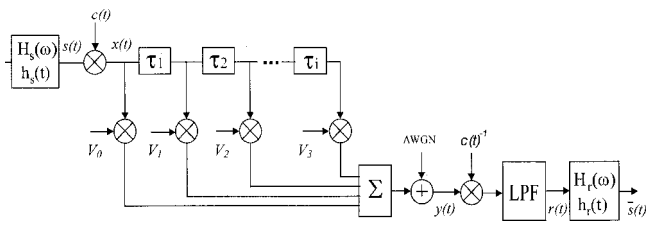


FIG. 2. Model of a UWA channel.

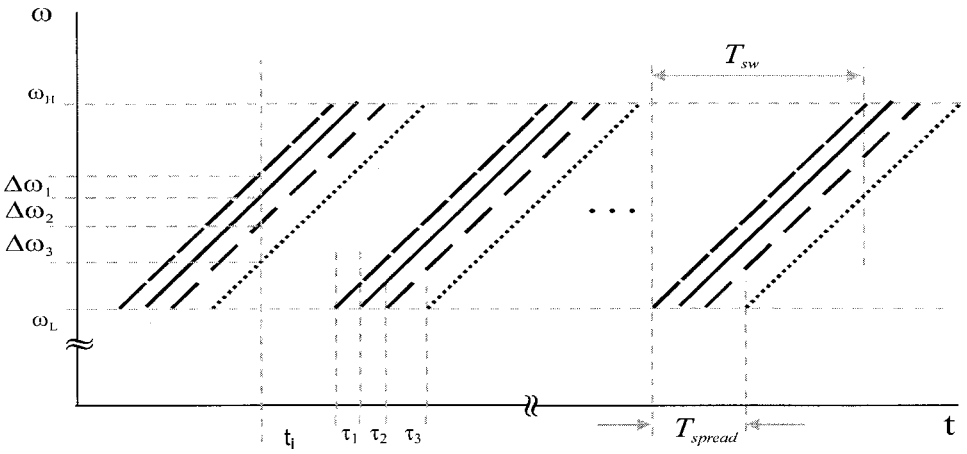


FIG. 3. Schematic drawing of a sonogram registered at receiver end.

If both the $c(t)$ and $s(t)$ have unit amplitudes, and, as noted, every coefficient V_i and delay element τ_i stays constant over the entire transmission line, then, after propagation along different paths in the underwater medium, the transmitted signal comes to the receiver in the form:

$$y(t)=V_0x(t)+\sum_i V_ix(t-\tau_i)+n(t), \tag{3}$$

where $x(t)$ is defined above, $x(t-\tau_i)$ is written as follows,

$$x(t-\tau_i)=s(t-\tau_i)\ \exp\left[j\left(\omega_L\left\{\frac{t-\tau_i}{T_{\rm sw}}\right\}T_{\rm sw}+m\left(\left\{\frac{t-\tau_i}{T_{\rm sw}}\right\}T_{\rm sw}\right)^2\right)\right], \tag{4}$$

and $n(t)$ is the white noise. It is evident that

$$\left\{\frac{t-\tau_i}{T_{\rm sw}}\right\}T_{\rm sw}=\begin{cases}t_c-\tau_{ci}, & t_c\geqslant\tau_{ci},\\ T_{\rm sw}+t_c-\tau_{ci}, & t_c<\tau_{ci},\end{cases} \tag{5}$$

where $t_c=\{t/T_{\rm sw}\}T_{\rm sw}$ is the cycle time defined in Eq. (2), and $\tau_{ci}=\{\tau_i/T_{\rm sw}\}T_{\rm sw}$ is a fractional part of the time delay related to the sweep duration $T_{\rm sw}$. Thus, every delayed arrival represented in the second member of Eq. (3) can be rewritten as

$$x(t-\tau_i)=\begin{cases}s(t-\tau_i)\ \exp\ [j(\omega_L(t_c-\tau_{ci})+m(t_c-\tau_{ci})^2)],\\ \quad t_c\geqslant\tau_{ci},\\ s(t-\tau_i)\ \exp\ [j(\omega_L(T_{\rm sw}+t_c-\tau_{ci})\\ \quad +m(T_{\rm sw}+t_c-\tau_{ci})^2)], \quad t_c<\tau_{ci}.\end{cases} \tag{6}$$

After transformation of Eq. (6), each delayed arrival can be then written as

$$x(t-\tau_i)=s(t-\tau_i)\ \exp\ [j(\omega_Lt_c+mt_c^2)]\times\exp\ [j(-\Delta\omega_it_c+\varphi_i)], \tag{7}$$

where

$$\Delta\omega_i=\begin{cases}2m\tau_{ci}, & t_c\geqslant\tau_{ci},\\ -2m(T_{\rm sw}-\tau_{ci}), & t_c<\tau_{ci},\end{cases}$$

signifies the frequency deviation of an ith multipath arrival provided by the delay τ_i, and

$$\varphi_i=\begin{cases}(m\tau_{ci}-\omega_L)\tau_{ci}, & t_c\geqslant\tau_{ci},\\ (\omega_H+\omega_L-2m\tau_{ci})\dfrac{T_{\rm sw}-\tau_{ci}}{2}, & t_c<\tau_{ci},\end{cases}$$

is the phase of the ith multipath arrival.

The term with $i=0$ in Eq. (3) represents an attenuated version of the original signal, and the other term signifies the multipath diversity of its delayed, attenuated, and frequency-shifted reproductions. Especially important in Eq. (7) is that at any instant all the interfering multipath arrivals have different frequencies spaced by $\Delta\omega_i$ from each other. This effect is shown in Fig. 3 (see frequency axis at time point t_i).

The same can be shown also for time-varying channels. Let us insert an additional member ω_d in the equations (3) and (4), which reflects the influence of Doppler onto the received signal. Then, the first term in Eq. (3) must be consequently rewritten as $x(t)\ \exp\ (j\omega_dt)$ and Eq. (4) must be complemented with a respective term containing the Doppler frequency so that

$$x(t-\tau_i)=s(t-\tau_i)\ \exp\ \left[j\left(\omega_L\left\{\frac{t-\tau_i}{T_{\rm sw}}\right\}T_{\rm sw}+m\left(\left\{\frac{t-\tau_i}{T_{\rm sw}}\right\}T_{\rm sw}\right)^2\right)\right]\exp(j\omega_d^i(t-\tau_i)), \tag{8}$$

where ω_d^i is the Doppler frequency encountered in an ith propagation path. The last exponent in Eq. (8) can reflect also time-varying phase/frequency shifts experienced by the signal in the channel. In this case, the ω_d^i is characterized with a time-dependent function specific for the ith path, induced, for example, due to internal waves, reflections from moving surface waves, motion, etc. When analyzing the expression in Eq. (8), one can conclude, however, that Doppler shifts belonging to different paths will not be coupled (will not interact) while the ω_d^i—frequency band enlargement of ith multipath signal—stays within certain boarders; namely so that a maximum value $\omega_{d\max}^i$ of the time-varying band-

width enlargement ω_d^i does not extend half of the frequency separation space between respective multipath arrivals (e.g., $\omega_{d\max}^i < \Delta\omega_i/2$). In this case, every arrival, even being strongly deformed with a respective time-varying Doppler term, stays within definite frequency range and does not influence another frequency band (any other multipath arrival); no intermodulation between differently varying Doppler terms belonging to different propagation paths takes place.

Hence, the most interesting value of the time-varying ω_d^i is its maximum value $\omega_{d\max}^i$, which implies restrictions on the resolution gain of the multipath arrivals. Generally, to attain a good resolution, the following inequality must satisfy $\Delta\omega_i \gg \omega_{d\max}^i$. Accomplishment of this inequality becomes easily possible in UWA channels profiting from relatively low sound speed in water and from setting the sweeps to really steep values.

The following interpretation is central for our method: by introducing the S2-carrier, a time spread of the signal in a multipath channel can be converted into a frequency reallocation of individual multipath arrivals over a definite spectrum area. By achieving an appropriate frequency separation of the multipath arrivals, considerable frequency-based ''rectification'' of the receiving signal can be attained, whereby each desirable spectrum component (multipath arrival) can be selected and processed individually.[13] Moreover, every multipath arrival—frequency band—is separated together with its individual time-varying (phase) deformations, obtained by the signal during its propagation over a certain path; the arrival does not contain any time-varying impairments specific for different propagation paths. An obvious fact is that a function of single-path time variations is usually not as intricate as one induced by interaction of time variation functions delivered to a receiver with multiple arrivals over differently varying propagation paths. Therefore, if necessary, an accurate tracking of (motion-induced) signal deformations becomes easily possible even by means of conventional and rather simple components like phase-tracking loops.

Hence, any time-varying multipath distorted signal can be represented with one single arrival, i.e., in the best form for processing and consequent evaluation of an encoded parameter.

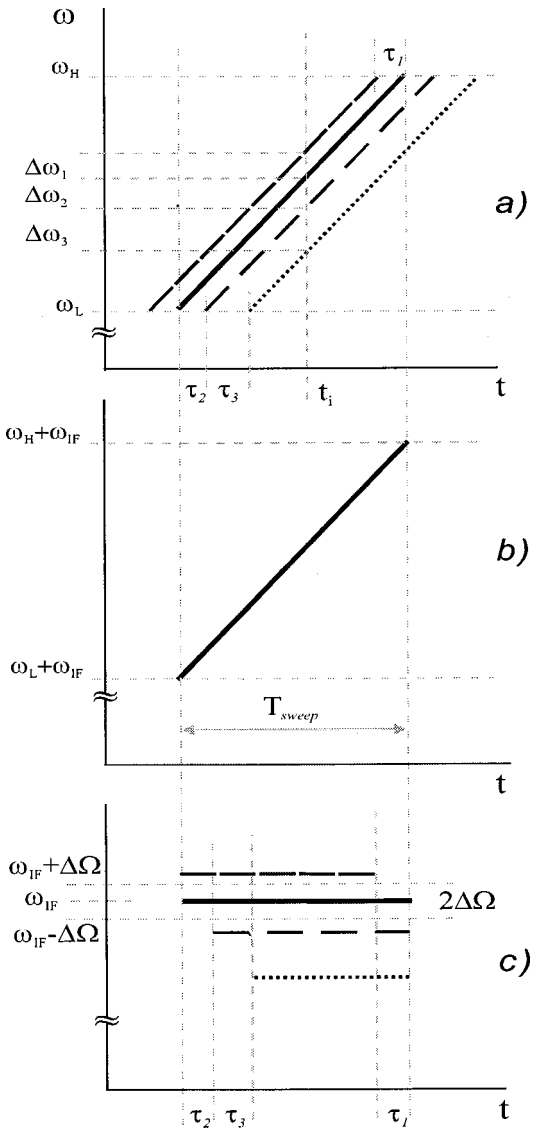


FIG. 4. Despreading: (a) a single sweep as part of a S2-carrier (4 multipath arrivals), (b) gradient-heterodyne signal, and (c) IF spectrum.

IV. SYSTEM DESIGN

A. Despreading

To simplify the demonstration of signal processing, we assume that only a few symbols are transmitted over the channel so that their total duration corresponds to only one single sweep.

Figure 4(a) schematically shows the sonogram of this signal coming to the receiver over four different propagation paths. Although only one single sweep cycle was sent, a multipath variety of four S2-carrier cycles shifted (in relation to each other) with certain delay intervals τ_i will be obtained by the receiver. These cycles are shown with different line styles, depicting apart from their different arrival instances also their different energy attributes. Because of the distinct attenuations over different propagation paths, one cycle from this variety may contain more energy than the others. Let this one be marked by a solid line. As it shows the best SNR, exactly this arrival is reasonable to separate from the others for further processing. However, it may be rather tricky to design a dynamic filter providing sufficiently sharp filtering to separate variable carrier frequencies.

Therefore, an appropriate transformation (despreading) must be primarily carried out, that is, the frequency variations must be removed. This transformation implies a multiplication of the received signal with an appropriately varying gradient-heterodyne signal having the same sweep cycle and the same slope of the frequency variation.

How to determine the S2-carrier cycle with most energy will be shown later. Here we assume the timing to be known from a previously accomplished training period. To distinguish the desired cycle from the others, we will refer to it as the ''main S2-cycle'' or ''main multipath signal,'' and the others as ''interfering S2-cycles'' or ''interfering multipath signals.''

The transformation begins with mixing the received sig-

nal [Fig. 4(a)] with the gradient-heterodyne signal [Fig. 4(b)]. With the correct timing parameter, generation of the heterodyne signal starts exactly at the instant when the main multipath signal arrives. Such a mixing produces two spectrum parts, where the lower one represents constant intermediate frequencies well suited for further processing. Figure 4(c) shows the spectrum of intermediate frequencies (IF) in the low-pass range. Every spectrum line in the range corresponds to an arrival of the S2-signal over a given propagation path. Apart from different time delays τ_i, every multipath arrival has also its unique frequency value (which is a very desirable result). This gives us a very good opportunity for combating the multipath interferences in frequency domains via usual band-pass filtering. After separation, each multipath arrival can easily be handled and independently processed in any conventional way for evaluation of its information parameter.

B. Initial estimation and specification of spreading and synchronization parameters

To initialize the event, a short synchronization preamble consisting of a series of sweeps is sent over an actual UWA channel.

The first parameter to be estimated is the timing of the main S2-cycle in a mixture of preamble's multipath arrivals. In mathematical terms, this can be obtained based on maximum likelihood estimation principles, as follows.

Let a received preamble sweep signal be denoted by u and the gradient-heterodyne signal by v. The frequency of the gradient-heterodyne signal is always the value of ω_{IF} higher than that of the transmitted one. The received signal is sampled and observed at times t. In our case, preamble signals are not encoded, i.e., all symbols have the same value.

As defined above [Fig. 4(c)], the main multipath signal (''main spectral line'') can be centered in the filter window $2\Delta\Omega$ by shifting, for example, the received signal in relation to the gradient-heterodyne signal. To spend less of the scan time, rather large shifting steps can be initially used for coarse alignment of the received signals in time. With regard to the $2\Delta\Omega$ window width, this step size must stay, however, under a given boundary to avoid the risk for the IF spectrum to ''skip'' over the $2\Delta\Omega$. So, the step t_{st} must be set slightly less than

$$t_{\mathrm{st\ max}}=\frac{2\Delta\Omega}{m}. \tag{9}$$

Assuming that the noise is Gaussian, the frequency of a strongest spectral line in every scan step is determined by

$$\hat{\omega}(\tau)=\arg \max_{\omega_{\mathrm{IF}}-\Delta\Omega<\omega<\omega_{\mathrm{IF}}+\Delta\Omega} A(\tau,\omega),$$
$$\text{with}\quad \tau=0,t_{\mathrm{st}},\ldots,T_{\mathrm{sw}}, \tag{10}$$

where the amplitude $A(\tau,\omega)$ is defined from the (discrete) Fourier transform as

$$A(\tau,\omega)e^{j\Phi(\tau,\omega)}=FT\{v(t+\tau)u^*(t)\},$$
$$\text{with}\quad \tau=0,t_{\mathrm{st}},\ldots,T_{\mathrm{sw}}. \tag{11}$$

[The phase information $\Phi(\tau,\omega)$ can also be used, when a coherent communication system is considered.] As a result, a vector of most powerful frequencies $\{\hat{\omega}(\tau)\}$, estimated by the scanning with different timing parameters, can be created. Finding the frequency of main multipath signal possessing largest energy concentration among all members of the vector, we can define the coarse timing parameter $\hat{\tau}_c$ as

$$\hat{\tau}_c=\arg \max_{\tau=0,\tau_{\mathrm{st}},\ldots,T_{\mathrm{sw}}} A(\tau,\hat{\omega}(\tau)). \tag{12}$$

After the coarse alignment in time has been accomplished, the timing uncertainty is within one t_{st} interval. Seeking for an accurate timing parameter, an element modifying the coarse timing to a precise value can be found from the frequency difference between the main spectral line and the central frequency of the $2\Delta\Omega$ window such as

$$\hat{\tau}_a=\frac{(\hat{\omega}_c-\omega_{\mathrm{IF}})}{m}. \tag{13}$$

Finally, the timing estimate is

$$\hat{\tau}=\hat{\tau}_c+\hat{\tau}_a. \tag{14}$$

The second parameter to be adjusted is the sweep gradient, which has to ensure a sufficient quality in separation of the main spectral line from the entire IF multipath spectrum. Depending on the channel's multipath time structure, the gradient must be set steep enough to guarantee that the distance between the main spectral line and its adjacent neighbor is large enough for its pure band-pass filtering. To verify this condition, an additional procedure examining the IF spectrum (multipath structure) was implemented. Striving for an opportunity to realize it in the real time, an algorithm representing a combination of two standard functions ''running maximum'' and ''running minimum'' (MatLab5.2, 1998) was applied. Scanning the entire vector with the IF power spectrum, the frequency indexes of all spectral lines together with their energy attributes could be determined. Finally, a feedback message is to be conveyed back to the transmitter, either to redefine the sweep gradient for a given data rate, or, in contrary, to specify another data rate for the gradient initially used.

C. Differential PSK receiver

After filtering out a suitable spectral line, decoding starts. In these experiments, a PLL local oscillator could be efficiently used since the deeply suppressed time-varying multipath arrivals could not seriously influence the loop stability. However, in the off-line processing, a simple differential four-point PSK (QPSK) receiver was used. Demodulation was performed by comparing phase from symbol to symbol. In every signaling interval, phase values were determined applying FFT analyses. The diagrams and BERs later shown were generated by such a demodulator.

V. EXPERIMENTAL RESULTS

In March and August 1999, some 4 GB data were acquired in shallow water channels in a lake (Baggersee near Bremen, Germany) and in the Baltic Sea (Eckenförde Bucht)

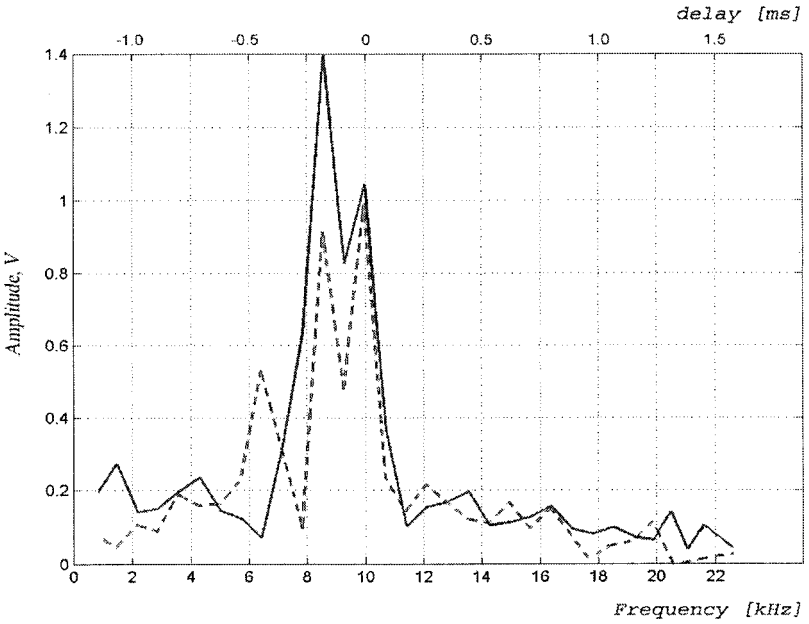


FIG. 5. Example of alterations in the supremacy between two principal arrivals (rate: 1000 symbols per second).

in order to test the method. In the lake, the communication distance was short (108 m), but because of rather small lake dimensions (maximum diameter 300 m), the interferences were especially severe. In the Baltic Sea, the distances were up to 2 km under a relatively flat sea surface, where the multipath structure was represented with a small number of independent interferers. In both cases, the transmitted signals were successfully decoded. The material obtained in the lake was, however, more representative. In the following section, some results of these Baggersee experiments will be presented (with permission of STN Atlas Elektronik GmbH).

The data shown was taken from one of the acquired 1-s data blocks, which were used for off-line processing to evaluate the method's performance.

A. Experimental setup

Transmit and receive transducers were mounted respectively at ~2 and ~13 m below the water surface from stationary platforms. The cite depth changed gradually from ~12 m at transmitter to ~17 m towards receiver. The horizontal radius was about 108 m. The lake conditions were rather uncalm; the wind strength reached about 5–6 m/s, and it was quite noisy because of heavy traffic over a highway near the lake. The source power has been varied, so that different SNRs were available for subsequent analysis. To imitate a source/receiver motion, the transmitting hydrophone was permanently moved during the experiment with a maximum speed of 2 m/s.

The S2-carrier occupied the frequency range from 42 up to 82 kHz, and had a sweep cycle time of 10 ms, thus making a frequency gradient of 4 kHz/ms.

For the broadband communication proposed, a new special broadband transmitter has been used: a model, which has a flat transmitting voltage response (TVR) in the range between 42 and 82 kHz with a maximum deviation of 3 dB from its central value.

As mentioned, this experiment was aimed solely at testing the physical feasibility of the approach, therefore, no high data rates were used: the modulation format was differential QPSK with the bit rate of 2 kb/s. Thus, the symbol duration was exactly 1 ms; every sweep of the sequence carried ten symbols.

The channel probe consisted of three sweeping pulses with an interval $0.5T_{sw}$ between them. The data block durations were from 0.4 up to 1.0 s long. The signal shape was rectangular. The received signal was sampled at a 320 kHz rate.

B. Despreading parameters

The received signal was multiplied with a locally generated version of the S2-carrier—a gradient heterodyne signal—having the same steepness (4 kHz/ms) and the same sweep cycle (10 ms) being produced, however, with an initial frequency of 10 kHz less than that of the transmitted signal (i.e., in range from $\omega_L - 10$ kHz to $\omega_H - 10$ kHz). Hence, the main spectral line of the IF spectrum, being obtained with the correct timing, ought to have its location into a filter window with the central frequency of 10 kHz, as well. A BPF filter, applied to prevent an influence from neighboring spectral lines, was of the least square FIR type and contained 20 coefficients. The filtering window width was 2 kHz, whereby the low and high cut frequencies were set on 9 and 11 kHz, respectively.

C. Channel characterization

To gain insight into the channel characteristics, a series of channel estimations was performed initially by transmitting unmodulated S2-carrier sweeps over the channel. After despreading and low-pass filtering, the IF spectrum reflected the multipath delay spread structure of the channel.

The channel was not constant during the observation time. When looking at the behavior of the principal arrival, it has become obvious that it was not always the same arrival that had the largest energy. This made it possible to have more than one principal arrival in this channel. Figure 5 re-

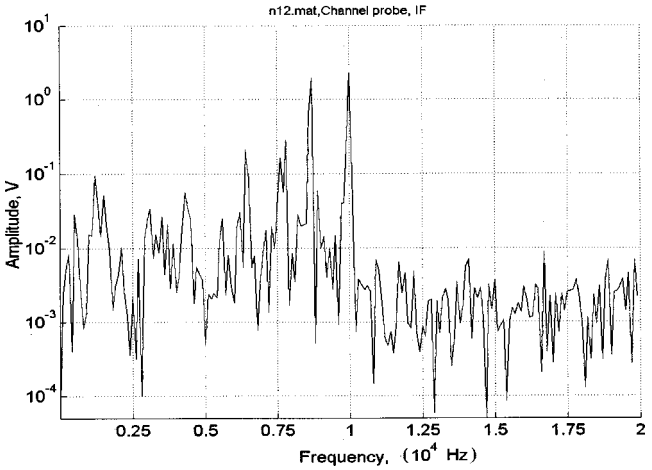


FIG. 6. Spectrum snapshot of the intermediate frequencies (rate: 200 symbols per second).

flects the alterations in the supremacy between two principal frequencies. An important detail is that such alteration can occur within a very short time (40 ms between these two snapshots). Power fluctuations, registered for the both multipath arrivals, were about 2–3 dB. The two dominating frequencies were adjacent and separated 1300 Hz from each other. Figure 6 shows the snapshot of the multipath structure with much better resolution obtained for an unmodulated sweep with the duration of 5 ms (in Fig. 5 the resolution was five times less). As one can see, the structure consisted of about ten multipath arrivals. The delay between the two principal arrivals received was about 320 μs, which was very close to the estimated delay of 325 μs for the surface reflected path if using simple geometry. The total multipath delay spread was estimated about 2.5 ms.

The frequency distances between multipath signals stayed stable during transmission of the data block, which would recall the work of many stations with close frequency spacing. Proceeding from the symbol rate and the distance between the two multipath arrivals—''two adjacent channels''—we normalized the frequency distance and calculated the relative frequency distance between these two interfering ''adjacent channel'' as $d=(1300\ \text{Hz})/(1000$ symbols per second$)=1.3$. Applying the filter to the main multipath signal and relying on calculations made for a case of the interchannel interference in Calandrino *et al.*,[14] this distance ought to be large enough to provide an error rate much below 10^{-4}. This was confirmed also by the experimental investigation of the carrier phase stability.

We observed a superimposition of 200 neighboring 1-ms long symbol intervals evaluated for the main spectral line in the IF range. Since the S2-carrier was transmitted without modulation, all phases of these symbol intervals ought to have approximately equal values, i.e., their differential values should be situated close to zero. The scatter plot of such differential phases is shown in Fig. 7. Since the signals were amplitude limited before their phase was evaluated, the scatter has the differential values located always the same distance from the plot center.

This figure shows utterly tolerable dispersion, whereby the SNR for this case was slightly more than 9 dB. The

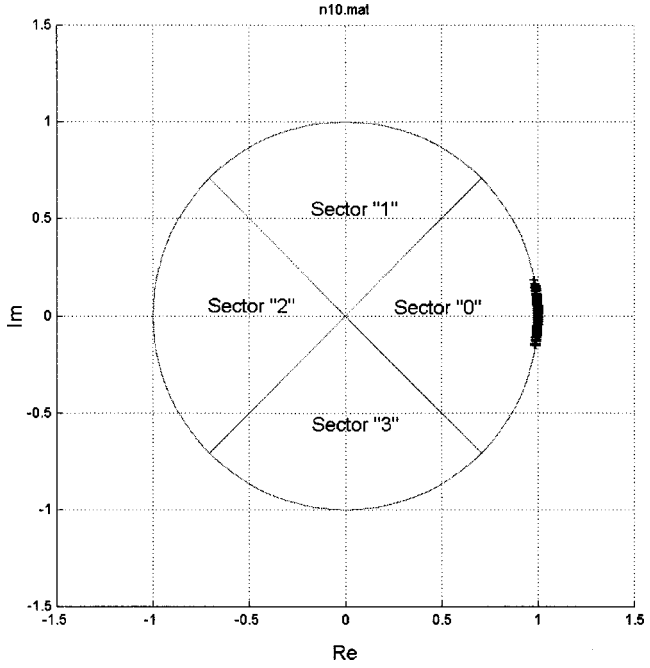


FIG. 7. Phase stability of the unmodulated S2-signals represented in the IF range.

overall dispersion was evaluated to be in order of ± 10 degrees.

D. Performance results

Figure 8 shows the results obtained with differential QPSK signals transmitted over the Baggersee channel using the S2-carrier.

The main multipath signal had again a SNR of slightly more than 9 dB. The component was filtered using the same linear FIR filter as above and phase evaluated. The scatter

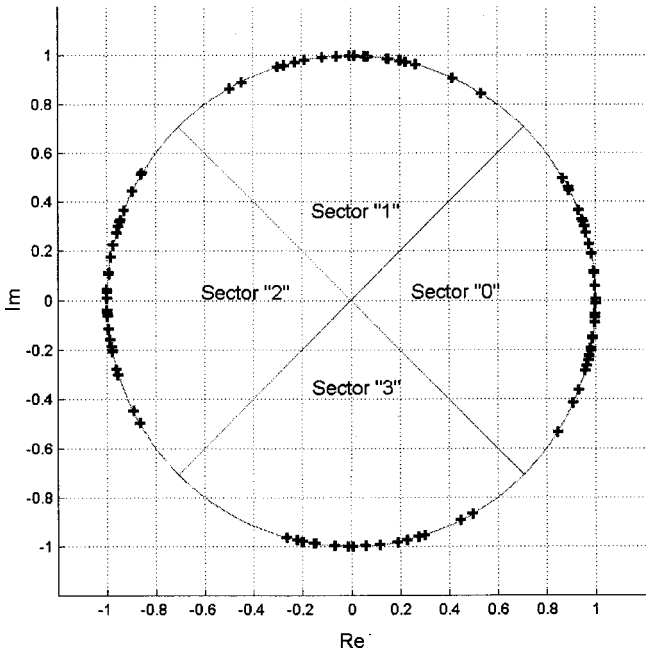


FIG. 8. Output phase scatter of the differential QPSK signals transmitted with the S2-carrier.

plot (Fig. 8) of the estimated data symbols on which the decisions were performed shows completely open eye pattern. No errors were detected in the block of 200 symbols. Generally, in ten trials each with 2000 symbols no error appeared.

Relatively large phase scatter can be explained with a residual interaction between multipath arrivals owing to still nonsufficient separation rate between them.

Although long-trial tests were still not carried out, we could at least roughly evaluate the error probability for this case. Letting the phase dispersions in every sector be normal around its central value, the calculated error probability P_e estimated was 5.8×10^{-5}, thus completely satisfying the usual UWA communication requirement.

VI. DISCUSSION AND CONCLUSIONS

The theoretical as well as the experimental results conclusively demonstrate the successful operation of the method in reducing the effects of multipath propagation in an underwater acoustic telemetry link. Especially in shallow water channels, communication can be substantially improved by the implementation of the S2C method.

The primarily important feature of our method is that filtering occurs not in time but in frequency domains. Using the property of acoustic signals to propagate under water with relatively low speed, the method allows us to extensively space all multipath arrivals in time-frequency area by converting their time delays into their frequency reallocations provided that sufficiently steep sweeps are applied. This is a significant advantage over the state of the art methods.

Although no large data rates were applied, it is straightforward to evaluate an achievable data rate for practical shallow water UWA channels. Let us consider, for example, a typical North Sea channel (mixed water), where it is usual to expect several multipath arrivals with the surface reflection as the strongest one. At distances of about 1–1.5 km and communication depths of 15–30 m, the time delay between direct path and such surface reflection would represent a quantity of usually 0.4–0.5 ms. By carrier sweep gradients of 10–12 kHz/ms, a resolution of the multipath signals would be around 5 kHz. Assuming a symbol rate of also 5 kHz, the speed of exchanges with QPSK signals would have the value of 10 kb/s. Considering the property of very weak correlation between sweeps being generated with opposite directions of frequency change (cross sweeps), it becomes possible to transmit at least two S2-carriers simultaneously in the same frequency range. Thus applying such cross-sweep configuration of the S2-carrier, a bit rate of 20 kb/s at distances of 1–2 km seems to be very realistic.

For any multipath structure, a suitable frequency gradient can be set in the training section, where two tendencies should be taken into account. The higher the desired resolution, the steeper sweeps must be produced. On the other hand, the steeper the sweeps, the shorter the cycle time of an S2-carrier. If the cycle time becomes too short, the current arrivals of the running cycle may interfere with tail arrivals of a previous cycle. A compromise has to be found, however, the sweep cycle T_{sw} must generally exceed the channel's delay spread. In the example of the North Sea channel, the steepness of 10 kHz/ms and the bandwidth of 50 kHz provide the carrier sweep cycle of 5 ms. Since the total delay spread in such a channel represents usually a value of no more than 5 ms, no interference between sweeps of neighboring cycles is to be expected.

It is worthwhile also mentioning that strong ISI encountered in our experimental channel can yield even better performance than in other channels, which do not exhibit the double principal arrival structure. The advantage of the method is the ability of the receiver to exploit natural signal redundancy inherent in consequence of the multipath propagation, e.g., by accurate separation and processing of several (even time-varying) arrivals simultaneously. In this case, much more energy of the received signal can be utilized for parameter evaluation providing increasing statistics for receiver decisions about arriving symbols.

A silent aspect of the approach consists of the ability to considerably improve the SNR by transmitting sweep-spread S2-signals owing to reducing influence of narrow-band noise. Thus, considerable distances can be achieved even when higher frequencies (up to 80–90 kHz) are in use.

ACKNOWLEDGMENTS

Thanks to the STN Atlas Elektronik, Germany, for support and for the opportunity to use its experimental facilities. We also thank Professor Dr. Ingo Rechenberg for hospitality and scientific support and Alexey Kebkal and Iván Sántibáñez-Koref for their technical assistance.

[1] J. A. Catipovich, "Performance limitations in underwater acoustic telemetry," IEEE J. Ocean. Eng. **15**(3), 205–216 (1990).
[2] M. Stojanivic, J. A. Catipoivich, and J. G. Proakis, "Phase-Coherent Digital Communications for Underwater Acoustic Channels," IEEE J. Ocean. Eng. **19**, 100–111 (1994).
[3] M. Stojanivic, "Recent Advances in High-Speed Underwater Acoustic Communication," IEEE J. Ocean. Eng. **21**, 125–136 (1996).
[4] R. Galvin and R. F. W. Coates, "Analysis of the performance of an underwater acoustic communication system and comparison with a stochastic model," in *Proc. Oceans '94*, Brest, France, September 1994, pp. III.478–III.482.
[5] G. S. Howe, P. Tarbit, O. Hinton, B. Sharif, and A. Adams, "Sub-sea acoustic remote communications utilizing an adaptive receiving beam former for multipath suppression," in *Proc. Oceans '94*, Brest, France, September 1994, pp. I.313–I.416.
[6] J. A. Neasham, D. Thompson, A. D. Tweedy, M. A. Lawlor, O. R. Hinton, A. E. Adams, and B. S. Sharif, "Combined Equalization and Beamforming to Achieve 20 kbits/s Acoustic Telemetry for ROVs," in *Proc. Oceans '96*, Ft. Lauderdale, FL, 1996, pp. 988–993.
[7] M. Stojanivic, J. A. Catipoivich, and J. G. Proakis, "Adaptive multichannel combining and equalization for underwater acoustic communications," J. Acoust. Soc. Am. **94**, 1621–1631 (1993).
[8] T. S. Yang and A. Al-Kurd M, "Environmental impact on phase coherent underwater acoustic communications," in *Proceedings 16th International Congress on Acoustics*, Seattle, WA, June 1998, Vol. 1, pp. 301–302.
[9] J. Spiesberger and P. Worcester, "Fluctuations of resolved acoustic multipath at long range in the ocean," J. Acoust. Soc. Am. **70**, 565–576 (1981).
[10] J. Spiesberger, P. J. Bushong, K. Metzger, Jr., and T. G. Birdsall, "Ocean acoustic topography: Estimating the acoustic travel tome with phase," IEEE J. Ocean. Eng. **14**, 108–119 (1989).
[11] O. R. Hinton, G. S. Howe, A. E. Adams, P. Tarbit, and B. Sharif, "Performance of a Stochastic Gradient Adaptive Beamformer for Sub-Sea Acoustic Communications," in *Proc. EUSIPCO'94*, Edinburgh, September 1994, pp. 1540–1543.

[12] S. P. D. Tarbit, G. Howe, O. Hinton, A. Adams, and B. Sharif, ''Development of a real-time adaptive equalizer for a high-rate underwater acoustic data communication link,'' in *Proc. Oceans '94*, Brest, France, September 1994, pp. I.307–I.312.
[13] R. Bannasch and K. Kebkal, Patent PCT/DE99/02628.
[14] L. Calandrino, G. Corazza, G. Crippa, and G. Immovilli, ''Intersymbol, Interchannel and Cochannel Interferencies in binary and quaternary PSK systems,'' Alta Freq. **XL**(5), 407–420 (1971).

Approximating the head-related transfer function using simple geometric models of the head and torso

V. Ralph Algazi and Richard O. Duda[a)]
CIPIC, Center for Image Processing and Integrated Computing, University of California, Davis, California 95616

Ramani Duraiswami, Nail A. Gumerov, and Zhihui Tang
Perceptual Interfaces and Reality Laboratory, Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland 20742

(Received 5 April 2002; accepted for publication 1 August 2002)

The head-related transfer function (HRTF) for distant sources is a complicated function of azimuth, elevation and frequency. This paper presents simple geometric models of the head and torso that provide insight into its low-frequency behavior, especially at low elevations. The head-and-torso models are obtained by adding both spherical and ellipsoidal models of the torso to a classical spherical-head model. Two different numerical techniques—multipole reexpansion and boundary element methods—are used to compute the HRTF of the models in both the frequency domain and the time domain. These computed HRTFs quantify the characteristics of elevation-dependent torso reflections for sources above the torso-shadow cone, and reveal the qualitatively different effects of torso shadow for sources within the torso-shadow cone. These effects include a torso bright spot that is prominent for the spherical torso, and significant attenuation of frequencies above 1 kHz in a range of elevations. Both torso reflections and torso shadow provide potentially significant elevation cues. Comparisons of the model HRTF with acoustic measurements in the horizontal, median, and frontal planes confirm the basic validity of the computational methods and establish that the geometric models provide good approximations of the HRTF for the KEMAR mannequin with its pinnae removed. [DOI: 10.1121/1.1508780]

PACS numbers: 43.64.Bt, 43.66.Qp, 43.66.Pn [LHC]

I. INTRODUCTION

A. Variation of the HRTF with elevation

Head-related transfer functions (HRTFs) are central to spatial hearing, and have been studied extensively (Blauert, 1997; Carlile, 1996; Wightman and Kistler, 1997). The HRTF depends not only on the position of the sound source relative to the listener, but also on the size and shape of the listener's torso, head, and pinnae. The resulting complexity makes its behavior difficult to understand.

In this paper, we investigate the HRTFs for distant sources using very simple geometric models of the head and torso to gain insight into various features observed in acoustically measured human HRTFs. The simplest informative model is the spherical-head model. Introduced by Lord Rayleigh almost a century ago (Strutt, 1907), it has been used by many researchers to explain how the head affects the incident sound field (Hartley and Fry, 1921; Kuhn, 1977, 1987; Brungart and Rabinowitz, 1999). Although this model provides only a crude approximation to a human HRTF, it yields a first-order explanation and approximation of how the interaural time difference (ITD) and the interaural level difference (ILD) vary with azimuth and range.

However, the spherical-head model does not provide any cues for elevation.[1] It is well established that the pinna provides the major source of elevation cues (Batteau, 1967; Gardner and Gardner, 1973; Wright *et al.*, 1974).[2] The effect of the pinna on the HRTF has been studied both experimentally (Mehrgardt and Mellert, 1977; Shaw, 1974, 1997; Wightman and Kistler, 1989) and computationally (Lopez-Poveda and Meddis, 1996; Kahana *et al.*, 1999; Kahana and Nelson, 2000; Katz, 2001). This work shows that the influence of the pinna is negligible below about 3 kHz, but is both significant and complicated at frequencies where the wavelength is short compared to the size of the pinna.

The torso also influences the HRTF and provides elevation dependent information (Kuhn and Gurnsey, 1983; Kuhn, 1987; Genuit and Platte, 1981). Although torso cues are not as strong as pinna cues perceptually, they appear at lower frequencies where typical sound signals have most of their energy. It has been shown that a simple ellipsoidal model of the torso can be used to calculate a torso reflection, and that such reflections provide significant elevation cues away from the median plane, even for sources having no spectral energy above 3 kHz (Algazi *et al.*, 2001a).

However, reflection is a short-wavelength or high-frequency concept, and modeling the effects of the torso by a specular reflection is only a first approximation. Furthermore, as the source descends in elevation, a point of grazing incidence is reached, below which torso reflections disappear and torso shadowing emerges. Rays drawn from the ear to points of tangency around the upper torso define a cone that we call the *torso-shadow cone* (see Fig. 1). Clearly, the specular reflection model does not apply within the torso-shadow cone. Instead, diffraction and scattering produce a qualitatively different behavior, characterized by the attenu-

a)Electronic mail: rod@duda.org

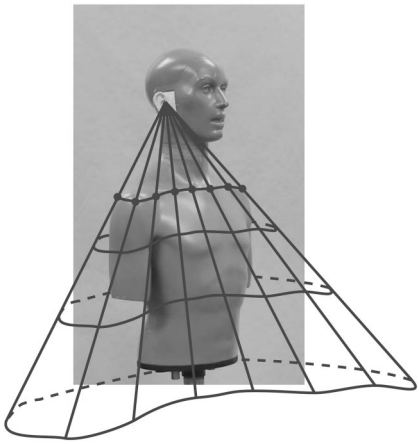


FIG. 1. The torso-shadow cone for the KEMAR mannequin. Rays from the ear are tangent to the torso at the points shown. At wavelengths where ray tracing is valid, specular reflections are produced by the torso for sound sources above the cone. The sound for sources within the cone is shadowed by the torso.

ation of high frequencies when the wavelength is comparable to or smaller than the size of the torso.

There are several reasons why the effect of the torso on the behavior of the HRTF for sources in the complete surrounding sphere has not been systematically measured or studied. First, the lengthy measurement process precludes asking human subjects to stand motionless, and seated measurements at low elevations are influenced by posture and/or the supporting chair, which introduces too many arbitrary variables. Although the use of a mannequin such as KEMAR (Burkhardt and Sachs, 1975) solves this particular problem, at very low elevations a truncated torso introduces meaningless artifacts of its own, and one would have to use a complete mannequin with intact arms and legs. Second, the torso effects appear at relatively low frequencies, where room reflections—even in anechoic chambers—make it hard to obtain accurate measurements. Third, it is experimentally difficult to place sufficiently large loudspeakers in the region directly below the subject. These obstacles have led to a lack of knowledge of HRTF behavior for sources in the torso shadow cone, a lack that may be responsible for the frequent observation that virtual sources synthesized with HRTFs rarely appear to come from really low elevations.

B. Methods for determining the HRTF for the snowman model

To gain a better understanding of the effects of the torso on the HRTF at all frequencies and elevations, a simple head-and-torso model called the snowman model was investigated. In its simplest form, the snowman model consists of a spherical head located above a spherical torso (Gumerov *et al.*, 2002). Unlike the isolated sphere, there is no elegant infinite-series solution for the scattering of sound waves by the snowman model. However, there are at least three ways to obtain the HRTF, all of which are employed in this paper:

(1) acoustic measurements,
(2) numerical computation using boundary-element methods, and
(3) numerical computation using multipole reexpansions.

Each of these approaches has its characteristic advantages and disadvantages, which are summarized in turn.

Acoustic measurements entail no mathematical idealizations, are accurate over much of the audible frequency range, and can produce both HRTFs and head-related impulse responses (HRIRs) equally easily. However, room reflections make it difficult to measure the response at very low frequencies, physical constraints can make it difficult to position a loudspeaker at very low elevations, and measurement and alignment errors make it difficult to get the repeatability needed to study systematically the effects of changing the values of snowman parameters.

By contrast, the computational methods used in this study work particularly well at low frequencies, can be used for any source location, and are well suited to systematic parametric studies. However, they employ idealized assumptions, require validation, and have time/accuracy tradeoffs that limit the highest frequencies that can be used. Being frequency-domain methods, they provide the HRTF directly, but they require the computation of the HRTF at a large number of linearly spaced frequencies to invert the Fourier transform and extract the HRIR.

Boundary-element methods (or similar finite-difference and finite-element methods) can be applied to an arbitrarily shaped boundary surface (Ciskowski and Brebbia, 1991). However, the continuous surface must be approximated by a discretely sampled mesh of points in three dimensions, spaced at roughly one-tenth of the shortest wavelength of interest. It is a challenge to obtain a sufficiently accurate mesh for the human torso, head and pinnae, even without taking the possible effects of hair into account. Furthermore, to determine the response at high frequencies requires very dense sampling and correspondingly long computation times (Katz, 2001).

The multipole reexpansion method used in this paper is similar to the T-matrix method (Waterman and Truell, 1961), and it extends the classical infinite series solution for a single sphere (Morse and Ingard, 1968) to scattering by multiple spheres. The technique used employs new expressions for reexpansion of multipole solutions (Gumerov and Duraiswami, 2001a). Coupled with a procedure for enforcing boundary conditions on the sphere surfaces, it can be used to solve multiple scattering problems in domains containing multiple spheres (Gumerov and Duraiswami, 2001b). No meshes are required. Although reexpansion requires the use of numerical methods to solve the linear equations that define the boundary conditions, space and frequency can be sampled with arbitrarily fine resolution. In the particular case where the spheres are coaxial, multipole reexpansion can be several orders of magnitude faster than boundary-element methods. However, in the current version, convergence problems limit the highest frequencies that can be investigated.

In this paper, each method is used for a different purpose. Acoustic measurements are used to obtain the HRTF for the KEMAR mannequin and to validate the numerical methods. Multipole reexpansion is used for systematic studies of the snowman with a spherical torso. The boundary-

FIG. 2. The physical snowman model, which is composed of a 4.15-cm-radius boccie ball resting on top of a 10.9-cm-radius bowling ball. The probe tube microphone is inside the boccie ball, with the probe tip flush with the surface. The bowling ball is supported by a 0.5-cm-radius cylindrical rod.

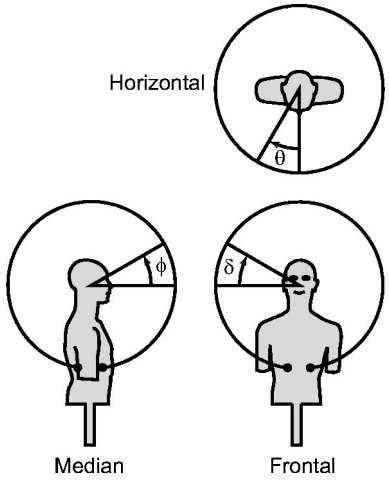


FIG. 3. The coordinate planes. In the horizontal plane, the azimuth θ ranges from 0 to 360 degrees. Support structures limit the range of the experimentally measurable elevation angles ϕ and δ. The entire 360-degree range is covered in the computational solutions.

element method is used for the snowman with an ellipsoidal torso. In the process, we (a) mutually validate the computational methods used, (b) identify the features of HRTFs for simple geometric models of the head and torso, (c) evaluate the adequacy of the snowman as an approximation to the human head and torso, and (d) use the snowman model to reveal the first-order effects of the torso and to identify possible localization cues.

II. METHODS

A. Measurement procedure

Acoustically measured HRTFs were obtained for two objects: the KEMAR mannequin shown in Fig. 1 and the physical snowman model shown in Fig. 2. Because head and torso effects are obscured by the presence of the pinnae, KEMAR's pinnae were removed and the exposed cavities were filled with putty and tape. Two Etymotic Research ER-7C microphones were placed inside the head, with the probe tips emerging at the entrance of the ear canals flush with the surface of the head. The Golay-code technique was used to measure the HRIRs (Zhou *et al.*, 1992). The test sounds were played through 3.2-cm-radius Bose Acoustimass™ Cube speakers mounted on a 1-m-radius hoop that was rotated about a horizontal axis through the midpoint of the interaural axis. The sampling rate for the measurements was 44.1 kHz. To remove room reflections, the resulting impulse responses were windowed using a modified Hamming window that eliminated everything occurring 2.5 ms after the initial pulse. The windowed responses were free-field equalized to compensate for the loudspeaker and microphone transfer functions. Because the small loudspeakers used were inefficient radiators at low frequencies, the low-frequency signal-to-noise ratio was poor, and it was not possible to completely restore the response below 500 Hz. As a result, measured HRTF values below 500 Hz should either be treated with suspicion or ignored.

To extend the useful range by another octave, the radius of the head of the physical snowman was made to be about half that of a human head, and the results were subsequently scaled in frequency accordingly. Specifically, the physical snowman model consisted of a 4.15-cm-radius boccie ball resting on top of a 10.9-cm-radius bowling ball, with a small collar added to keep the head from rolling off. The model was supported by a 0.5-cm-radius metal rod. The boccie ball was drilled to accommodate one ER-7C microphone, with the probe tip emerging at a horizontal diameter.

For both KEMAR and the physical snowman model, measurements were made in the horizontal, median and frontal planes (see Fig. 3). The hoop was rotated in uniform steps of 360 degrees/128≈2.8 degrees. For horizontal-plane measurements, the azimuth θ covered a full 360 degrees. For median-plane and frontal-plane measurements, the hoop support structure limited the elevation angles ϕ and δ to the interval from -81.6 to $+261.6$ degrees, so that no measurements could be made in a ± 8.4-degree cone directly below the subject.

B. Computational procedures

Both of the computational methods used in this paper solve the Helmholtz equation (the Fourier transform of the wave equation) at a specified frequency in an infinite domain containing one or more scattering bodies. The Helmholtz equation is given by

$$\nabla^2 p + k^2 p = 0, \tag{1}$$

where p is the Fourier transform of the acoustic pressure, $k = \omega/c$ is the wave number, ω is the circular frequency, and c is the speed of sound. The incident pressure field p_{inc} is typically prescribed as the field from an isotropic point source, and the goal is to compute the scattered field, $p_{\text{scat}} = p - p_{\text{inc}}$, subject to boundary conditions at the surfaces of the scatterers and at infinity. We assume that the surfaces are "sound hard" ($\partial p/\partial n = 0$), and that the scattered sound field is outgoing at infinity.

When there is a single spherical scatterer, the scattered field at a point specified by the spherical coordinates (r, θ, ϕ) can be written in the form

$$p_{\text{scat}} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} h_l(kr) Y_{lm}(\theta, \phi), \tag{2}$$

where $h_l(\cdot)$ is the lth-order spherical Hankel function, $Y_{lm}(\cdot,\cdot)$ are the spherical harmonics, and the coefficients a_{lm} are determined by requiring that $p = p_{\text{inc}} + p_{\text{scat}}$ satisfies the boundary conditions on the sphere. Such a solution was used by Duda and Martens (1998) to represent the HRTF of a spherical head. If there are N spheres in the domain, one can exploit the linearity of the Helmholtz equation and write the solution as

$$p_{\text{scat}} = p_1 + p_2 + \cdots + p_N, \tag{3}$$

where

$$p_j = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm}^j h_l(kr_j) Y_{lm}(\theta_j, \phi_j). \tag{4}$$

Each of the functions p_j is centered at the corresponding sphere, and is expressed in a local spherical coordinate system. These series are truncated at some finite number of terms, and the coefficients a_{lm}^j are found by requiring that the boundary conditions at the surface of each sphere be satisfied. The procedure for doing this using multipole translation and reexpansion is presented in Gumerov and Duraiswami (2001b).

When the scattering surfaces are not spherical and the multipole reexpansion technique cannot be used, the boundary element method is used instead. This method works by using Green's identity to write Eq. (1) as an integral equation for the acoustic pressure. This equation specifies the pressure at a *field* point X on the surface of the acoustic domain as

$$C^X p^X = \int_{\Gamma_Y} \left[G^{XY} \frac{\partial p^Y}{\partial n^Y} - \frac{\partial G^{XY}}{\partial n^Y} p^Y \right] d\Gamma_Y, \tag{5}$$

where Γ is the surface of the acoustic domain, n is the unit outward normal vector to the acoustic domain at a surface (*source*) point Y, G is the free-space Green's function, and C is the jump term that results due to the treatment of the singular integral involving the derivative of the Green's function. The Green's function for the three-dimensional free-space problem, expressed in terms of the wave number k and the distance r between the source and field points, is

$$G^{XY} = \frac{\exp\{-ikr\}}{4\pi r}, \tag{6}$$

where $i = \sqrt{-1}$ is the complex constant. The surface of the scatterers is discretized using plane triangular elements. The equation is written at each boundary element, and a linear system of equations is obtained, which can be symbolically represented as

$$[F]\{P\} = [G]\left\{\frac{\partial P}{\partial n}\right\}, \tag{7}$$

where $[F]$ and $[G]$ are matrices whose coefficients are obtained by evaluating integrals involving $\partial G/\partial n$ and G kernels, respectively; $\{P\}$ is the vector of acoustic pressures at the surface nodes, and $\{\partial P/\partial n\}$ is the vector of normal derivatives of the pressure. Imposition of the boundary conditions leads to a system of linear equations that can be solved for the pressure.

Both the multipole reexpansion method and the boundary-element procedure yield the magnitude and phase of the HRTF at a particular frequency f. Computational time/accuracy tradeoffs limit the maximum allowable value of ka, where a is the head radius and $k = 2\pi f/c$ is the wave number, c being the speed of sound. For the snowman models that were investigated, the maximum useable value of ka was approximately 10. For the standard 8.75-cm head radius, this corresponds to a maximum frequency of about 6 kHz. To be conservative, all HRTF calculations were limited to exactly 5 kHz.

To obtain the HRIRs, the HRTFs for 500 frequencies uniformly spaced from 0 Hz to 5 kHz were calculated, and the ifft function in MATLAB™ was used to calculate the inverse discrete Fourier transform. Because this procedure implicitly assumes that values of the HRTF above 5 kHz are all zero, direct use of the inverse transform leads to significant and distracting Gibbs phenomenon ripples in the impulse response. For graphical display, these ripples were removed by applying a standard Hamming window to the magnitude spectrum, leaving the phase unchanged (Oppenheim and Schafer, 1969). Like low-pass filtering, windowing smooths the impulse response, reducing the height and increasing the width of pulses. For the window used, a unit impulse that would have a duration of 0.2 ms when sharply band limited to 5 kHz has its peak height approximately halved and its duration approximately doubled. Although this results in some loss of information, it greatly increases the clarity of the graphs.

III. THE FRONTAL-PLANE HEAD-AND-TORSO HRTF

To investigate the behavior of the HRTF, we start with the familiar spherical-head model and subsequently consider a sequence of progressively more complex cases. Because the results for the frontal plane exhibit a greater variety of behavior, we focus on it first.

A. The spherical head model

Figure 4 shows two image representations of the magnitude of the right-ear, frontal-plane HRTF for a sphere having the standard 8.75-cm head radius. These particular images were created using the algorithm presented in Duda and Martens (1998), but the same results were also produced by both the multipole reexpansion code and the boundary element code. Brightness corresponds to dB magnitude as shown by the grayscale bar at the top of the image. In Fig. 4(a), each vertical line corresponds to the frequency response at a particular elevation angle, δ. At low frequencies the response is 0 dB for any elevation angle. The largest response (which is approximately 6 dB at high frequencies) occurs at $\delta = 0$ degrees, where the source points directly at the right ear. As expected, the response is generally large on the ipsilateral side (-90 degrees $< \delta <$ 90 degrees) and small on the contralateral side (90 degrees $< \delta <$ 270 degrees). However, on the contralateral side the sphere exhibits a "bright spot," which appears in Fig. 4(a) as a bright vertical streak centered at $\delta = 180$ degrees. The dark bands on each side of this

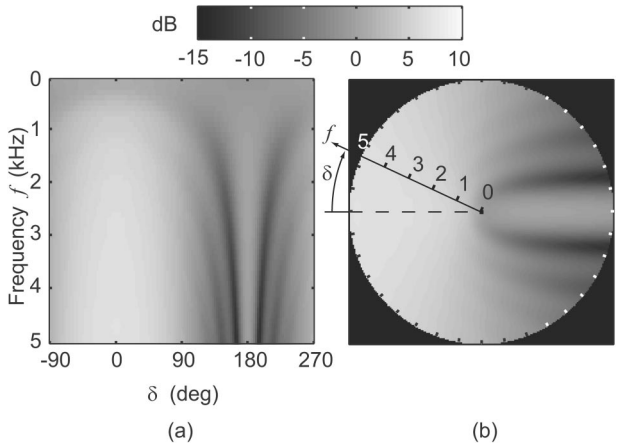


FIG. 4. Image representations of the magnitude of the HRTF for an ideal rigid sphere. The magnitude in dB is represented by brightness. The data is for the right ear of an 8.75-cm-radius sphere. The images are for the frontal plane (see Fig. 3). In (a), the left half of the image corresponds to the ipsilateral side, and the right half to the contralateral side. Note the bright spot that appears as a vertical streak at $\delta = 180$ degrees. (b) shows the same data in polar coordinates. The elevation angle δ corresponds directly to the frontal view in Fig. 3. Thus, once again the left half corresponds to the ipsilateral side and the right half to the contralateral side, but the bright spot appears as a broad horizontal streak.

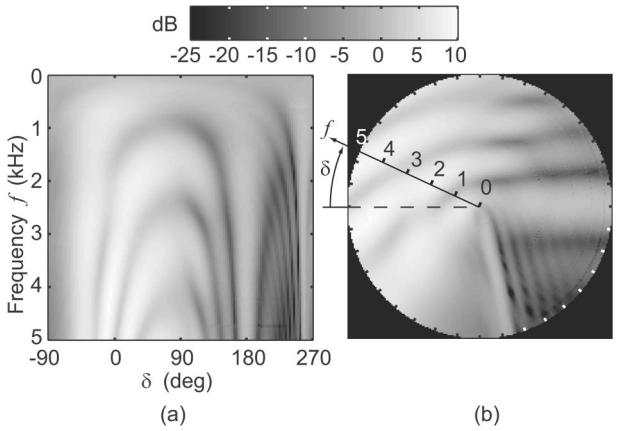


FIG. 6. The computed HRTF for the physical snowman model shown in Fig. 2, scaled for a head radius of 8.75 cm. The arch-shaped notches that are symmetric about $\delta = 90$ degrees are due to specular reflections from the upper torso. The deeper notches around 210 to 250 degrees are caused by torso shadow. A torso bright spot can be seen around $\delta = 255$ degrees.

streak are interference patterns whose regularity is due to the perfect symmetry of the sphere; more irregularly shaped surfaces have the same general behavior, but the interference patterns become smeared.

Figure 4(b) is a useful alternative remapping of the information in Fig. 4(a). In this polar plot, frequency ranges from 0 to 5 kHz along any radial line. The center of the image corresponds to $f = 0$, where the response is exactly 0 dB. The frequency response for incident sound waves arriving at an angle δ is found along the radial line at the angle δ as shown. This puts the HRTF display into direct correspondence with the coordinate system shown in Fig. 3.

The HRIR, which includes both the magnitude and the phase response of the HRTF, is particularly useful for revealing multipath components of the response. Figure 5(a) shows a family of HRIR curves corresponding to Fig. 4. Note that the ipsilateral response (-90 degrees $< \delta <$ 90 degrees) is not only stronger than the contralateral response (90 degrees $< \delta <$ 270 degrees), but it also occurs sooner. The approximately 0.7 ms difference in the arrival times at $\delta = 0$ degrees and $\delta = 180$ degrees is the maximum ITD. The bright spot appears as a local maximum in the response at $\delta = 180$ degrees. In the vicinity of the bright spot, one can see a second pulse that follows the first pulse [see Fig. 5(b)]. It is this second pulse that is the source of the interference patterns seen in the frequency domain. A rough interpretation is that one pulse is composed of wave components traveling around the ipsilateral side of the sphere, and the other is composed of components traveling around the contralateral side; their in-phase confluence at $\delta = 180$ degrees is the source of the bright spot.[3]

B. The physical snowman model

We now examine the effects produced by the introduction of the torso. Using the same 8.75-cm head radius, a 23-cm spherical torso is added directly below and tangent to the head. This results in a ratio of head size to torso size that is the same as that for the physical snowman shown in Fig. 2.

The frontal plane HRTF computed by the multipole reexpansion method is shown in Fig. 6. Comparison of Figs. 4(a) and 6(a) reveals two major differences. First, three arch-shaped notches centered near $\delta = 80$ degrees appear on the ipsilateral side. The lowest-frequency notch occurs around 1 kHz at $\delta = 80$ degrees, where the response dips to -5 dB. As was shown in Algazi *et al.* (2001a), these elevation-dependent notches are comb-filter interference patterns

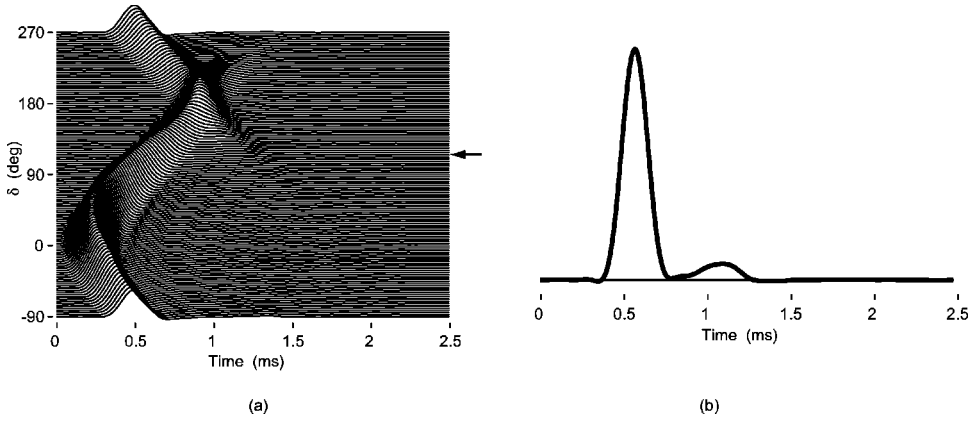


FIG. 5. (a) The HRIR for the sphere. The waveforms have a 5-kHz bandwidth, and the spectrum was smoothed with a Hamming window before inversion. The bright spot appears as a local maximum at $\delta = 180$ degrees. (b) An expanded plot of the impulse response at $\delta = 150$ degrees. This illustrates that near the bright spot the impulse response is bimodal. The weaker peak can be attributed to waves that travel around the contralateral side of the sphere. This second pulse is the source of the interference patterns seen in Fig. 4.

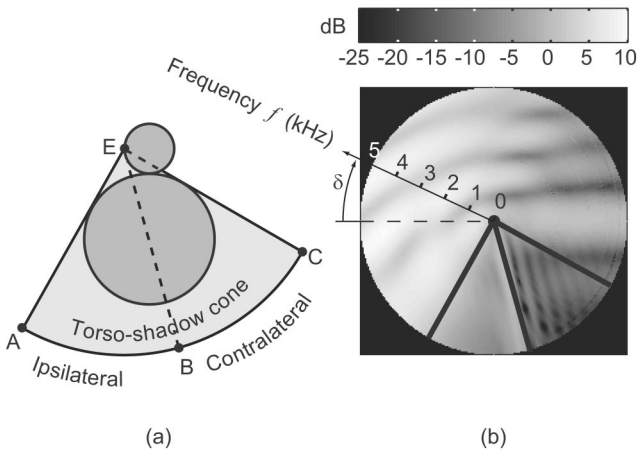


FIG. 7. Boundaries of the torso-shadow cone for the physical snowman model. The bisecting ray BE through the center of the torso to the right ear shown in (a) defines the direction of the torso bright spot. The tangent rays AE and CE are shown superimposed on the computed HRTF in (b). There is a close correspondence between the geometrical boundaries and the region of reduced response.

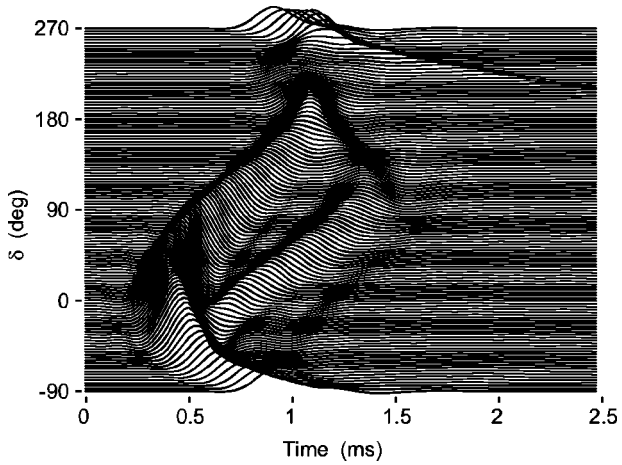


FIG. 8. The computed HRIR for the physical snowman. The torso reflection is prominent when the elevation δ is between -30 and 150 degrees. The head bright spot can be seen near $\delta=180$ degrees, and the torso bright spot near $\delta=255$ degrees. The response for δ between 200 and 250 degrees is flattened and broadened by torso shadow. The long thin ''tail'' in this region is responsible for the strong interference notches seen in Fig. 6(b). A symmetric tail for δ between -90 and -40 degrees is barely visible.

caused by torso reflection. They extend throughout the audible frequency range, and provide cues for elevation away from the median plane.

A second major difference is the appearance of deeper and more closely spaced notches on the contralateral side below the horizontal plane, where δ ranges from roughly 195 to 250 degrees. These low-elevation notches, which form a pattern of parallel lines in Fig. 6(b), fall in the torso-shadow cone; combined with head shadow, they cause the response for frequencies above 1 kHz to be much lower on the contralateral side than on the ipsilateral side.

Somewhat surprisingly, the lowest response does not occur when the source is directly below ($\delta=-90$ or $+270$ degrees). Instead, another bright spot appears at very low elevations. This ''torso bright spot'' is particularly clear in Fig. 6(b), where it forms a bright radial ridge near $\delta=255$ degrees. Thus, the snowman model exhibits two bright spots, one due to the head around $\delta=180$ degrees, and one due to the torso around $\delta=255$ degrees.

There is a simple explanation for the torso bright spot. If a sound source below the torso were directed along a line from the center of the torso to the location of the right ear, and if the head did not disturb the sound field, a bright spot would be formed on the contralateral side of the torso and would strongly ''illuminate'' the right ear. For the dimensions of the physical snowman, the elevation angle for this line is 254.6 degrees, which is consistent with this interpretation.

The torso-shadow cone for the physical snowman is shown drawn to scale in Fig. 7(a). The ipsilateral limit is defined by the ray AE tangent to the torso on the ipsilateral side, and the contralateral limit is defined by the ray CE tangent to the torso on the contralateral side. BE, the ray through the center of the torso to the right ear, bisects the torso-shadow region. Sources in the ipsilateral zone are shadowed by just the torso, while sources in the contralateral zone are shadowed by both the torso and the head. In Fig. 7(b) these three rays are shown superimposed on the polar HRTF plot of Fig. 6(b). Although the correspondence is not perfect, AE is closely aligned with the edge of reduced ipsilateral response, BE is closely aligned with the torso bright spot, and CE is roughly aligned with the edge of the even more reduced contralateral response. Thus, the geometric torso-shadow cone is in good agreement with the zones of reduced response. A similar geometric argument helps to explain why the torso reflection notches attain their lowest frequencies near $\delta=80$ degrees. If the head is removed, the diagram in Fig. 7(a) is symmetric about the ray from B to E. Outside of the torso shadow cone, a pulse of sound directed to the ear at E is followed by a subsequent torso reflection. For pulses directed along the tangent ray from A to E or along the tangent ray from C to E, the delay between the initial pulse and the reflection is zero. By symmetry, the maximum time delay occurs for an overhead ray directed from E to B, and this leads to the lowest frequency for the interference notch. For the dimensions of the physical snowman, this overhead ray would be found at $\delta=74.6$ degrees. Although the head disturbs the symmetry, this argument explains why the angle for the lowest notch frequency is biased below 90 degrees.

Additional insight can be obtained by comparing the HRIRs for the sphere and the snowman. The 5-kHz bandwidth HRIR, computed by inverting the multipole reexpansion, is shown in Fig. 8. Comparing these results to those for the isolated sphere [Fig. 5(a)], three prominent differences can be seen. First, the torso reflection is clearly present in the snowman HRIR in the general range of elevations from about -30 to $+150$ degrees. The maximum time delay between the main pulse and the torso reflection occurs when the source is overhead, and is approximately $\Delta T=0.7$ ms. In the frequency domain, this corresponds to the first notch that occurs at $f_0=1/(2\Delta T)=700$ Hz. This value is in good agreement with the location of the lowest frequency arch at $\delta=80$ degrees in Fig. 6(a).

Second, the response in the interval from 200 to 250

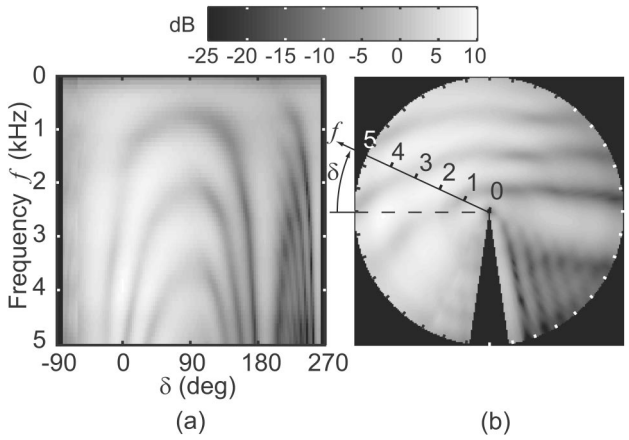


FIG. 9. The measured HRTF for the physical snowman model shown in Fig. 2, scaled for a head radius of 8.75 cm. These results should be compared with the computational results in Fig. 6.

degrees is significantly broader and flatter for the snowman than for the isolated sphere. This corresponds to the presence of torso shadowing on the contralateral side. Third, in this same interval there is a long, thin ''tail'' whose arrival time changes rapidly with elevation. Like the torso reflection, this second pulse produces notch filter interference patterns, and is the source of the deep notches that appear in the contralateral torso shadow zone in Fig. 6. This presence for this tail can be explained by imagining that waves traveling over the torso from the contralateral side of the torso-shadow cone can be divided into two groups. The group that arrives first travels over the contralateral side [along the line CE in Fig. 7(a)], where it is further shadowed by the head, while the group that arrives later takes a longer path over the ipsilateral side [along the line AE in Fig. 7(a)], and is not subject to head shadow. Although this is a crude explanation of a complicated diffraction and scattering phenomenon, it also explains why the same ''tail'' is much less prominent when the source is on the ipsilateral side, where the first arriving group of waves is not subject to head shadow but the second arriving group (the ''tail'') is attenuated by head shadow.

IV. VALIDATION OF THE MULTIPOLE REEXPANSION TECHNIQUE

To confirm the validity of the multipole reexpansion method and of the characteristics of the snowman HRTF just presented, we now compare those results to the results of acoustic measurements of the physical snowman model shown in Fig. 2. Figure 9 shows the measured HRTFs, scaled in frequency to match the standard 8.75-cm head radius. As was mentioned in Sec. II A, experimental constraints limited the elevation angles to the interval from -81.6 to $+261.6$ degrees; regions for which no data are available are shown in solid black. Ignoring those regions, a comparison of Figs. 6 and 9 shows a generally very good correspondence between the computational and the acoustic results.

In particular, the torso reflections and the bright spot for the head are in excellent agreement. There is also good general agreement regarding the behavior in the contralateral torso shadow region, including the torso bright spot. The main difference is that the interference notches are a bit deeper and more closely spaced in the computational data than in the measured results. Because such interference patterns are quite sensitive to small changes in path lengths, this difference could be caused by any of a number of imperfections in the experimental conditions, such as the presence of a supporting rod, the presence of the collar, or small angular misalignments.

Comparison of the HRIRs reveals some differences that are not as evident in the frequency domain. As was mentioned in Sec. II A, it was not possible to compensate for the low response of the loudspeakers below 500 Hz without exposing low-frequency noise. The measured data is effectively a high-pass-filtered version of the true responses, with a 3-dB corner frequency around 500 Hz. This is revealed by a slight darkening at the top of Fig. 9(a) that is not visually prominent in the frequency domain, but that shows up clearly in the time domain as a negative overshoot following the main pulse. The high-pass filtering hides the low-frequency noise that is present in the measured data.

Thus, in the comparison of the computed and the measured impulse responses shown in Fig. 10, the computed response was filtered by a single-pole high-pass filter with a 500-Hz corner frequency to introduce a comparable distortion to the curve [see Fig. 10(a)]. In addition, a Hamming window was used to band-limit the measured impulse responses to 5 kHz to account for the lack of high-frequency energy in the computed results [see Fig. 10(b)]. With these corrections understood, there is again very good agreement between the computed and the measured data. This gives confidence that the multipole reexpansion technique is providing correct HRTF values in both magnitude and phase.

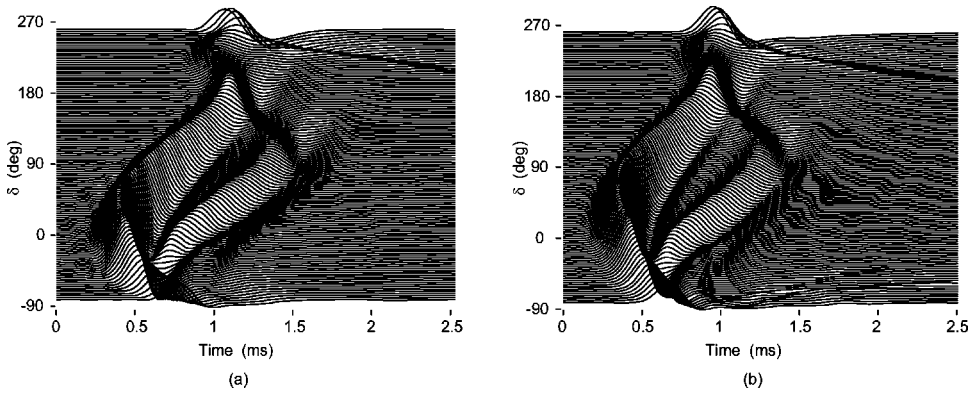


FIG. 10. Comparison of the (a) computed and (b) measured impulse responses for the physical snowman. The computed HRIR was high-pass filtered to simulate the effects of loss of low-frequency information in the measured data. This accounts for the negative overshoot following the main pulse. Similarly, the measured HRIR was low-pass filtered (bandlimited to 5 kHz) to simulate the absence of high-frequency information in the computed results. The range of elevations is limited to the range for which measurements are available.

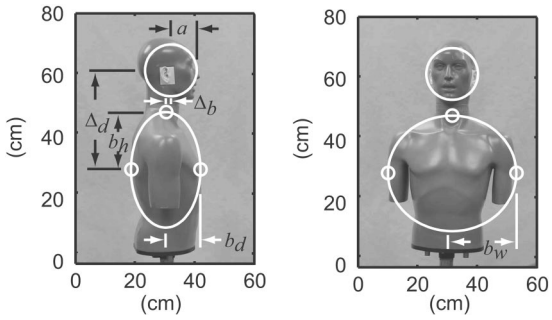


FIG. 11. Fitting a spherical head and ellipsoidal torso to the KEMAR mannequin.

V. EVALUATING THE SNOWMAN MODEL

To investigate how well a human HRTF can be approximated by the HRTF of the snowman model, we decided to compare the HRTF for the KEMAR mannequin shown in Fig. 1 to two snowman models, one with a spherical torso and one with an ellipsoidal torso. This process involves two steps: (a) geometrically fitting the snowman models to KEMAR, and (b) comparing the resulting HRTFs. Each process is explained in turn.

A. Fitting the snowman model to KEMAR

The optimum snowman model for KEMAR is the one whose HRTF best matches KEMAR's HRTF. However, there is no obvious or well-established measure of error in comparing two HRTFs. As a consequence, it was decided simply to match anthropometric characteristics. The head and torso components were fit separately, placing more emphasis on fitting the upper torso than the lower torso. Although better results could be obtained with a more sophisticated fitting procedure, the method used had the virtue of being easy to understand and apply.

The spherical head is defined by its center and radius. The center was located at the center of a bounding box for KEMAR's head (see Fig. 11). When the regression equation in Algazi *et al.* (2001b) was used to estimate the optimum head radius, a value of 8.70 cm was obtained, which is remarkably close to the standard 8.75-cm value. To fit the torso, the five control points shown as open circles in Fig. 11 were used.[4] This was almost sufficient to define an ellipsoidal torso model whose principal axes were parallel to the coordinate axes, but left one degree of freedom free. This was resolved by arbitrarily placing the center of the ellipsoid at the intersection of the lines joining the left/right and front/back control points.

The resulting spherical head and ellipsoidal torso are shown in Fig. 11. The dimensions for this ''best-fitting'' model were as follows:

Head radius, a 8.7 cm

Torso half height, b_h 19.3 cm

Torso half width, b_w 21.5 cm

Torso half depth, b_d 11.6 cm

Torso geometric mean radius, b 16.9 cm

Torso displacement from head, back, Δ_b 1.7 cm

Torso displacement from head, down, Δ_d 30.9 cm.

The resulting snowman model provides just a first approximation to KEMAR. The spherical head is wider, shorter, and shallower than the real head. The ellipsoidal torso matches the real torso fairly well on top, but is badly mismatched below the control points. This was an intentional compromise, and was made for three reasons. First, an ellipsoid that would provide a better match to the lower torso would necessarily sacrifice the more important upper region. Second, in the torso shadow region where the match is poor, KEMAR is no longer a realistic model for a human. Third, at very low elevations, the wavelengths that are unattenuated by torso shadow are sufficiently long that there is reason to hope that a crude fit will be adequate and that the specific shape is relatively unimportant. Finally, the head is not tangent to the torso, leaving a gap where the neck should be. Although a more realistic model would include a neck component, it was omitted in the interest of simplicity.

For an even simpler approximation, the head was centered over the torso, and the ellipsoid was approximated by a sphere whose radius (16.9 cm) was the geometric mean of the semiaxes of the ellipsoid, so that the volume was unchanged. The resulting spherical torso is a coarser approximation, exhibiting a symmetry about the vertical polar axis that is not realistic. However, it provides a model that is easy to understand and evaluate.

B. Comparing the KEMAR and the snowman HRTFs

We now compare the pinnaless KEMAR HRTF to that of the spherical-torso and ellipsoidal-torso models. In all cases, the KEMAR HRTFs were measured acoustically, the spherical-torso results were computed using the multipole reexpansion method, and the ellipsoidal-torso results were computed using the boundary-element method. We compare frontal-plane results in some detail, and compare horizontal-plane and median-plane results more briefly. We start by comparing the KEMAR HRTF to the HRTF for the spherical-torso approximation, and then examine the improvement that the ellipsoidal-torso model can provide.

1. Frontal plane

The frontal-plane KEMAR HRTF is shown in Fig. 12(a), and the corresponding spherical-torso HRTF is shown in Fig. 12(b). The gross characteristics are very similar. In particular, in both cases one can see the presence of a strong response on the ipsilateral side, torso reflection notches at upper elevations, head shadowing on the upper contralateral side, and even stronger torso shadowing on the lower contralateral side. Thus, the model provides an approximation to KEMAR that captures important, first-order effects. However, there are numerous differences as well, caused either by the simplicity of the model or the choices made in fitting the model to KEMAR.

On the upper contralateral side, the head bright spot is almost as well defined for KEMAR as it is for the model. However, it occurs around $\delta = 155$ degrees for KEMAR versus 180 degrees for the model, which is a 25-degree shift.

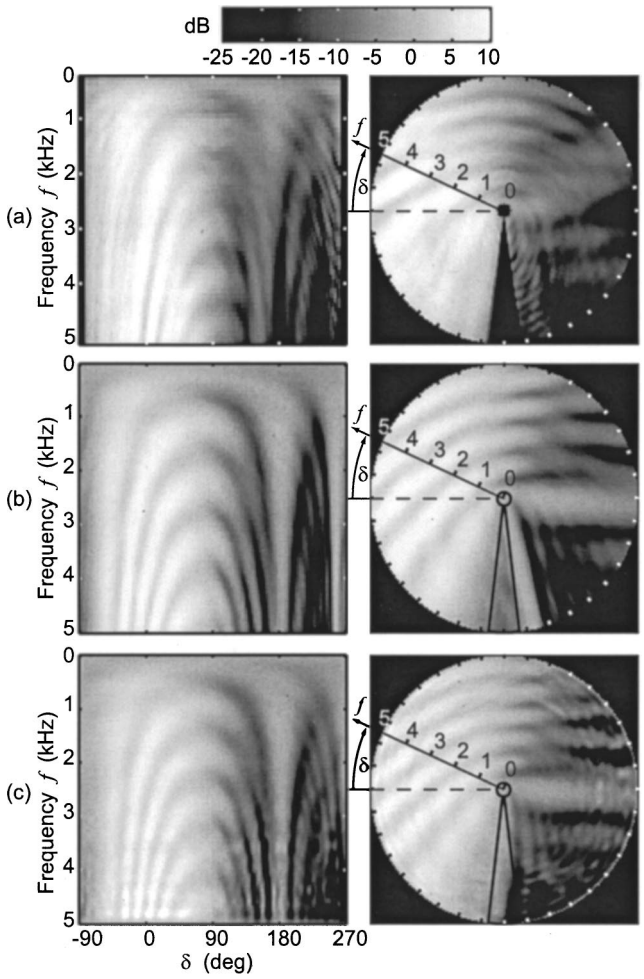


FIG. 12. Measured and computed frontal-plane HRTFs. (a) KEMAR with the pinnae removed. (b) The spherical-torso snowman approximation. (c) The ellipsoidal-torso snowman approximation.

The reason for this shift can be traced to the way that the spherical head was fit to KEMAR's head. The actual top of KEMAR's head is about 3 cm above the top of the sphere (see Fig. 11). If the spherical head were to be shifted up 3 cm without moving the z coordinate of the ears in space, the ears would no longer be across a diameter, but would be down by $\sin^{-1}(3/8.7) \approx 20$ degrees, which would introduce the corresponding shift in the position of the bright spot and would account for most of the discrepancy. However, shifting the head up results in a poorer fit to the data at low elevations. If this difference is important, a better solution would be to use an ellipsoidal head model (Duda *et al.*, 1999).

On the upper ipsilateral side, the torso-reflection arches for KEMAR and the model are very similar, with the highest-frequency fifth arch being even better defined for KEMAR than for the model. However, there is a systematic shift in the elevation at which the notches reach their lowest frequency, being around $\delta = 95$ degrees for KEMAR and around 75 degrees for the model. This shows up clearly as a 20-degree rotation of the torso reflection notches in the polar plots. Using the dimensions for the model, the heuristic argument given in Sec. III B predicts that the elevation for the lowest frequency should appear at $\delta = 90$ degrees $-\tan^{-1}(8.7/30.9) \approx 74$ degrees, which is in close agreement with what is observed. The exact reason why this point is different for KEMAR is not completely understood, but it is probably a combination of the effect of the neck and a mismatch between the surface orientation of the ellipsoid and the top of the torso in the vicinity of the neck.

Although a small zone directly below KEMAR could not be measured, the ipsilateral torso shadow that is visible matches the model well. The contralateral torso shadow is more complex for KEMAR than it is for the model, and—not surprisingly—there is no sign of a torso bright spot. The new interference patterns that appear for $\delta > 240$ degrees are undoubtedly due to scattering from the truncated lower torso, and are not expected to be encountered in human HRTFs. However, the general presence of strong torso shadow on the contralateral side for both KEMAR and the snowman model presents a strong contrast with the behavior of the spherical head model alone (cf. Fig. 4).

The boundary-element method was used to compute the frontal-plane HRTF for the ellipsoidal-torso model, and the results are shown in Fig. 12(c). As might be expected, the results are closer to those for the spherical-torso model than to KEMAR. There are two primary effects of stretching the spherical torso into an ellipsoid: (a) introducing angular anisotropy, and (b) breaking up the perfect symmetry that leads to the torso bright spot. Because there is little difference between the height and width of the ellipsoid, the first effect is not very prominent in a frontal-plane response. The second effect can be seen, however, in that the interference patterns in the contralateral torso shadow region are not as sharply defined for the ellipsoidal torso as for the spherical torso. Indeed, in this region, the polar plot for the ellipsoidal torso seems to be intermediate between that for the spherical torso and that for KEMAR.

2. Horizontal plane

The torso has relatively little effect on the HRTF in the horizontal plane. The HRTF for the isolated spherical head looks exactly the same in the horizontal plane as it does in the frontal plane (see Fig. 4). The horizontal-plane results for KEMAR, the spherical-torso snowman, and the ellipsoidal-torso snowman are shown in Fig. 13. The effects of torso reflection are more irregular in the KEMAR data than in the models. The difference between torso width and torso depth for the ellipsoidal-torso snowman introduces an anisotropy that can be seen as a flattening of the torso reflection arches in Fig. 13(c), but the effect is small. Perhaps the greatest difference between KEMAR and the models is the relative weakness of the head bright spot in KEMAR. As was explained in the discussion of frontal-plane results, this can be attributed to the fact that KEMAR's ears are displaced below the center of the head.

3. Median plane

In the median plane, the HRTF for the isolated spherical head is uninteresting, because there is no variation with the elevation angle, ϕ. The median-plane results for KEMAR, the spherical-torso snowman, and the ellipsoidal-torso snowman are shown in Fig. 14. Above the horizontal plane, all

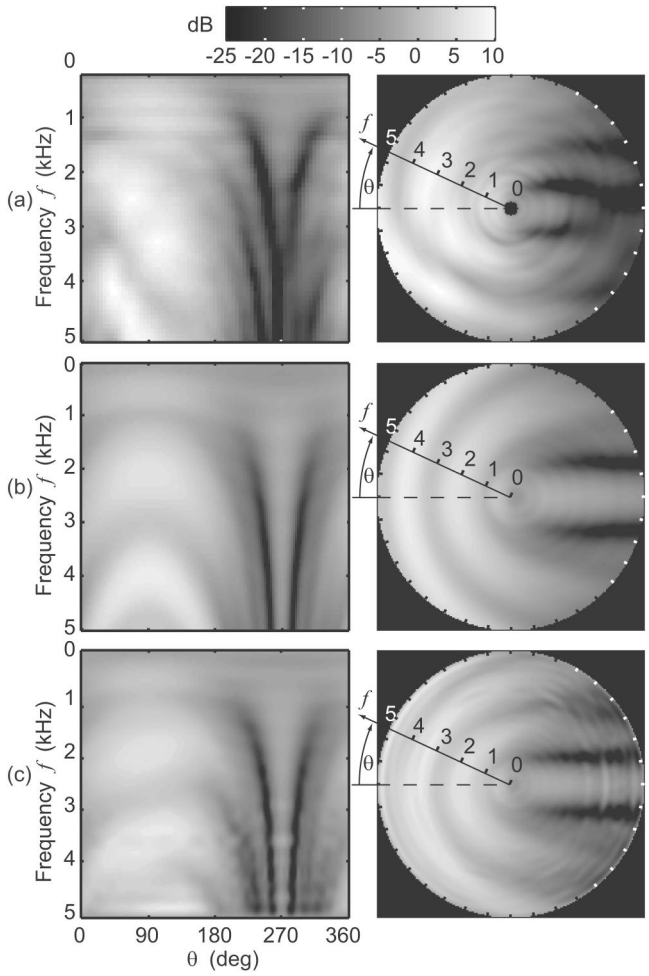


FIG. 13. Measured and computed horizontal-plane HRTFs. (a) KEMAR with the pinnae removed. (b) The spherical-torso snowman approximation. (c) The ellipsoidal-torso snowman approximation.

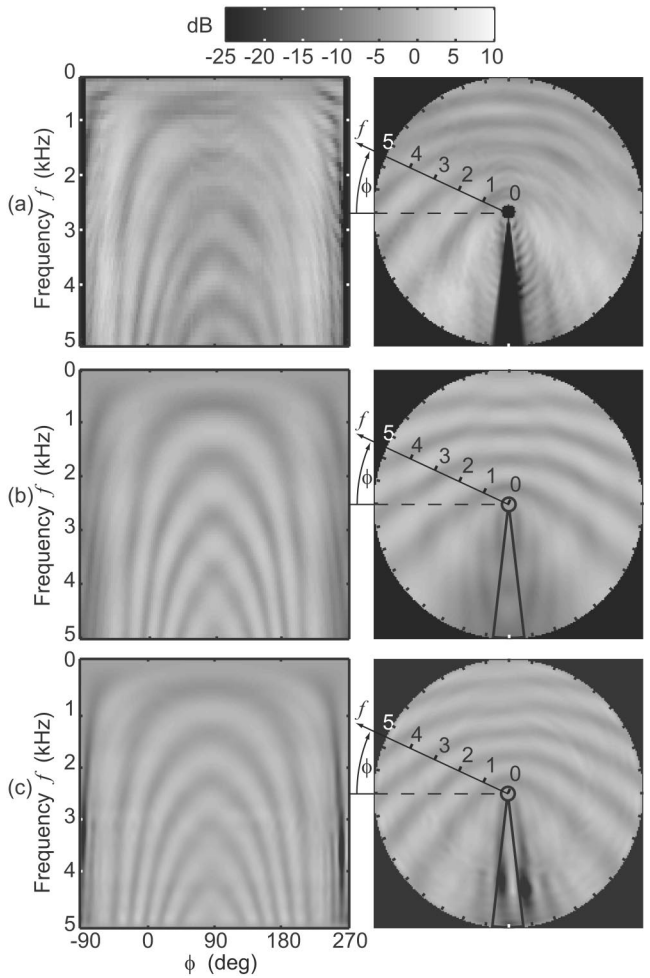


FIG. 14. Measured and computed median-plane HRTFs. (a) KEMAR with the pinnae removed. (b) The spherical-torso snowman approximation. (c) The ellipsoidal-torso snowman approximation.

three results are very similar, with the slightly flatter torso reflection arches for the ellipsoidal model providing a slightly better fit to the KEMAR data. At lower elevations, the differences between the spherical torso and the ellipsoidal torso models are greater than might be expected. In particular, the ellipsoidal-torso model exhibits two moderately deep interferences notches, with a bright spot at the point directly below, while the spherical-torso model exhibits only a general shadowing. The low-elevation shadowing is relatively weak compared to what is observed in the frontal plane. This is explained by the fact that head and torso shadow both occur in the frontal plane, while only torso shadowing occurs in the median plane. The ellipsoidal-torso results, although less easily explained, seem to be closer to what is observed in the KEMAR data. However, once again KEMAR's sharply truncated torso introduces some complex interference patterns at low elevations that are not really relevant to human HRTFs.

VI. DISCUSSION AND CONCLUSIONS

The results of this investigation show that simple models of the head and torso can explain the major features found in the pinnaless KEMAR HRTF for distant sources. The addition of either a spherical or an ellipsoidal torso to the standard spherical-head model changes the HRTF significantly, bringing the behavior of the model substantially closer to that of a pinnaless KEMAR. In particular, the torso introduces reflections when the source is at high elevation, and shadow when the source is at low elevations. These phenomena are not seen with the isolated spherical head model. More elaborate geometric models should produce even better approximations, but the simplicity of the spherical-torso snowman model facilitates systematic studies.

Our previous work showed that torso reflections provide relatively weak but genuine elevation cues, particularly when the source is away from the median plane (Algazi *et al.*, 2001a). The elevation cues provided by torso shadow have not been subjected to systematic psychoacoustic tests, but informal listening experiments using HRIRs generated by the models indicate that torso shadow does increase the sense that a virtual auditory source is at a low elevation. Because these elevation cues occur at low frequencies, they are particularly important for sources such as footsteps or thunder that have little high-frequency energy.

Simple head and torso models have uses in addition to providing insight into HRTF behavior. As part of a structural model of the HRTF, they can be customized to particular

individuals by matching their parameters to anthropometry (Algazi *et al.*, 2001d). In addition, they can be used to compensate acoustically measured HRTFs, filling in important low-frequency information that is difficult to measure experimentally.

There are several open issues that need further investigation. Perhaps most important, it is clear that the detailed behavior of human HRTFs at low elevations is sensitive to posture. However, relatively little is known as to how sensitive people are to the acoustic changes that accompany postural changes. Another open question concerns the effect of the neck on the HRTF and its perceptual importance. Finally, the sensitivity of the results to displacements of the head relative to the torso and displacements of the ears relative to the center of the head need to be better understood, particularly because these displacements have been found to be quite substantial for human subjects (Algazi *et al.*, 2001c). Although the geometric modeling approach cannot answer the perceptual questions, it offers an attractive way to investigate the effects of different components of the body on HRTFs.

ACKNOWLEDGMENTS

The authors would like to thank Dennis Thompson for his help with much of the experimental work. Support was provided by the National Science Foundation under Grant Nos. IIS-00-97256 and ITR-00-86075. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the National Science Foundation.

[1]To be more precise, the ITD and ILD for the spherical-head model are constant on a cone of confusion. In the interaural-polar coordinate system, a cone of confusion is a cone of constant interaural-polar azimuth θ_{ip} (Algazi *et al.*, 2001a). Thus, the ITD and ILD uniquely determine the interaural-polar azimuth, but provide no information about the interaural-polar elevation ϕ_{ip}. A simple coordinate conversion shows that the interaural-polar azimuth is related to the standard vertical-polar azimuth θ_{vp} and vertical-polar elevation ϕ_{vp} by $\theta_{ip} = \sin^{-1}[\sin\theta_{vp}\cos\phi_{vp}]$. Thus, strictly speaking, the ITD and ILD do not determine either θ_{vp} or ϕ_{ip}, but instead determine the product $\sin\theta_{vp}\cos\phi_{vp}$. However, the main point is that the spherical-head model does not provide sufficient cues to determine both the azimuth and the elevation.

[2]Butler (1975) gives an excellent survey of the extensive research on pinna cues. In addition to the static acoustic cues that are captured by the HRTF, there are also dynamic elevation cues generated by head motion (Perrett and Noble, 1997). This paper is concerned with static cues only.

[3]This interpretation is presented by Duda and Martens (1998). The impulse responses shown in that paper were computed for a 22.05-kHz bandwidth, and their higher temporal resolution makes it easier to see the two pulses that appear in the vicinity of the bright spot.

[4]These control points were drawn from earlier work on creating an HRTF database for human subjects in which operational definitions for 17 measurements of the head and torso, including the head height and neck height, were provided (Algazi *et al.*, 2001c). This unambiguously specified the locations of four of the torso control points: the left and right arm points and the front and back torso points. The definition for the top of the torso used with the HRTF database gave a point that was too low for the current study. For this paper, the top of the torso was located by going down from the center of the bounding box for the head by the amount of one-half the head height plus one third of the neck height. Although this procedure is arbitrary, it is well defined and can be applied to subjects other than KEMAR.

Algazi, V. R., Avendano, C., and Duda, R. O. (**2001a**). ''Elevation localization and head-related transfer function analysis at low frequencies,'' J. Acoust. Soc. Am. **109**, 1110–1122.

Algazi, V. R., Avendano, C., and Duda, R. O. (**2001b**). ''Estimation of a spherical-head model from anthropometry,'' J. Audio Eng. Soc. **49**(6), 472–478.

Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. (**2001c**). ''The CIPIC HRTF database,'' in *WASPAA'01* (Proc. 2001 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics), New Paltz, New York, pp. 99–102.

Algazi, V. R., Duda, R. O., Morrison, R. P., and Thompson, D. M. (**2001d**). ''Structural composition and decomposition of HRTFs,'' in *WASPAA'01* (Proc. 2001 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics), New Paltz, New York, pp. 103–106.

Batteau, D. W. (**1967**). ''The role of the pinna in human localization,'' Proc. R. Soc. London, Ser. B **168**, 158–180.

Blauert, J. P. (**1997**). *Spatial Hearing* (revised edition) (MIT, Cambridge, MA), pp. 50–93, 373–374.

Brungart, D. S., and Rabinowitz, W. R. (**1999**). ''Auditory localization of nearby sources. Head-related transfer functions,'' J. Acoust. Soc. Am. **106**, 1465–1479.

Burkhardt, M. D., and Sachs, R. M. (**1975**). ''Anthropometric manikin for acoustic research,'' J. Acoust. Soc. Am. **58**, 214–222.

Butler, R. A. (**1975**). ''The influence of the external and middle ear on auditory discriminations,'' in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer Verlag, New York), pp. 247–260.

Carlile, S. (ed.) (**1996**). *Virtual Auditory Space: Generation and Applications* (Landes, Austin, TX), pp. 38–45.

Ciskowski, R. D., and Brebbia, C. A. (eds.) (**1991**). *Boundary Element Methods in Acoustics* (Computational Mechanics Publications, Southampton).

Duda, R. O., and Martens, W. L. (**1998**). ''Range dependence of the response of a spherical head model,'' J. Acoust. Soc. Am. **104**, 3048–3058.

Duda, R. O., Avendano, C., and Algazi, V. R. (**1999**). ''An adaptable ellipsoidal head model for the interaural time difference,'' in *ICASSP'99* (Proc. IEEE International Conference in Acoustics Speech and Signal Processing), Vol. II, pp. 965–968.

Gardner, M. B., and Gardner, R. S. (**1973**). ''Problem of localization in the median plane: Effect of pinna cavity occlusion,'' J. Acoust. Soc. Am. **53**, 400–408.

Genuit, K., and Platte, H. J. (**1981**). ''Untersuchungen zur Realisation einer richtungsgetreuen Übertragung mit elektroakustischen Mitteln (Investigations on the implementation of directionally faithful transmission by electroacoustical means),'' *Fortschritte der Akustik*, FASE/DAGA '81 Berlin (VDE-Verlag, Berlin), pp. 629–632.

Gumerov, N. A., and Duraiswami, R. (**2001a**). ''Fast, exact, and stable computation of multipole translation and rotation coefficients for the 3-D Helmholtz equation,'' UMIACS Technical Report TR-2001-44, University of Maryland, College Park, MD (http://www.umiacs.umd.edu/~ramani/pubs/multipole.pdf).

Gumerov, N. A., and Duraiswami, R. (**2001b**). ''Multiple scattering from N spheres using multipole reexpansion,'' UMIACS Technical Report TR-2001-72, University of Maryland, College Park, MD (http://www.umiacs.umd.edu/~ramani/pubs/multisphere.pdf).

Gumerov, N., Duraiswami, R., and Tang, Z. (**2002**). ''Numerical study of the influence of the torso on the HRTF,'' in *ICASSP'02* (Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing).

Hartley, R. V. L., and Fry, T. C. (**1921**). ''The binaural localization of pure tones,'' Phys. Rev. **18**, 431–442.

Kahana, Y., and Nelson, P. A. (**2000**). ''Spatial acoustic mode shapes of the human pinna,'' preprint 5218, *109th Convention Audio Engineering Society*, Los Angeles, CA.

Kahana, Y., Nelson, P. A., Petyt, M., and Choi, S. (**1999**). ''Numerical modeling of the transfer functions of a dummy-head and of the external ear,'' Paper s52654, *AES 16th International Conference on Spatial Sound Reproduction*, Rovaniemi, Finland.

Katz, B. F. G. (**2001**). ''Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation,'' J. Acoust. Soc. Am. **110**, 2440–2448.

Kuhn, G. F. (**1977**). ''Model for the interaural time differences in the azimuthal plane,'' J. Acoust. Soc. Am. **62**, 157–167.

Kuhn, G. F. (**1987**). ''Physical acoustics and measurements pertaining to directional hearing,'' in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer Verlag, New York), pp. 3–25.

Kuhn, G. F., and Guernsey, R. M. (**1983**). ''Sound pressure distribution about the human head and torso,'' J. Acoust. Soc. Am. **73**, 95–105.
Lopez-Poveda, E. A., and Meddis, R. (**1996**). ''A physical model of sound diffraction and reflections in the human concha,'' J. Acoust. Soc. Am. **100**, 3248–3259.
Mehrgardt, S., and Mellert, V. (**1977**). ''Transformation characteristics of the external human ear,'' J. Acoust. Soc. Am. **61**, 1567–1576.
Morse, P. M., and Ingard, K. U. (**1968**). *Theoretical Acoustics* (Princeton U. P., Princeton, NJ), pp. 418–441.
Oppenheim, A. V., and Schafer, R. W. (**1969**). *Discrete-Time Signal Processing* (Prentice Hall, Englewood Cliffs, NJ), pp. 447–450.
Perrett, S., and Noble, W. (**1997**). ''The effect of head rotations on vertical plane sound localization,'' J. Acoust. Soc. Am. **102**, 2325–2332.
Shaw, E. A. G. (**1974**). ''The external ear,'' in *Handbook of Sensory Physiology*, Vol. V/I: Auditory System, edited by W. D. Keidel and W. D. Neff (Springer Verlag, New York), pp. 455–490.
Shaw, E. A. G. (**1997**). ''Acoustical features of the human external ear,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum, Mahwah, NJ), pp. 25–47.
Strutt, J. W. (Lord Rayleigh) (**1907**). ''On our perception of sound direction,'' Philos. Mag. **13**, 214–232.
Waterman, P. C., and Truell, R. (**1961**). ''Multiple scattering of waves,'' J. Math. Phys. **2**, 512–537.
Wightman, F. L., and Kistler, D. J. (**1989**). ''Headphone simulation of free-field listening. I: Stimulus synthesis,'' J. Acoust. Soc. Am. **85**, 858–867.
Wightman, F. L., and Kistler, D. J. (**1997**). ''Factors effecting the relative salience of sound localization cues,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum, Mahwah, NJ), pp. 1–23.
Wright, D., Hebrank, J. H., and Wilson, B. (**1974**). ''Pinna reflections as cues for localization,'' J. Acoust. Soc. Am. **56**, 957–962.
Zhou, B., Green, D. M., and Middlebrooks, J. C. (**1992**). ''Characterization of external ear impulse responses using Golay codes,'' J. Acoust. Soc. Am. **92**, 1169–1171.

Optimizing the acquisition time of the middle latency response using maximum length sequences and chirps

Steven L. Bell, Robert Allen, and Mark E. Lutman
Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

(Received 21 June 2002; accepted for publication 20 July 2002)

The middle latency response (MLR) may be used as an indicator of anesthetic depth but has been criticized due to its long acquisition time. This study explores methods for optimizing recording of the MLR to maximize signal-to-noise ratio (SNR) and hence reduce acquisition time. The first experiment investigates the effects of increasing stimulation rate beyond conventional values and also using higher rates by means of maximum length sequences (MLS). The second experiment compares the use of click and chirp stimuli to elicit the MLR, both at conventional and MLS stimulation rates. For all conditions total recording duration is fixed at 185 s and stimulation level is fixed at 60 dB SL. It was found that SNR increases progressively with rate using conventional click stimulation until the theoretical rate limit is reached at the reciprocal of the response duration. The SNR improvement is equivalent to increasing test speed by a factor of 3. Using MLS stimulation, the SNR increases further until a maximum is reached at a rate of 167 clicks/s, equivalent to a fivefold test speed improvement relative to a conventional recording at 5 clicks/s. The use of chirp stimuli designed to compensate for the frequency dependent cochlear traveling wave delay produces an increase in wave V–Na amplitude at all recording rates. For the later latency waves of the response an increase in amplitude is seen for MLS, but not for conventional chirp trains. The optimum SNR was obtained using chirp stimuli at a MLS rate of 167 opportunities/s. It is concluded that the combination of chirps and MLS can reduce acquisition time to less than one-tenth of that required for conventional stimulation at 5 clicks/s for the same SNR. This would confer material benefits for estimating anesthetic depth using MLR. [DOI: 10.1121/1.1508791]

PACS numbers: 43.64.Ri, 43.64.Yp [LHC]

I. INTRODUCTION

Reducing the recording time of auditory evoked potentials is almost always desirable, because there is generally limited time available and it may be necessary to obtain a large number of recordings under different conditions. The middle latency response (MLR) auditory evoked potential is usually elicited using click stimuli and occurs approximately 20 to 70 ms after the stimulus. It varies in amplitude from tenths of a microvolt to a few microvolts and is embedded in spontaneous EEG having an amplitude typically of 10 to 30 μV (Elkfafi *et al.*, 1997). Thus, the signal-to-noise-ratio (SNR) is less than 1:10 (-20 dB). The MLR waveform has a characteristic pattern with six features (Hall, 1992). These are three negative waves (N_a, N_b, and N_c) and three positive peaks (P_a, P_b, and P_c). For reference, a typical MLR is shown Fig. 1.

A potential application of the MLR is as indicator of depth of anesthesia in patients undergoing surgery (Thornton *et al.*, 1989; Beer *et al.*, 1996). Although the exact relationship between MLR components and patient awareness is still controversial (Jessop and Jones, 1992), Thornton *et al.* (1989) have demonstrated that the latencies and amplitudes of the midlatency waves change in response to both inhaled and intravenous agents. With increasing anesthetic dose, latency increases and amplitude decreases. Furthermore, N_b latencies shorter than about 45 ms appear to correspond to awareness as demonstrated with the isolated forearm technique (Tunstall, 1977). Also, the configuration of the waveform changes so that at light stages of anesthesia all three peaks are seen, at a moderate depth of anesthesia only P_a and P_b are seen, and at a deeper level of anesthesia than is desirable, none of the peaks is seen.

If the MLR is a reliable indicator of anesthetic depth, it may be possible to use the MLR as the input to a control system to adjust anesthetic dose (Kenny and Mantzaridis, 1999; Allen and Smith, 2001). In this case, it is desirable to maximize SNR, to give the best possible estimate for control. The traditional method for improving the SNR of auditory evoked potentials is to use a synchronized average of many successive response epochs, where the onset of the stimulus triggers the synchronization process. Theoretically, for n averages and Gaussian noise the SNR increases by a factor $\sqrt{n}$. Typically, to obtain an MLR trace when monitoring depth of anesthesia in surgery, a rolling average of 1000 epochs is used which is regularly updated (say, every 50 epochs). The responsiveness of the change in this moving average estimator to a change in anesthesia is crucial. At 5.95 clicks/s, averaging 1000 epochs takes 168 s to fully update the average. This is too slow to monitor potentially rapid changes in depth of anesthesia. The MLR has therefore been criticized for this purpose due to acquisition times of 3 to 5 min (Heier and Steen, 1996).

One way in which the acquisition time of the MLR may be improved is simply by increasing the click rate. In con-

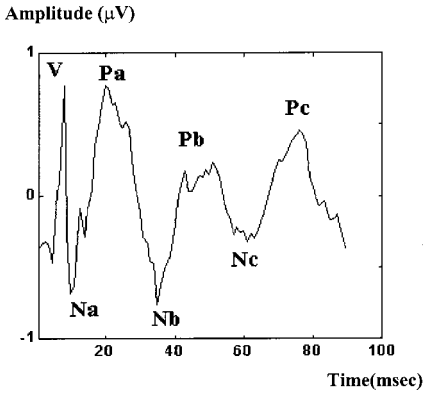


FIG. 1. A typical MLR.

ventional MLR recording, successive responses do not overlap. For an analysis time window of 70 ms, a conventional click stimulation pattern with a rate of up to 14 clicks/s might be used. Above this rate, responses would overlap. However, most studies that have attempted to measure depth of anesthesia using the MLR have used lower click rates such as 6 clicks/s (e.g., Elkfafi *et al.*, 1997). It is unclear whether this is because the MLR is thought to adapt and reduce in amplitude at higher click rates, or whether authors have simply followed the protocol of an early paper investigating depth of anesthesia using the MLR (Thornton *et al.*, 1989).

Another way in which the acquisition time of the MLR may be improved is by using an alternative pattern of click stimulation such as the maximum length sequence (MLS) technique (Eysholdt and Schreiner, 1982). MLS allows a higher stimulation rate to be used than for conventional MLR by overlapping successive recording epochs. A deconvolution technique is then used to extract the responses which overlap in time. An MLS of order b (where b is an integer) is of length 2^b-1 (stimulus opportunities) and contains $2^{(b-1)}$ stimuli. MLS has been used to improve the recording time of transient evoked otoacoustic emissions (Thornton, 1993) and it has been demonstrated that MLS can been used to record auditory evoked potentials such as the auditory brainstem response (ABR) (Burkard, Shi, and Hecox, 1990; Burkard, 1991; Lina-Grand, 1994; Jiang, 1999) and the MLR (Picton, Champagne, and Kellett, 1992; Musiek and Lee, 1997).

Bell, Lutman, and Allen (2001) have demonstrated that MLS can reduce the acquisition time of the MLR by a factor of 4 when compared to conventional click stimulation at a rate of 6 clicks/s. However, it is unclear whether the estimated test time reduction they achieved using MLS is greater than could be achieved by simply increasing the conventional click rate. Indeed, Van Veen and Lasky (1994) have demonstrated that, for the ABR, MLS reduces acquisition time no more than can be achieved by using the optimal conventional recording rate.

A further technique that may improve the acquisition of auditory evoked potentials has been suggested by Dau *et al.* (2000) and involves the use of chirp stimuli to compensate for temporal dispersion occurring in the response of the basilar membrane for stimuli at different frequencies. Click stimuli are commonly used to elicit auditory evoked potentials, as they have a wideband frequency spectrum and so stimulate a large portion of the cochlea. However, a click will take time to travel along the basilar membrane, so low-frequency regions at the apex of the cochlea will be stimulated about 10 ms later than high-frequency regions at the base of the cochlea. The response of the auditory nerve to a click will therefore be spread out in time. Dau *et al.* (2000) generated chirps that compensate for the place–frequency mapping of the cochlea, sweeping from low to high frequency within 10 ms. They demonstrated that ABR elicited using such chirps have a larger amplitude than ABR elicited using clicks. Rupp *et al.* (2001) have demonstrated a similar effect of a chirp stimulus on the amplitude of the middle latency auditory evoked field.

The amplitude, and hence the SNR, of recordings made with dichotic stimulation should be greater than that from monaural stimulation (Kadobayashi *et al.*, 1984). Dichotic stimulation should therefore result in a reduction in test time. The inclusion of conditions with both dichotic and monaural stimulation would assess the magnitude of this effect.

The aims of this study were therefore as follows: (1) to investigate the extent to which the acquisition time of the MLR might be improved by increasing the rate for conventional click stimulation; (2) to map the rate-adaptation curve for MLS stimulation with a view to optimization of test speed; (3) to investigate whether using the chirp stimuli of Dau *et al.* (2000) might further improve the acquisition time of the MLR, both for conventional stimulation rates and for higher MLS stimulation rates; and (4) to investigate the dichotic/monaural stimulation advantage when using conventional and chirp stimulation. Two experiments were completed for the study as described below.

II. METHOD

Subjects with normal hearing participated and the MLR was recorded at a fixed stimulus level (60 dB SL) for equal acquisition times. The SNR of each recording was estimated directly using the F_{sp} technique (Elberling and Don, 1984), which estimates the signal from the power in the final average and the noise from the variance from epoch to epoch of a single sample point at a fixed latency. The SNR was used to infer the relative test speeds of recordings as described below.

The study took the form of two experiments. The first investigated the effects of varying conventional and MLS recording rate on the MLR. In order to map the rate-adaptation effect fully, it was necessary to use two subject groups. The recordings from the first subject group did not map the adaptation function fully with MLS rate, so a further set of recordings at higher recording rates was made on a second subject group. The second experiment compared recordings made with click and chirp stimuli at two conventional and two MLS rates.

A. Subjects

Subjects were taken from the student population at the University of Southampton. For experiment 1 there were two groups: Group 1 comprised 6 males and 8 females and group 2 comprised 6 males and 5 females. Experiment 2 comprised

one group of 5 males and 5 females (not all the same subjects were available for both experiments). To minimize effects of age and hearing impairment, subjects included in the study were required to be otologically normal. They were aged between 18 and 30 years with hearing threshold levels better than 20 dB throughout the range 250–8000 Hz in each ear and all had normal tympanograms. No subjects had a history of ear disease or undue noise exposure. The mean age was 24.9 years for experiment 1 and 24.4 years for experiment 2.

B. Experimental design

1. Experiment 1

All recordings were made using rectangular click stimuli with a duration of 0.1 ms. For group 1, the majority of recordings involved dichotic stimulation, although some monaural stimuli were included. Six conventional stimulation rates of 5, 7, 9, 11, 13, and 15 clicks/s were used and six MLS order-4 stimulation rates of 50, 70, 90, 111, 143, and 167 stimulus opportunities/s. Monaural recordings were also made at 5 clicks/s and 90 stimulus opportunities/s to allow a comparison between monaural and dichotic stimulation. For group 2 all recordings involved dichotic stimulation and recordings were made at conventional rates of 5, 17, and 19 clicks/s and MLS rates of 70, 111, 143, 167, 200, 250, and 333 opportunities/s.

2. Experiment 2

Recordings were made at conventional rates of 5 and 15 clicks/s and MLS order-4 rates 143 and 167 opportunities/s. These represent the slowest conventional rate from experiment 1, the fastest conventional rate which allows an analysis window of around 70 ms, and the two MLS rates which produced the greatest SNR in experiment 1. All recordings were made with dichotic stimulation using the same click stimuli as for experiment 1 and using the chirp stimuli described below.

C. Stimuli

For both experiments, a stimulus level of 60 dB SL (sensation level) was used. Here, ''dB SL'' refers to the level of the click or chirp above the behavioral threshold level for the same stimulus (see below for details). In both experiments, the order of stimulation was randomized among subjects.

The same duration of 185 s was used for all recordings. Artifact rejection was applied off-line to exclude epochs that may contain excessive noise. For each recording, a criterion was calculated so that so that the extreme 10% of epochs ($\pm$2%) were rejected. This means that the same proportion of epochs was excluded for all subjects, thus maintaining parity of acquisition time.

1. Specification of the chirp stimuli

The chirps used in this experiment were the approximate[1] chirps defined by Dau *et al.* (2000). They had a nominal frequency range from 100–10 000 Hz and were generated using the same MATLAB algorithm as Dau *et al.* (see Fig. 2, upper panel). The chirps were replayed via insert earphones to obtain the MLR. In order to verify the acoustic chirps, the output from the insert earphone was delivered to an artificial ear complying with IEC 711, the microphone of which was, in turn, connected to a Brüel & Kjær 2112 spectrometer. The frequency spectra of the chirps and clicks are shown in the lower panel of Fig. 2.

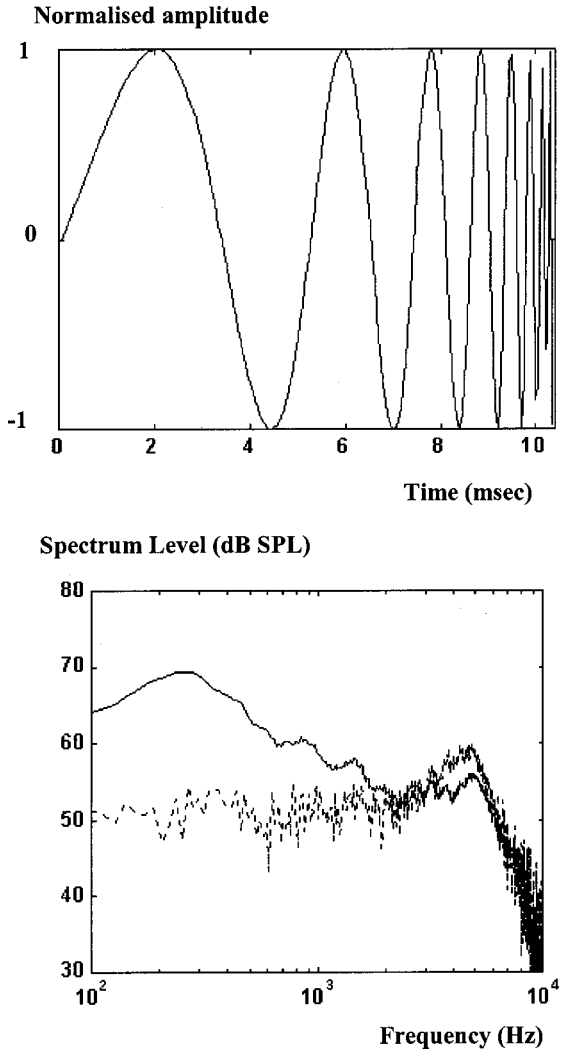


FIG. 2. Top panel: The waveform of the approximate chirp. Bottom panel: The frequency spectra of the approximate chirp (continuous line) and the click (dotted line) measured acoustically at 60 dB above the mean threshold level. Noise was reduced by averaging 100 epochs. A 93.8-dB SPL sound calibrator was used as a reference.

D. Apparatus

The subject was seated in a soundproof booth while click or chirp stimuli were delivered through insert earphones (Etymotic ER-3A for experiment 1 and ER-2 for experiment 2) via a Kamplex KC50 audiometer. For monaural presentation, stimuli were delivered to the right ear and, to prevent cross hearing, a broadband masking noise at 52-dB SPL was delivered via the left earphone. Relaxed subjects produce less electrophysiological noise and hence they were seated in a comfortable chair with a head rest. All subjects were awake and reading. To reduce electrical interference, screened cables were used and the biological amplifier (see below) was placed near the subject.

Stimuli were presented at a level of 60 dB SL to each ear, where sensation level (SL) is referenced to the detection threshold for clicks or chirps presented at a rate of 5/s. For group 1, the mean detection threshold for clicks was 38-dB peak equivalent (pe) SPL for the right ear and 40-dB pe SPL for the left ear. For group 2, the corresponding values were 39- and 38-dB pe SPL for clicks. For chirps the thresholds were 20- and 22-dB pe SPL.

All stimuli, including MLS, were produced using a computer-controlled CED μ1401 laboratory interface containing analog-to-digital and digital-to-analog converters. Responses were recorded synchronously by the same interface from recording electrodes connected via a CED 1902 biological amplifier. When recording using MLS, the sequences need to be produced contiguously (having no gaps between sequences). For each experimental condition, an appropriate sequence of stimuli was loaded into a memory buffer in the CED μ1401 and this was replayed cyclically at a sample rate of 10 kHz for experiment 1 and 20 kHz for experiment 2 (to ensure frequencies in the chirp stimulus were not aliased). The output sequence that was repeated could consist either of a single click or chirp followed by zeros for conventional recording, or an MLS of clicks or chirps of order 4, which contains 15 stimulus opportunities (8 of which are clicks or chirps). The length of the buffer was varied (and correspondingly the spacing of the stimuli within the buffer) to define the stimulation rate. Each click consisted of a 0.1-ms square wave output from the CED μ1401. The length of the chirps was 10.4 ms. The assumption has been made that the measurement system is linear, so the input sequence and the output response can be convolved in time. This assumption is required because at MLS stimulation rates of 143 and 167 opportunities/s some of the chirp stimuli will overlap in time.

The CED μ1401 sampled the outputs of the biological amplifiers at a rate of 10 kHz for experiment 1 and 20 kHz for experiment 2. Samples were placed into buffers which were the same length as the continuously repeating stimulus output buffer. Two buffers were used cyclically and their contents were alternately transferred to the controlling computer for averaging. Hence, while one buffer was acquiring data, the other was being transferred and analyzed. This process was effected by software specifically written for this purpose. Data were subsequently downsampled to 1 kHz for off-line analysis in MATLAB.

For recording, the bandpass filters were set to 15 Hz (high-pass) and 250 Hz (low-pass). Despite attempts to reduce 50-Hz mains interference on recordings, the FFTs of the raw data from both studies showed a significant 50-Hz component. Recordings were therefore zero-phase filtered in MATLAB using a 5th-order 50-Hz notch filter before further analysis. The effect of the filtering on the MLR waveforms was negligible.

1. Electrode configuration

For both experiments, auditory evoked responses to the clicks were picked up by three silver–silver chloride electrodes placed on the skin: active electrode on the vertex of the head, ground electrode on the forehead, and reference electrode on the nape of the neck after skin preparation. Recordings were made with impedances of less than 5 kΩ at 1 kHz between all pairs of the electrode array.

An important consideration when recording the MLR is the position of the reference electrode (Hall, 1992). With the reference electrode placed on the mastoid, postauricular muscle (PAM) interference is seen as a large -ve going trough at approximately 15 ms. The magnitude of the response is too large to be neural in origin and the PAM interference leads to an apparent N_b latency shift to 45 ms. PAM interference is reduced by using a nonencephalographic reference electrode placement (the nape of the neck in the current work).

E. Determination of signal-to-noise ratio

Sampled data comprised contiguous epochs from the recording electrodes synchronized with stimulation and streamed to disk. This allowed the use of the F_{sp} statistic to estimate SNR. For the present study, the F_{sp} statistic was defined by

$$F_{sp} = \text{var (S)}/\text{var (SP)},$$

where var (S) is the variance within the averaged MLR between 25 and 45 ms after the onset of the stimulus (which was chosen to correspond with the N_b region of the MLR) and var (SP) is the variance of a single point 35 ms after stimulus onset calculated across all the epochs recorded and divided by the number of epochs. This is a power SNR estimate which tends to 1 as SNR tends to zero; a rms measure of SNR is defined by $\sqrt{F_{sp}-1}$ [see Elberling and Don, 1984, Eq. (14)]. Sininger (1993) demonstrated a linear increase of F_{sp} with the number of averages taken. The rms SNR measure, $\sqrt{F_{sp}-1}$, should increase as the square root of the number of averages. This was verified by examining pilot data.

Elberling and Don (1984) suggest that the F_{sp} statistic is a better estimate of SNR than the $\pm$ difference measure (Wong and Bickford, 1980) when the signal may contain an unknown random variability due to low-frequency components. As the MLR contains substantial low-frequency components, the F_{sp} statistic was used in preference.

III. RESULTS

A. Experiment 1

1. Variation in P_a–N_b wave amplitude with stimulus rate

The variation in $P_a - N_b$ amplitude with rate for conventional stimulation is shown in the upper panel of Fig. 3 for groups 1 and 2. For conventional click rates (5 to 19 Hz) all stimulus opportunities are filled. The lower panel of Fig. 3 shows the variation with rate for MLS. For both subject groups, repeated measures analysis of variance (ANOVA) was performed separately with conventional stimulation or MLS rate as the within subject factor. For group 1, there was no significant effect of conventional rate but there was a significant effect with MLS shown by the somewhat sigmoid trend in the lower panel of Fig. 3. This was manifest as a significant cubic term in the polynomial fit of the ANOVA.

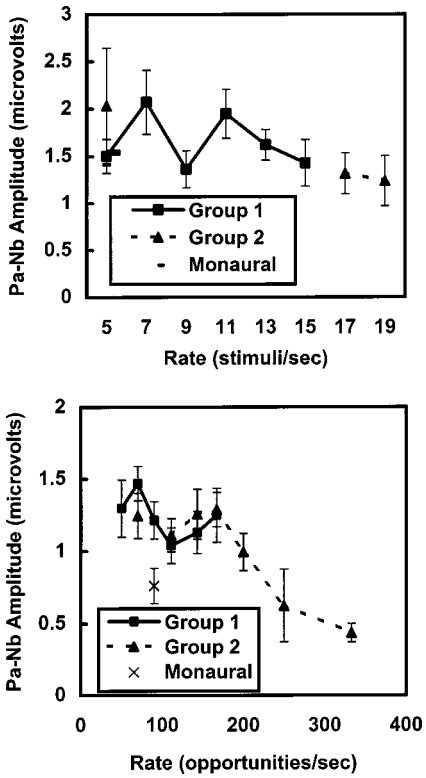


FIG. 3. Upper panel shows P_a-N_b amplitude as a function of click rate for conventional stimulation. Lower panel shows P_a-N_b amplitude as a function of click rate for MLS stimulation. Error bars represent one standard error of the mean.

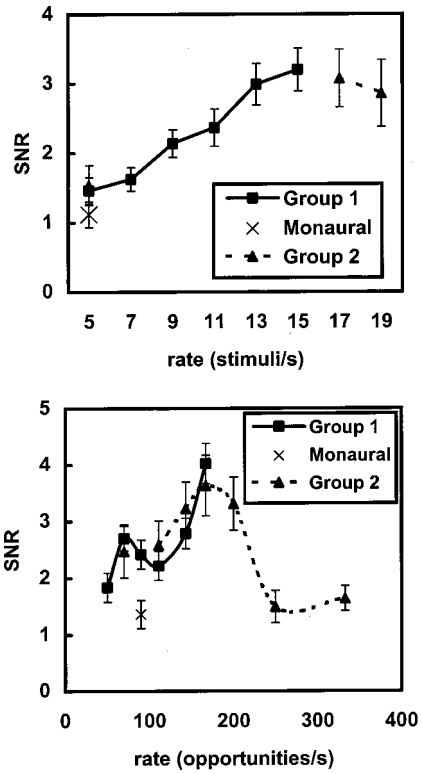


FIG. 4. Upper panel shows SNR as a function of click rate for conventional stimulation. Lower panel shows SNR as a function of click rate for MLS stimulation. Error bars represent one standard error of the mean.

Paired samples t-tests revealed no significant difference between monaural and dichotic wave amplitudes for stimulation at 5 clicks/s, where the variance in amplitude measurements was large. However, there was a significant difference ($p<0.01$) at 90 opportunities/s with monaural stimulation giving lower wave amplitude.

For group 2, only the MLS recordings shows a significant rate effect. The ANOVA involved polynomial contrasts to examine the effect of rate. For group 2 there were significant linear and quadratic effects shown by the accelerating downward trend in the lower panel of Fig. 3. It appears that there is little adaptation of response amplitude with conventional stimulation at rates below 20/s, but there is with MLS where rates are higher. As the MLS stimulation rate increases above 167 opportunities/s, the response amplitude decreases substantially. Repeated measures analysis of variance shows no overall effect of rate on wave latency for either conventional or MLS stimulation for either group (latencies are not shown).

2. Variation in SNR with stimulus condition

The upper and lower panels of Fig. 4 show the variation in SNR with rate from experiment 1 for conventional stimulation and MLS stimulation respectively (the data from the two groups have been overlaid). For conventional stimulation, SNR increases with rate. Repeated measures ANOVA for group 1 shows the linear effect of rate to be significant ($p<0.001$). For MLS stimulation, significant linear and cubic effects of rate are shown for group 1 ($p<0.001$) and a significant quadratic effect is evident for group 2 ($p<0.005$), where the SNR increases for rates up to around 143–167 clicks/s and then decreases at higher rates. The reason for this is that the averaged noise reduces with increasing number of stimuli to a greater extent than the reduction in signal amplitude shown in the lower panel of Fig. 3. A significant reduction in SNR occurs for monaural compared to dichotic stimulation ($p<0.001$) at 90 opportunities/s, but not at 5 clicks/s.

3. Variation in wave detectability with stimulus condition

N_a, P_a, and N_b wave detectability was assessed for all recordings made. The first author determined whether each wave could be identified. For click stimuli, wave detectability increases with the SNR. Only the recordings made using MLS stimulation at 167 opportunities/s had 100% detectability of the N_b wave for both subject groups.

B. Experiment 2

Experiment 2 investigated the effect of a chirp stimulus (Dau *et al.*, 2000) on the amplitude, latency and SNR of the MLR. A comparison was made between recordings made with clicks and chirps at two conventional and two MLS recording rates.

1. Variation in wave amplitude with stimulus type

The upper, middle, and lower panels of Fig. 5 show wave $V-N_a$, N_a-P_a, and P_a-N_b amplitudes for click and chirp stimuli at different rates. For the conventional click rates (5 and 15 Hz) all stimulus opportunities are filled. For

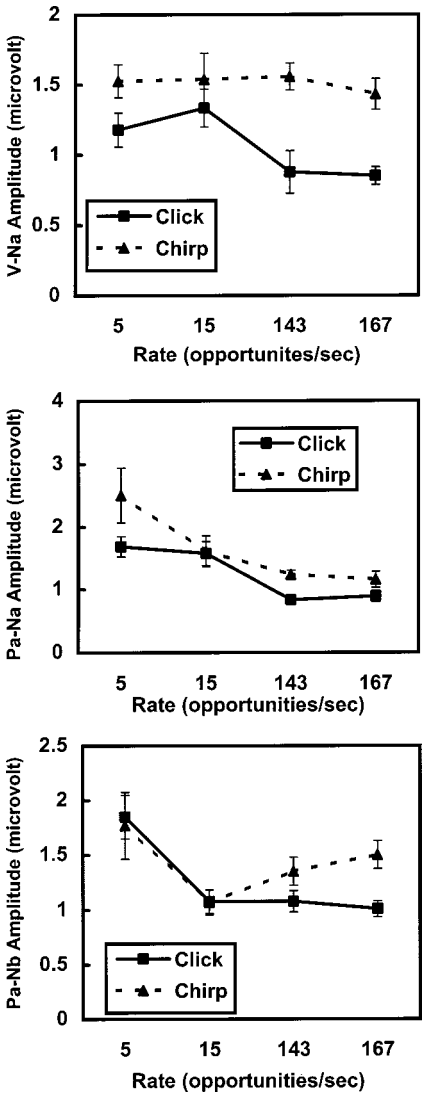


FIG. 5. Upper panel shows $V-N_a$ amplitude as a function of stimulus type. Middle panel shows N_a-P_a amplitude as a function of stimulus type. Lower panel shows P_a-N_b amplitude as a function of stimulus type. Error bars represent one standard error of the mean.

each wave, a repeated measures ANOVA was performed with stimulus type (click or chirp) and rate as within-subject factors. Only wave $V-N_a$ amplitude showed a significant overall effect of stimulus type ($p<0.05$). There is significant increase in amplitude for chirps over clicks at rates of 5, 143, and 167 opportunites/s as shown by paired-samples t-tests. The average increase in wave $V-N_a$ amplitude for chirps over clicks is 48%.

For the N_a-P_a and P_a-N_b waves, ANOVAs were performed separately for conventional and MLS stimulation. A significant increase in amplitude with chirp stimulation occurs at MLS stimulation rates ($p<0.001$ for N_a-P_a and $p<0.01$ for P_a-N_b) but not at conventional stimulation rates. The only significant effect of rate on wave amplitude is a decrease in P_a-N_b amplitude ($p<0.05$) with increasing conventional rate.

2. Variation in wave latency with stimulus type

Latencies for wave V of the ABR and the N_a, P_a, and N_b waves of the MLR were measured by the first author from each of the recordings made. Repeated measures ANOVA shows that the effect of rate is only significant for wave V ($p<0.01$), where latency increases with rate. The average increase in latency for chirps compared with clicks is 8.9 ms. This is slightly less than the full duration of the chirp (10.4 ms).

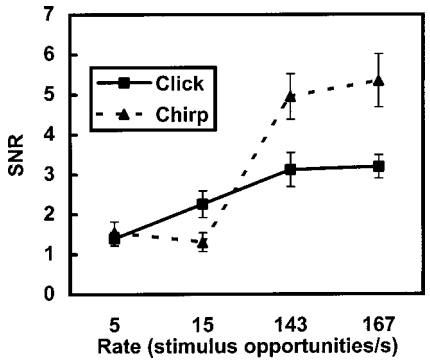


FIG. 6. SNR as a function of stimulus type. Error bars represent one standard error of the mean.

3. Variation in SNR with stimulus type

SNR was estimated for regions corresponding to the N_b region of the MLR (25–45 ms after the stimulus onset for clicks and 35–55 ms for chirps). The variation in SNR for the N_b region with stimulus type and rate is shown in Fig. 6.

Repeated measures ANOVA shows significant effects of both stimulus type and rate on SNR for the N_b region of the MLR ($p<0.001$), with a significant interaction between rate and stimulus type ($p<0.001$). The highest SNR is seen at 167 opportunities/s for both stimulus types. Paired t-tests revealed significant differences between clicks and chirps at 15, 143, and 167 stimulus opportunities/s ($p<0.05$). At 143 and 167 opportunities/s the SNR was significantly greater for chirps than clicks, but at 15 stimulus opportunities/s the SNR was significantly reduced. The pattern of increase in SNR with chirps over clicks for MLS stimulation, but not for conventional stimulation, is consistent with a corresponding increase in P_a-N_b wave amplitude for chirps over clicks with MLS.

4. Wave detectability

Wave detectability was rated by the first author for each of the stimulation conditions in experiment 2. Overall wave detectability is slightly higher for clicks than for chirps. Note that N_b wave presence is 100% for both stimulus types at MLS rates of 143 and 167 stimuli/s.

IV. DISCUSSION

A. Implications for test speed improvement by increasing conventional stimulation rate

One of the aims of this study was to investigate whether increasing the conventional recording rate could improve the acquisition of the MLR. Figures 4 and 6 show that SNR increases significantly with conventional rate. As acquisition time can be assumed to be inversely proportional to the square of the SNR, the relative acquisition time, and hence

test speed, relative to 5 clicks/s can be estimated. The test speed increase is almost linear with rate. For example, in experiment 2 where recordings at 5 and 15 clicks/s had a mean SNR of 1.39 and 2.25, respectively, the ratio of test speeds is $(2.25/1.39)^2 = 2.64$. This ratio approaches 3, which is what would be expected if the response does not adapt with increasing rate.

The limit to increasing conventional rate is set by overlapping of successive responses, which occurs when the response duration exceeds the reciprocal of the rate used. To identify the latency of a peak in an auditory response, it is necessary to view some of the response beyond the peak. If the N_b latency were to increase to around 50 ms due to increasing depth of anesthesia (Thornton, 1989), the analysis time window should not be shortened much below approximately 70 ms for correct identification of the wave. At 15 clicks/s, the maximum time window is 67 ms. This is around the upper limit for conventional stimulation rate which still allows identification of the N_b wave. The limitation imposed by the analysis window length is more restrictive when chirp stimuli are used, because additional time must be allowed for the chirp itself, which has a duration of 10.4 ms.

A comparison of monaural and dichotic stimulation was made for group 1 of experiment 1 at a conventional rate of 5 stimuli/s and an MLS rate of 90 stimulus opportunities/s. For both rates, monaural stimulation resulted in a smaller SNR than dichotic stimulation. This was significant at 90 opportunities/s but not at 5 stimuli/s, where the variance in measurements was greater. The corresponding predicted test time reductions were 1.7 times at 5 stimuli/s and 3.2 times at 90 stimulus opportunities/s.

B. The test speed improvement with MLS stimulation

It appears that a rate of 143–167 opportunities/s represents the best trade-off between a reduction in averaged noise due to an increasing number of averages and a decrease in response amplitude due to adaptation of the response with rate. This is consistent with the P_a–N_b amplitudes shown in the lower panel of Fig. 3, where wave amplitude is seen to drop off above 167 opportunities/s. The optimal MLS stimulation rate found in this study is somewhat higher than that found by Bell, Allen, and Lutman (2001). However, the electrode placement used in the previous study may have resulted in PAM interference on the responses recorded (see above), resulting in increased variability of recordings.

The SNR at 167 opportunities/s from experiment 2 is 3.19 (Fig. 6). If this is compared to the mean SNR of 1.39 at 5 clicks/s from group 1, the estimated test speed improvement is 5.3 times. This is approximately twice the improvement in test speed that can be achieved by increasing the conventional stimulation rate. This would appear to contradict the findings of Van Veen and Lasky (1994), who suggested that MLS cannot be used to improve the SNR of recordings above that of a conventional recording. However, their findings were based on data from other studies and they may not have taken into account the limitation of having a sufficient time window for wave analysis. At 167 click/s, the analysis time window available is 90 ms using an MLS of order 4.

The rate-adaptation curve shown in the lower panel of Fig. 3 shows a similar pattern to that found for adaptation of the ABR with MLS rate by Leung *et al.* (1998). They reported that SNR increased with MLS rate up to 200 opportunities/s at a level of 60 dB nHL. At higher rates, there was no increase in SNR due to adaptation of wave *V* of the ABR with rate. It would appear that a similar pattern of rate adaptation occurs for the MLR as for the ABR, and this may represent an underlying neural process of adaptation. However, the limit on conventional stimulation rate imposed by the required analysis time window is lower for the MLR (around 15 stimuli/s) than for the ABR (around 100 stimuli/s). Hence, the benefit of using MLS is greater for MLR than for the ABR [the ABR can be recorded at much higher conventional rates than the MLR, so little benefit is seen on ABR by increasing to the optimum MLS rate of 200 stimuli/s Leung *et al.* (1998)].

C. Latency variation with MLS stimulation

There was found to be no significant variation in wave latency with either increasing MLS or conventional recording rates. This finding contradicts previous studies which appear to have found a change in wave latency with rate (Bell, Allen, and Lutman, 2001; Picton *et al.*, 1992). However, both those studies used a mastoid reference electrode at A1. As mentioned above, PAM interference appears to produce a shift in wave latency and this effect may account for apparent latency shifts in those two studies.

D. The use of chirp stimuli to elicit the MLR

For experiment 2, the chirp stimuli described above were used at both conventional and MLS stimulation rates. For the wave V–N_a amplitude, an increase is seen for chirps over clicks at all stimulation rates. This is consistent with the findings of Dau *et al.* (2000), who investigated the use of chirp stimulation to obtain the ABR. However, for the N_a–P_a and P_a–N_b amplitudes, no significant improvement in response amplitude is seen at conventional rates, but a significant improvement is seen at MLS rates. This somewhat surprising finding requires further investigation. Possible explanations may be that the response of the higher regions of the auditory pathway to chirps are different for irregular pseudorandom stimuli than for regular stimuli. Alternatively, the adaptation of the later latency components of the MLR at high MLS rates may be different for clicks than for chirps.

The pattern of results for SNR was consistent with that for wave amplitude. For conventional stimulation rates, no benefit is seen for the chirp over the click stimulus. Indeed, the SNR is significantly lower for the chirp than the click at 15 stimuli/s. At both MLS stimulation rates, however, the chirp stimuli produced a significantly greater SNR than the click stimuli. The greatest SNR of 5.35 was obtained at 167 chirps/s. If this is compared to the SNR of 1.39 at 5 clicks/s, it represents an estimated reduction in test duration by a factor of 14.9.

As expected, the chirp resulted in an increase in wave latency of approximately 10 ms, which is the length of the

chirp stimulus. This is consistent with the concept that the chirp evoked response is not simply an onset response, but is a synchronous response across the nerve-fiber array which occurs when sufficient time has elapsed for the frequencies in the chirp to travel to the point of maximal excitation on the basilar membrane.

E. The potential of the MLR for assessing anesthetic depth

If the MLR is to be used as a measure of anesthetic depth, an important consideration is whether it is possible to record an MLR from all subjects undergoing surgery. Overall the wave detectability was high in both experiments, which is encouraging. However, it may not be possible to record an MLR from subjects with impaired hearing, or from young children in whom the MLR response is not fully developed (Hall, 1992). In both experiments, it was possible to record an N_b wave from all subjects at 167 opportunities/s. This was not true for any of the conventional recording rates and is consistent with the recordings at 167 opportunities/s having a higher SNR than the conventional recordings. Of course, wave detectability may be increased if the averaging process is continued for a longer time period, but in these experiments the recording time was fixed to represent the practical time constraints for assessing anesthetic depth.

Middle-ear pressure has been shown to vary with some anesthetic agents (O'Neill, 1985). If the MLR is to be used with anesthetics that alter middle-ear pressure, it will be necessary to investigate the effect that this pressure change has on the MLR. It will also be necessary to investigate whether other parameters such as subject age, body temperature, and hearing threshold can have a significant effect on the latencies of the MLR waves.

V. SUMMARY AND CONCLUSIONS

This study investigated the effect of varying the conventional stimulation rate, varying the MLS stimulation rate, and using chirp stimuli instead of clicks on the MLR. MLR recordings can be reliably recorded using conventional stimulation rate chirps, clicks at high MLS stimulation rates, and chirps at high MLS stimulation rates. The same duration of 185 s was used for all recordings. It was possible to record an N_b wave from all subjects at an MLS stimulation rate of 167 stimulus opportunities/s.

SNR was found to increase almost linearly with increasing conventional click rate, so a reduction in test time of 3 times can be expected by increasing the conventional click rate from 5 to 15 click/s. The length of the analysis time window available for correctly identifying an N_b wave, which may be shifted in latency due to the action of anesthetics, limits the amount by which the conventional click rate can be increased to approximately 15/s.

Monaural stimulation resulted in a reduction in response amplitude compared to dichotic stimulation. The correspondingly lower SNR would therefore result in an increased test time.

SNR increases with MLS stimulation rate up to 167 opportunities/s, beyond which SNR reduces. The rate-adaptation pattern for MLS stimulation is similar to that reported by Leung *et al.* (1998) for the ABR. The test speed improvement that can be expected with MLS clicks is 5.3 times that at 5 clicks/s, which is almost twice as much as can be achieved by simply increasing the conventional click rate.

The use of chirps causes a shift in wave latency of around 10 ms, corresponding to the length of the chirp itself. This is consistent with the findings of Dau *et al.* (2000), who hypothesize that the chirp elicits a synchronous response across the auditory-nerve fibers and is not simply an onset response. This means that a longer analysis time window is necessary to correctly identify waves.

For the wave $V-N_a$ region of the MLR, use of the chirp produces a significant increase in wave amplitude compared with click stimulation at all rates, consistent with the findings of Dau *et al.* (2000). However, for the later waves of the MLR, an increase in wave amplitude with chirp stimulation is seen at MLS stimulation rates, but not at conventional stimulation rates. This effect may warrant further investigation. The pattern is repeated in the SNR analysis, where replacing the click with a chirp produces an increase in SNR using MLS stimulation, but not using conventional stimulation.

The improvement in SNR using MLS chirps at 167 opportunities/s corresponds to an estimated reduction in test time by more than an order of magnitude compared to a conventional recording at 5 clicks/s. This might, for example, correspond to a reduction in test time from 185 to under 20 s, which represents a valuable improvement if N_b latency is to be used to assess depth of anesthesia in the operating theater.

ACKNOWLEDGMENTS

Thanks are due to Torsten Dau, who kindly provided a MATLAB file to generate the ''approximate chirps'' he used in his paper,[1] and to Roger Thornton for his comments on a earlier version of this paper. The support of SLB by the EPSRC is also acknowledged.

[1]Dau *et al.* (2000) used two types of chirp: an ''exact chirp'' that simulates the situation where the propagation constant of the traveling wave on the basilar membrane is frequency dependent, and an ''approximate chirp'' in which the propagation constant is the same for all frequencies. They show that the frequency-independent expression used for the approximate chirp produces errors at low frequencies near the cochlear window, but that the effect is not large.

[2]When chirp stimuli were used, the analysis time window was moved to 35–55 ms and the single point used to calculate F_{sp} was taken 45 ms after the start of the chirp to account for the length of the chirp.

Allen, R., and Smith, D. (**2001**). ''Neuro-fuzzy closed-loop control of depth of anasthesia,'' Artif. Intell. Med. **21**, 185–91.

Beer, N. A. M., Van Hooff, J. V., Brunia, C. H. M., Cluitmans, P. J. M., Korsten, H. H. M., and Beneken, J. E. W. (**1996**). ''Midlatency auditory evoked potentials as indicators of perceptual processing during general anasthesia,'' Br. J. Anaesth. **77**, 617–624.

Bell, S. L., Allen, R., and Lutman, M. E. (**2001**). ''The feasibility of maximum length sequences to reduce acquisition time of the middle latency response,'' J. Acoust. Soc. Am. **109**, 1073–1081.

Burkard, R. (**1991**). ''Human brain-stem auditory evoked responses obtained by cross correlation to trains of clicks, noise bursts, and tone bursts,'' J. Acoust. Soc. Am. **90**, 1398–404.

Burkard, R., Shi, Y., and Hecox, K. E. (**1990**). ''Brain-stem auditory-evoked responses elicited by maximum length sequences: Effects of simultaneous masking noise,'' J. Acoust. Soc. Am. **87**, 1665–1672.
Dau, T., Wegner, O., Mellert, V., and Kollmeier, B. (**2000**). ''Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion,'' J. Acoust. Soc. Am. **107**, 1530–1540.
Elberling, C., and Don, M. (**1984**). ''Quality estimation of averaged auditory brainstem responses,'' Scand. Audiol. **13**, 187–197.
Elkfafi, M., Shieh, J. S., Linkens, D. A., and Peacock, J. E. (**1997**). ''Intelligent signal processing of evoked potentials for anasthesia monitoring and control,'' IEE Proc.: Control Theory Appl. **144**, 354–360.
Eysholdt, U., and Schreiner, C. (**1982**). ''Maximum length sequences—A fast method for measuring brainstem evoked responses,'' Audiology **21**, 242–250.
Hall, J. W. (**1992**). *Handbook of Auditory Evoked Responses* (Allyn and Bacon, Boston).
Heier, T., and Steen, P. A. (**1996**). ''Assessment of anesthesia depth,'' Acta Anaesthesiol. Scand. **40**, 1087–1100.
Jessop, J., and Jones, J. G. (**1992**). ''Evaluation of the actions of general anasthetics in the human brain,'' Gen. Pharmacol. **23**, 927–35.
Jiang, Z. D., Brosi, D. M., and Wilkinson, A. R. (**1999**). ''Brainstem auditory evoked response recording using maximum length sequences in term neonates,'' Biol. Neonate **76**, 193–199.
Kadobayashi, I., Kira, Y., Toyoshima, A., and Nishijima, H. A. (**1984**). ''A study of auditory middle latency responses in relation to electrode combinations and stimulus conditions,'' Audiology **23**, 509–519.
Kenny, G. N., and Mantzaridis, H. (**1999**). ''Closed-loop control of propofol anasthesia,'' Br. J. Anaesth. **8**, 223–228.
Leung, S., Slaven, A., Thornton, A. R. D., and Brickley, G. J. (**1998**). ''The use of high stimulus rate auditory brainstem responses in the estimation of hearing threshold,'' Hear. Res. **123**, 201–205.
Lina-Granada, G., Collet, L., and Morgon, A. (**1994**). ''Auditory-evoked brainstem responses elicited by maximum-length sequences in normal and sensorineural ears,'' Audiology **33**, 218–236.
Musiek, F. E., and Lee, W. W. (**1997**). ''Conventional and maximum length sequences middle latency response in patients with central nervous system lesions,'' J. Am. Acad. Audiol **8**, 173–180.
O'Neill, G. (**1985**). ''Prediction of post-operative middle ear pressure changes after general anesthesia with nitrous oxide,'' Acta Oto-Laryngol. **100**, 51–57.
Picton, T. W., Champagne, S. C., and Kellett, A. J. C. (**1992**). ''Human auditory potentials recorded using maximum length sequences,'' Electroencephalogr. Clin. Neurophysiol. **84**, 90–100.
Rupp, A., Uppenkamp, S., Gutschalk, A., Dau, T., Patterson, R., and Scherg, M. (**2001**). ''On the representation of peripheral neural activity patterns in primary auditory cortex,'' Association for Research in Otolaryngology, Midwinter meeting.
Sininger, Y. S. (**1993**). ''Auditory brain stem response for objective measures of hearing,'' Ear Hear. **14**, 23–30.
Thornton, A. R. D. (**1993**). ''High rate otoacoustic emissions,'' J. Acoust. Soc. Am. **94**, 132–136.
Thornton, C., and Newton, D. E. F. (**1989**). ''The auditory evoked response: A measure of depth of anasthesia,'' in *Balliere's Clinical Anasthesiology: Depth of Anasthesia*, edited by J. G. Jones (Balliere-Tindall, London, 1989), pp. 559–585.
Thornton, C., Barrowcliffe, M. P., Konieczko, K. M., Ventham, P., Doro, C. J., Newton, D. E. F., and Jones, J. G. (**1989**). ''The auditory evoked response as an indicator of awareness,'' Br. J. Anaesth. **63**, 113–115.
Tunstall, M. E. (**1977**). ''Detecting wakefulness during general anasthesia for Caesarean section,'' Br. Med. J. **21**, 1321.
Van Veen, B. D., and Lasky, R. E. (**1994**). ''A framework for assessing the relative efficiency of stimulus sequences in evoked response measurements,'' J. Acoust. Soc. Am. **96**, 2235–2243.
Wong, P. K. H., and Bickford, R. G. (**1980**). ''Brain stem auditory evoked potentials: The use of noise estimate,'' Electroencephalogr. Clin. Neurophysiol. **50**, 25–34.

Primitive stream segregation of tone sequences without differences in fundamental frequency or passband

Brian Roberts[a)]
School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, England

Brian R. Glasberg and Brian C. J. Moore
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

(Received 26 December 2001; revised 31 May 2002; accepted 1 August 2002)

Peripheral-channeling theorists argue that differences in excitation pattern between successive sounds are necessary for stream segregation to occur. The component phases of complex tones comprising unresolved harmonics ($F0=100$ Hz) were manipulated to change pitch and timbre without changing the power spectrum. In experiment 1, listeners compared two alternating sequences of tones, A and B. One sequence was isochronous (tone duration=60 ms, intertone interval=40 ms). The other began isochronously, but the progressive delay of tone B made the rhythm irregular. Subjects had to identify the sequence with irregular rhythm. Stream segregation makes this task more difficult. A and B could differ in passband (1250–2500 Hz, 1768–3536 Hz, 2500–5000 Hz), component phase (cosine, alternating, random), or both. Stimuli were presented at 70 dB SPL in pink noise. Dissimilarity in either passband or phase increased discrimination thresholds. Moreover, phase differences raised threshold even when there was no passband difference. In experiment 2, listeners judged moment-by-moment the grouping of long ABA-ABA-... sequences. The measure was the proportion of time a sequence was heard as segregated. The factors that increased segregation were very similar to those that increased threshold in experiment 1. Overall, the findings indicate that substantial stream segregation can occur without differences in power spectrum. It is concluded that differences in peripheral channeling are not a requirement for stream segregation.
[DOI: 10.1121/1.1508784]

PACS numbers: 43.66.Ba, 43.66.Mk, 43.66.Hg [NFV]

I. INTRODUCTION

The auditory system is often exposed to a mixture of sounds arising from independent sources in the environment, which must be partitioned into separate perceptual streams. This process involves the perceptual separation of sounds that are temporally interleaved (sequential grouping), temporally overlapping (simultaneous grouping), or both. Perceptual properties like melody, rhythm, and tempo depend on sequential grouping. Ortmann (1926) proposed that the grouping of musical tones into melodies is governed by similarity in frequency. Experimental studies using sequences of alternating high- and low-frequency pure tones have since confirmed his insight (Miller and Heise, 1950; Bregman and Campbell, 1971; van Noorden, 1975). These studies have shown that such sequences can be heard either as arising from a single source or from more than one source. Stream segregation is most likely when there is a large difference in frequency between high and low tones, and when they are heard in rapid succession. Bregman (1990) argues that this reflects the likelihood that they have arisen from a common source.

Hartmann and Johnson (1991) proposed that stream segregation can be explained largely in terms of differences in the ''peripheral channeling'' of sounds by frequency and by ear (i.e., sounds producing a more similar pattern of peripheral excitation are less likely to segregate from each other). This hypothesis arose from a set of experiments that they conducted using an interleaved melody task similar to that devised by Dowling (1973). On each trial, two familiar melodies were interleaved note by note and subjects were asked to identify them. The melodies were contrasted using a range of acoustic attributes (e.g., ear of presentation, temporal envelopes, spectral composition) and the effects on accuracy were measured. Hartmann and Johnson observed that differences in peripheral channeling were the dominant factor governing stream segregation. Recent attempts to develop computational models based on this idea have proved successful in accounting for a range of streaming effects (e.g., Beauvois and Meddis, 1996; McCabe and Denham, 1997).

A. Are peripheral-channeling cues a prerequisite for stream segregation?

Many studies of sequential grouping using complex tones have shown that differences in timbre and pitch can lead to stream segregation (e.g., Wessel, 1979; Singh, 1987; Bregman *et al.*, 1990; Iverson, 1995). However, differences in timbre are typically produced by spectral changes (e.g., compactness, tilt, formant frequencies) and the presence of resolved harmonics (Shackleton and Carlyon, 1994) leads to changes in the excitation pattern as F0 is changed, even

[a)]This study was carried out in autumn 2000, while the first author was on sabbatical in Cambridge. Electronic mail: b.roberts@bham.ac.uk

when all tones occupy a common spectral passband. Hence, the results of most of these studies are explicable in terms of peripheral channeling. For example, studies using sequences of musical instrument sounds have shown that the best predictor of their segregation is similarity of spectral center-of-gravity (Iverson, 1995; Wessel, 1979). However, streaming can also be affected by amplitude envelope. Iverson (1995) found that sequences of musical instrument tones were rated as more segregated when their amplitude envelopes were more dissimilar. Singh and Bregman (1997) observed that a smaller separation in *F*0 was required for sequences of two harmonic complexes to segregate when they differed in amplitude envelope (rise/fall 5/95 ms or 95/5 ms) than when they did not. Nonetheless, this result does not show that streaming can occur in the absence of peripheral-channeling cues, because the *F*0 separation required for stream segregation was sufficient to result in different peripheral channels being excited by the two sounds.

It is important to note that studies of auditory sequential organization do not provide measures of a single underlying process. Bregman (1990) distinguishes two processes by which a subset of sounds may be heard as separate perceptually from the rest. One of these, *primitive stream segregation*, is characterized as a bottom-up process that is automatic and obligatory. The other, *schema-based selection*, is characterized as an attentional process that uses knowledge of the form of familiar sounds to select them. This distinction is consistent with the observations of van Noorden (1975). He examined the grouping of a sequence comprising alternating high- and low-frequency tones that differed in tone onset-to-onset time (tone repetition time, or TRT) and in frequency separation. He made a distinction between the temporal coherence boundary (TCB) and the fission boundary on a plot of frequency separation against TRT. The sequence could not be heard as a single stream above the TCB, presumably because of the automatic operation of primitive segregation. Below the fission boundary, the sounds could not be heard as two streams. Between these two boundaries, either percept could be heard. Primitive stream segregation is influenced by TRT and frequency separation. At longer TRTs, larger separations are required to induce streaming. In contrast, sounds can be selected by schema-driven processes if their frequency separation is larger than a minimum critical value, of a few semitones, almost regardless of TRT. Contrary to Bregman's (1990) distinction, Carlyon *et al.* (2001) have recently presented evidence that the build up of primitive stream segregation is influenced by attention. However, the idea that primitive stream segregation, once established, is automatic and obligatory in character has not been challenged.

Vliegen and Oxenham (1999) and Cusack and Roberts (2000, experiment 1) demonstrated that listeners can select one set of sounds (an interleaved melody) from another (a set of distractors) in the absence of differences in peripheral channeling. The former study eliminated spectral cues by using complex tones with a common passband and comprising only unresolved harmonics. The notes of the melody (and distractors) were defined by changes in *F*0. The latter study eliminated spectral cues by using narrow-band stimuli (pure tones or noises) within a fixed frequency range. Notes were defined by changing center frequency, and could be either the same or contrasting in timbre. One problem with interpreting the results of these studies is that schema-based selection, rather than enhanced primitive stream segregation, might be responsible for success in identifying the melody. To distinguish between the two processes, a task is required for which good performance depends on the integration of the sequence of sounds (i.e., they must be heard as a single stream). If the introduction of an acoustic difference leads to obligatory and automatic stream segregation, then worse performance should result despite the efforts of the listeners.

B. Temporal discrimination as a measure of stream segregation

Many studies have shown that stream segregation can impede the perception of temporal relationships within a sequence of sounds. For example, it has long been known that listeners are poor at judging the order of disparate sounds presented in rapid sequences (e.g., Broadbent and Ladefoged, 1959; Warren *et al.*, 1969). Bregman and Campbell (1971) presented listeners with repeating sequences of six pure tones, comprising three high-frequency tones interleaved with three low-frequency tones. It was found that rapid presentation led to the perception of separate high- and low-pitched streams, and that listeners were good at judging the temporal order of a subset of three tones only if they were all heard as part of the same stream. These effects are distinct from findings of elevated thresholds for detecting a temporal gap between two isolated tones when they are dissimilar (see Vliegen *et al.*, 1999; Grose *et al.*, 2001, for reviews). In particular, van Noorden (1975) found that performance depended on frequency interval much more strongly for long sequences of tones than for short ones. This is consistent with results showing that stream segregation builds up over time (Bregman, 1978; Anstis and Saida, 1985). The auditory system seems to begin with a bias towards integration, but gradually accumulates evidence that may cause streaming to occur (Bregman, 1990).

Cusack and Roberts (2000, experiment 2) demonstrated, using a temporal discrimination task, that primitive stream segregation could be enhanced substantially by differences in timbre based on periodicity cues, but did not try to show that this factor could induce streaming in the absence of spectral cues. Vliegen *et al.* (1999) were the first authors who tried to do this. They used an ABA-ABA-... sequence, in which tone B could be either exactly at the temporal midpoint between two successive tones A or slightly delayed. In one set of conditions, the complex tones comprised high, unresolved, harmonics with a varying *F*0 and a fixed spectral passband. Tone A had a fixed *F*0 of 100 Hz and tone B had an *F*0 in the range 1–18 semitones higher. By using unresolved harmonic complexes, Vliegen *et al.* (1999) aimed to ensure that the excitation pattern showed only minimal changes ($<$0.5 dB; Glasberg and Moore, 1990) over the range of *F*0's used. Each sequence consisted of four ABA triplets, flanked on each side by two triplets from which tone B was removed (A_A). The initial tones A were introduced to encourage a tendency for stream segregation to build up. In the reference

stimulus, the position of tone B was exactly at the midpoint between two tones A. In the comparison stimulus, tone B was delayed, so that the rhythm became irregular when the sequence was heard as integrated. The effect of increasing the *F*0 difference between tones A and B from 1 to 18 semitones was significant, but small (mean threshold increased by about 4 ms). In contrast, increasing the frequency difference between A and B over the same range in a pure-tone control condition raised the mean threshold by about 11 ms. Varying the passband of the complexes produced a change in threshold similar to that seen in the pure-tone condition.

Vliegen *et al.* (1999) suggested that periodicity differences are less potent than spectral differences in producing primitive stream segregation, and hence are more easily overcome in a task requiring integration. In contrast, periodicity differences are clearly salient enough to be exploited successfully in a task requiring selection (Vliegen and Oxenham, 1999). However, there is some reason for caution in interpreting Vliegen *et al.*'s results:

First, the TRT was 120 ms within an ABA triplet. Van Noorden (1975) showed that the separation between high and low tones at the TCB increases markedly for TRTs above 100 ms. Hence, primitive stream segregation is more likely to be observed for TRTs shorter than those used by Vliegen *et al.* (1999). Second, the leading tones A may have been only partly successful in overcoming an initial bias towards integration. Only four 90-ms tones A were heard in the first 900 ms before the set of ABA triplets began. So few tones, with such a long TRT, are unlikely to have caused a near-maximum build up of streaming by the time a temporal discrimination was required. Hence, the degree of segregation arising from *F*0 differences may have been underestimated. Third, the offset-to-onset time between consecutive tones A was 150 ms within a triplet, but only 120 ms between triplets. This would have produced a more complex rhythm when tones A and B segregated than that produced by equal time intervals, which may have increased the difficulty of the discrimination. Fourth, the differing effects of changes in passband or *F*0, which produce changes in timbre and pitch, respectively, may have influenced Vliegen *et al.*'s findings. It is difficult to equate these perceptual differences. Perhaps the set of pitches was less heterogeneous than the set of timbres, and hence appeared less effective as a cue for streaming. A different selection of periodicity cues might produce greater heterogeneity and hence provide clearer evidence that primitive stream segregation can occur in the absence of spectral cues. The experiments reported here sought to overcome these problems.

II. EXPERIMENT 1

A. Method

1. Overview

This experiment measured temporal discrimination performance for sequences of alternating complex tones. The properties of these tones were varied in ways intended to influence their stream segregation and hence to affect the ability of listeners to perform the discrimination. In its purest form, the peripheral-channeling hypothesis states that stream segregation cannot occur in the absence of differences in peripheral excitation. Therefore, a rigorous test of the hypothesis requires a manipulation that can produce a substantial change in the pitch or timbre of a sound with only a minimal change in the excitation pattern evoked. The approach taken here was to manipulate the phase relations of a set of unresolved harmonics in order to change pitch (value and strength) and timbre without changing the *F*0. Previous studies have shown that the pitch sensation induced by unresolved-harmonic complexes can be used to identify melodies and to discriminate musical intervals (Moore and Rosen, 1979; Houtsma and Smurzynski, 1990; Vliegen and Oxenham, 1999).

Shackleton and Carlyon (1994) reported that the pitch generated by a set of unresolved harmonics of 100 Hz increased by an octave when the component phases were changed from cosine to alternating (odd harmonics in sine phase and even harmonics in cosine phase). This effect occurs because the phase alternation generates a substantial secondary peak in the temporal envelope within each period. Bilsen (1973) found that randomizing the phases of a set of unresolved harmonics, which eliminates regular peaks in the temporal envelope, weakened greatly the sensation of pitch and created a harsher, more noiselike, percept. These changes in pitch and timbre are very salient, but do not lead to any change in the excitation pattern as computed from the power spectrum. Hence, even if frequency resolution is greater than is typically assumed (see Moore, 2002), any residual ripple in the excitation patterns evoked by these stimuli should not confound the results of the experiment (however, see Sec. IV A for further discussion). A clear demonstration that temporal discrimination for successive tones with the same passband and *F*0 can be impaired simply by introducing a phase difference between them would indicate that a difference in peripheral channeling is not a minimum condition for primitive stream segregation.

To interpret the magnitude of phase effects on stream segregation, this study also explored the effect of differences in peripheral channeling by including conditions in which the stimuli had different passbands. It was predicted that temporal discrimination would deteriorate as the difference in spectral center-of-gravity between consecutive tones increased.

2. Subjects

Six subjects (five young adults and the second author) took part in this experiment. All subjects reported having normal hearing and having received some musical training. Only the second author had previously taken part in auditory perception experiments.

3. Stimuli and conditions

Each stimulus was an alternating sequence of complex tones (ABABAB...) with an added background noise to mask any combination tones (e.g., Smoorenburg, 1972; Pressnitzer and Patterson, 2001). All tones had a constant *F*0 of 100 Hz and comprised only high, unresolved harmonics. Nine types of complex tone were created through the binary combina-

tion of three passbands and phase relations. Harmonic complex tones were generated by adding components in cosine phase (c), alternating phase (a), or random phase (r). Following Shackleton and Carlyon (1994), alternating-phase stimuli were composed of odd harmonics in sine phase and even harmonics in cosine phase. Components within the passband were set to equal amplitude, and those outside were attenuated by 80 dB/oct. All components for which the required attenuation exceeded 60 dB were removed. The spectral slopes of the stimuli were not steep enough to evoke strong edge pitches (Kohlrausch and Houtsma, 1992). The three passbands, each one octave wide in terms of their −3 dB points, were 1250–2500 Hz (low, L), 1768–3536 Hz (medium, M), and 2500–5000 Hz (high, H). Note that the corner frequencies of band M correspond to the geometric means of their counterparts in bands L and H. The choice of frequencies also ensured that the lowest harmonic at full amplitude was the 13th, and hence it can be assumed that all stimuli comprised only unresolved harmonics (see Shackleton and Carlyon, 1994). A one-octave passband is wide enough to ensure that a set of unresolved harmonics generates a clear pitch (at least in the cosine- and alternating-phase conditions).

The total duration of each tone was 60 ms, including 10-ms raised-cosine onset and offset ramps. In a reference sequence, the silent interval after each tone was 40 ms, giving an onset-to-onset time between consecutive tones of 100 ms. This value is very effective at inducing streaming in alternating pure-tone sequences (Bregman and Campbell, 1971; van Noorden, 1975). Each reference sequence comprised 12 AB cycles of 200 ms each, giving an overall duration of 2.4 s. Test sequences differed from reference sequences only in that some of the tone B tokens were delayed to make the sequence anisochronous (see Sec. II A 4).

There were two main sets of sequences, one in which tone A was always Hc (high band, cosine phase) and one in which it was always Lc (low band, cosine phase). Each of these tones A was paired with all nine combinations of passband and component phase for tone B. Hence the experimental factors were band difference and phase difference. There was also a control set of four sequences (Ha-Ha, La-La, Hr-Hr, Lr-Lr) to check that changes in pitch strength and value, or in timbre, produced by changing the component phases did not have any unforeseen effects on performance. Hence, there was a grand total of 22 stimuli.

All complex tones were presented at 70 dB SPL, giving a level per component of 57.4 dB for M-band stimuli. To mask simple difference tones, a pink noise was generated (Ivie, IE-20B) with a spectrum level of 27 dB at 100 Hz. This level was chosen on the basis of Pressnitzer and Patterson's (2001) study of combination tones.[1] The noise was low passed at 10 kHz using a 90 dB/oct filter (Kemo VBF8, mk4), and was presented continuously throughout a set of runs. Although this noise would not have masked cubic difference tones, any such tones should have had a negligible effect, owing to their coincidence with components in the skirts of the passband.

All stimuli were generated ''on the fly'' and played out at a sampling rate of 50 kHz using a 16-bit digital-to-analog converter (TDT DD1) and anti-aliasing filter. A new randomization was used each time a random-phase stimulus was generated. Stimuli were then passed through a programmable attenuator (TDT PA4), a headphone buffer (TDT HB6), a manual attenuator (Hatfield 2125), and were presented diotically over Sennheiser HD580 earphones in a double-walled sound-attenuating chamber (Industrial Acoustics).

4. Procedure

Thresholds for detecting anisochrony in the tone sequences were measured using a two-interval forced-choice (2IFC) method with a three-down one-up adaptive staircase (Levitt, 1971). This procedure estimates the 79.4% correct point on the psychometric function. A trial consisted of a reference and a test sequence, presented in random order across the two intervals. In the reference sequence, the position of each tone B was set to the exact midpoint in time between its neighboring tones A. The test sequence was configured in the same way for the first six AB cycles, but tone B was then progressively delayed in equal steps over the next four cycles, and the cumulated delay was maintained for the final two cycles (see Fig. 1). This type of sequence was first used by Cusack and Roberts (2000, experiment 2). Tone B was delayed, rather than advanced, owing to van Noorden's (1975, p. 59, Fig. 6.3) finding that listeners were more sensitive to that change.

Listeners used keys on a response box to start each trial and to indicate in which interval tone B was delayed with respect to the midpoint. Feedback was given after each trial. The initial set of isochronous cycles was intended to allow streaming to build up substantially before the progressive delay was introduced. When perceptible, the delay led to the impression of a change from a regular to an irregular rhythm. The introduction of the delay progressively over four AB cycles ensured that attending solely to the set of tones B was a poor listening strategy. The silent gap between the two intervals of each trial was set to 1 s. Bregman's (1978) data indicated that much of the cumulative effect of repetition on streaming decays in the first second of silence. The gap between a response and the beginning of the next trial was also set to a minimum of 1 s.

The maximum delay that could be imposed on tone B without causing a temporal overlap with the following tone A was 40 ms, and this value was used as the starting value for the adaptive staircase. This delay was typically well above threshold for the listeners in this study. The delay was varied on a logarithmic scale. Up to the first two turnpoints, the step size was a factor of 1.414. It was then reduced to 1.189 and held constant for another four turnpoints. Threshold estimates were defined as the geometric mean of the delay at the last four turnpoints. Whenever the standard deviation of the log values exceeded 0.2 (no occurrences beyond initial training), that run was rejected and replaced. Whenever a subject chose the wrong interval ten times when the delay was maximal (40 ms), the run was terminated and 40 ms was taken as the estimate. The estimate for any completed run in which the correct responses fell below 79.4% was also set to 40 ms. Together, these instances amounted to about 10% of runs. The geometric mean of three estimates was taken as the

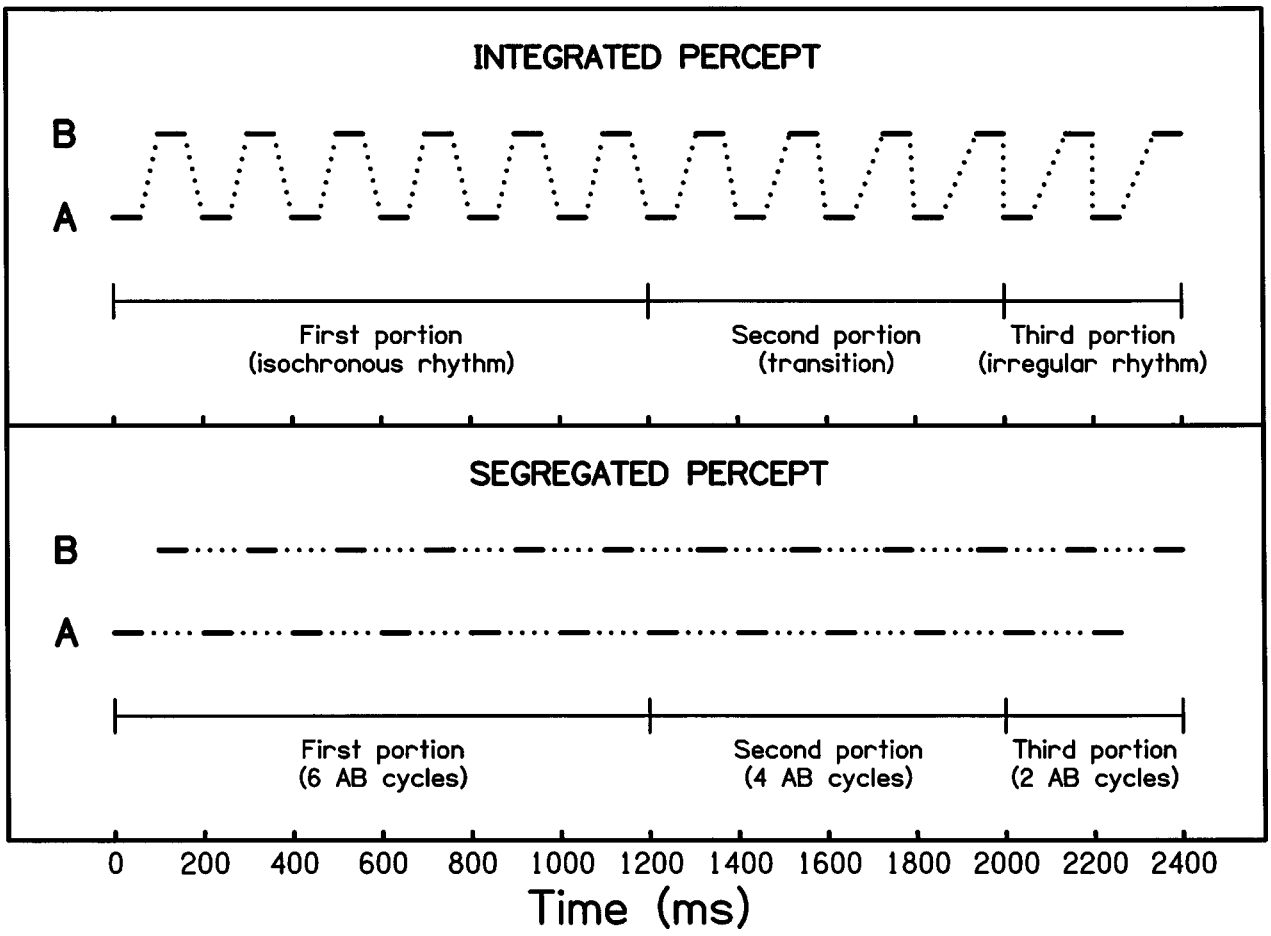


FIG. 1. Schematic of the test sequences used in experiment 1. Each complex tone is depicted using a short solid line and tones heard as belonging to the same auditory stream are linked using dotted lines. The upper and lower panels represent a sequence heard as integrated (one stream) or segregated (two streams), respectively. The progressive delay of tone B is difficult to detect when tones A and B are heard as arising from independent sources.

threshold. The only exceptions were instances in which the standard deviation of the three log values exceeded 0.2 (about 5% of runs), in which case an extra run took place and the threshold was taken as the geometric mean of all four estimates.

The main experiment was divided into four sessions. The first three sessions comprised 18 runs each, and the fourth comprised the remaining 12 runs plus any extras required (see above). The running order was randomized differently for each subject. Any extra runs required were sequenced in their original order. A typical session lasted about $1\frac{1}{2}$ h. Beforehand, subjects attended a combined screening and training session. This session began with three runs of the Hc-Hc combination, for which temporal discrimination was expected to be good. An extra run was undertaken if one of the three runs failed to meet the standard deviation criterion. No more than one extra run was required for any subject. The criterion for passing the screening procedure was a threshold of 20 ms or less. All subjects passed this criterion. Each subject then received training comprising a representative sample of eight runs in a predetermined order (Hc-Ha, Hc-Hr, La-La, Lc-Hc, Lc-Hr, Lr-Lr, Hc-Lc, and Hc-La).

B. Results

Figure 2 displays temporal discrimination thresholds for different combinations of tones A and B. There are clear effects of band difference and phase difference on threshold. Thresholds for the control sequences (Ha-Ha, Hr-Hr, La-La, Lr-Lr), plus Hc-Hc and Lc-Lc, are shown in Table I. The table shows that phase relations *per se* have little effect on threshold, and so the observed effect of *differences* in phase between tones A and B is genuine. The data were assessed using a three-way repeated-measures analysis of variance (ANOVA). The independent variables were band difference (0, 1, or 2 steps for L, M, and H pairings), phase difference (c-c, c-a, c-r), and passband of tone A (H or L). Data for the control sequences were not included in the analysis. The main effect of band difference was highly significant [means: 0=22.2 ms, 1=28.7 ms, 2=32.7 ms; $F(2,10)=70.86$, $p<0.001$], and so was the main effect of phase difference [means: c-c=21.7 ms, c-a=30.4 ms, c-r=31.4 ms; $F(2,10)=41.80$, $p<0.001$]. In contrast, the main effect of the passband of A was not significant [means: H=28.2 ms, L=27.5 ms; $F(1,5)=0.13$, $p>0.5$]. Only one interaction term, band difference × phase difference, was significant [$F(4,60)=8.76$, $p<0.001$]. This result probably reflects a ceiling effect on performance. First, the introduction of a phase difference is unlikely to have much effect on performance when a difference in spectral band has already induced substantial segregation. Indeed, when this happens, we assume that performance is governed instead by the detection of temporal delays *within* the set of tones B. Second, the procedure re-

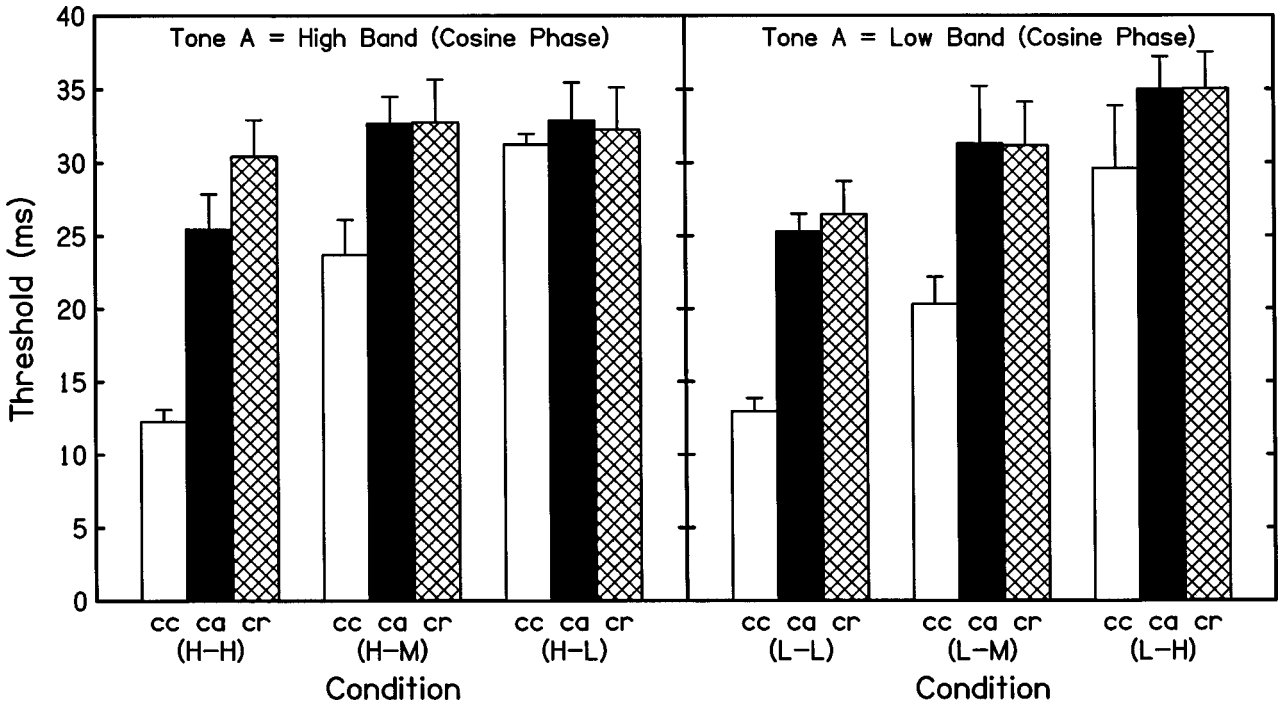


FIG. 2. Mean threshold for six subjects, with intersubject standard errors, for detection of the delay of tone B. Left and right panels show the results for conditions in which tone A was high and cosine (Hc) or low and cosine (Lc), respectively. The phase of tone B (cosine, alternating, or random) is depicted using open, filled, or hatched histogram bars, respectively, in each panel. The results for band differences of 0, 1, and 2 steps are depicted by sets of histogram bars arranged from left to right in each panel. A delay of 40 ms was the maximum possible value.

turned a threshold estimate of 40 ms (the maximum delay possible) on any trial for which a listener could not reliably discriminate the stimuli.

Pairwise comparisons were performed using the least-significant-difference (LSD) test on those factors identified as significant by the ANOVA.[2] All comparisons within the band-difference factor showed significant effects on threshold [(0 vs 1 steps: $t(10)=7.30$, $p<0.001$); (0 vs 2 steps: $t(10)=11.80$, $p<0.001$); (1 vs 2 steps: $t(10)=4.50$, $p<0.002$)]. This reflects the substantial differences in spectral center-of-gravity between tones A and B resulting from the passband differences. Pairwise comparisons within the phase-difference factor showed that changing tone B from cosine phase to either alternating or random phase increased the threshold significantly [$t(10)=7.49$, $p<0.001$; $t(10)=8.29$, $p<0.001$]. It did not matter which of these two phases was chosen for tone B [$t(10)=0.80$, ns]. The effect of component phase was dependent on the band difference between tones A and B. When the band difference was two steps, phase differences did not have any significant effect on threshold, undoubtedly due to the ceiling effects discussed earlier. Of particular interest is the effect of phase differences in the absence of a difference in band. As for the main effect, changing tone B from cosine phase to either alternating or random phase increased threshold significantly [$t(60)=7.32$, $p<0.001$; $t(60)=9.09$, $p<0.001$], and it did not matter which of these two phases was chosen for tone B [$t(60)=1.77$, ns]. The same pattern was seen for a band difference of one step.

TABLE I. Mean thresholds (ms), with intersubject standard errors, for the control sequences and corresponding cosine-phase stimuli used in experiment 1. The results for different bands and component phases (tone A−tone B) are shown by columns and rows, respectively.

	H-H	L-L
c-c	12.28 (0.81)	12.95 (0.92)
a-a	12.08 (1.16)	11.70 (0.40)
r-r	12.94 (1.31)	12.11 (0.87)

C. Discussion

The most important result is the dependence of threshold on phase differences between tones A and B when there is no difference in passband. This finding indicates, for the first time, a *substantial* effect of phase differences on stream segregation in the *absence* of differences in the power spectrum between successive tones. In fact, thresholds roughly doubled in size (ms) when a phase difference was introduced. Our finding contrasts strikingly with the weak effect on stream segregation of differences in *F*0 reported by Vliegen *et al.* (1999). In addition, the finding that band difference also has a substantial effect on threshold is consistent with the well-established importance of peripheral channeling in stream segregation (e.g., Hartmann and Johnson, 1991). Thresholds for sequences in which tones A and B were identical were similar to those observed for the most closely corresponding stimuli in Vliegen *et al.*'s study.

III. EXPERIMENT 2

The results of experiment 1 support the idea that primitive stream segregation can occur in the absence of spectral cues. Although temporal discrimination thresholds have the advantage of being a performance measure, they provide

only an indirect measure of auditory organization. Direct measures can be obtained using subjective report. Grimault *et al.* (2000) recently used this approach to show that sequences of unresolved harmonic complexes [as defined by the criterion of Shackleton and Carlyon (1994)] can undergo stream segregation in the absence of spectral cues. Like Vliegen *et al.* (1999), they found that this effect is modest compared to that produced by spectral differences. For example, consider their findings when the F0's of tones A and B were set to 250 and 176 Hz, respectively. The degree of segregation observed was 87% when the common passband was low (125–625 Hz), where the *F*0 difference between A and B led to substantial changes in excitation pattern, but only 10% when the common passband was high (3900–5400 Hz), where the *F*0 difference did not lead to substantial changes in excitation pattern.

This apparent difference in strength of stream segregation may in part reflect the type of subjective measure used by Grimault *et al.* (2000). Their listeners were required to make a binary judgment (one or two sources) at the end of sequences of 4 s. The build up of streaming, though considerable after 4 s, continues over a much longer time scale (Anstis and Saida, 1985). Also, continuous monitoring of the perceptual organization of a sequence is likely to be more sensitive than a single binary decision. Therefore, experiment 2 was designed to provide a more sensitive subjective measure of streaming. By using the same six subjects as had taken part in experiment 1, we were also able to explore the correlation between the data obtained from the two experiments. As far as we are aware, this approach has not been taken before.

A. Method

1. Overview

Listeners were asked to judge the extent of stream segregation for a repeating sequence of sounds in a task similar to that used by Anstis and Saida (1985). These authors used sequences of alternating low- and high-frequency pure tones, and observed that a range of separations led to an ambiguous percept that ''flipped'' back and forth between a single stream and separate low- and high-pitched streams. They asked listeners to indicate the perceived state of a sequence (one or two streams) throughout a trial, and used the proportion of time that the sequence was heard as integrated as their measure of streaming. Cusack and Roberts (1999) modified the sequence structure to make the judgment easier. Instead of a simple alternating sequence (Anstis and Saida, 1985), a more complex one like that used by van Noorden (1975) was employed (i.e., ABA-ABA-...). For this kind of sequence, differences in stream segregation lead to a difference in the rhythm perceived, providing a salient cue for the listener. We adopted this approach to measure the extent of stream segregation as a function of band difference and phase difference.

2. Stimuli and conditions

The sequence structure was like that of the reference sequence used in experiment 1, except that every alternate tone B was removed and replaced by silence. Hence, each stimulus comprised a sequence of tone triplets (ABA-ABA-ABA-....) with a tone offset-to-onset time of 140 ms between triplets. Each ABA- cycle was 400 ms in duration, and each sequence consisted of 75 cycles, giving a total duration of 30 s. These sequences were not modified by introducing delays, and so remained regular throughout. The set of stimuli tested was reduced from 22 to 18 by excluding the control sequences used in experiment 1. It was considered sufficient to include the sequences Hc-Hc and Lc-Lc to show that segregation does not occur when tones A and B are the same. Stimuli were again presented at 70 dB SPL and accompanied by pink noise.

3. Procedure

On each trial, subjects continuously monitored the perceptual organization of the sequence using a single key (flip-flop configuration) on a response box. The sequence evoked a distinctive ''galloping'' rhythm when heard as integrated, but was heard as two independent and monotonous sequences when segregated, with stream A having twice the tempo of stream B. Each trial began with an assumption of integration (Bregman, 1978). Subjects were given feedback on their current response (one or two streams). Following Anstis and Saida (1985), subjects were told not to try to hear the sequences in any particular way, but just to listen and respond to spontaneous changes in percept. Subjects were also told that they may occasionally hear a ''mixed'' percept, in which they can hear separate A and B sequences, but still experience a ''galloping'' sensation. Should this happen, they were asked to decide which of the two percepts was dominant. There was a gap of 3 s from the end of a trial until the next trial was ready to begin. Once the start key was pressed, there was a further 1 s until the next tone sequence began. Hence, the minimum gap between sequences was 4 s, which in most cases should have been sufficient for the complete decay of any built-up tendency to stream (Bregman, 1978).

The experiment was run in a single session of about 2 h. First, subjects received a demonstration in which they heard four ABA- sequences. These were like the experimental stimuli, except that tones A and B were pure tones. Tone A was fixed at 1000 Hz, and tone B was set successively to 1050, 1600, 1200, and 1400 Hz. These values were chosen to induce percepts that included well integrated, well segregated, and ambiguous. Second, subjects were given one set of these stimuli, in the same order, in which they became familiar with using key presses to categorize the current state of integration/segregation throughout a tone sequence. Third, subjects received a single randomized block of the 18 experimental stimuli for practice. Finally, subjects received six blocks of the experimental stimuli, giving a total of 108 trials. The order of these trials was randomized differently for each subject. Each estimate of the proportion of time a stimulus was heard as segregated was the mean of the six values obtained. These proportions were also computed for 5-s slices spanning the full duration of each sequence, so that the evolution of stream segregation across time could be examined.

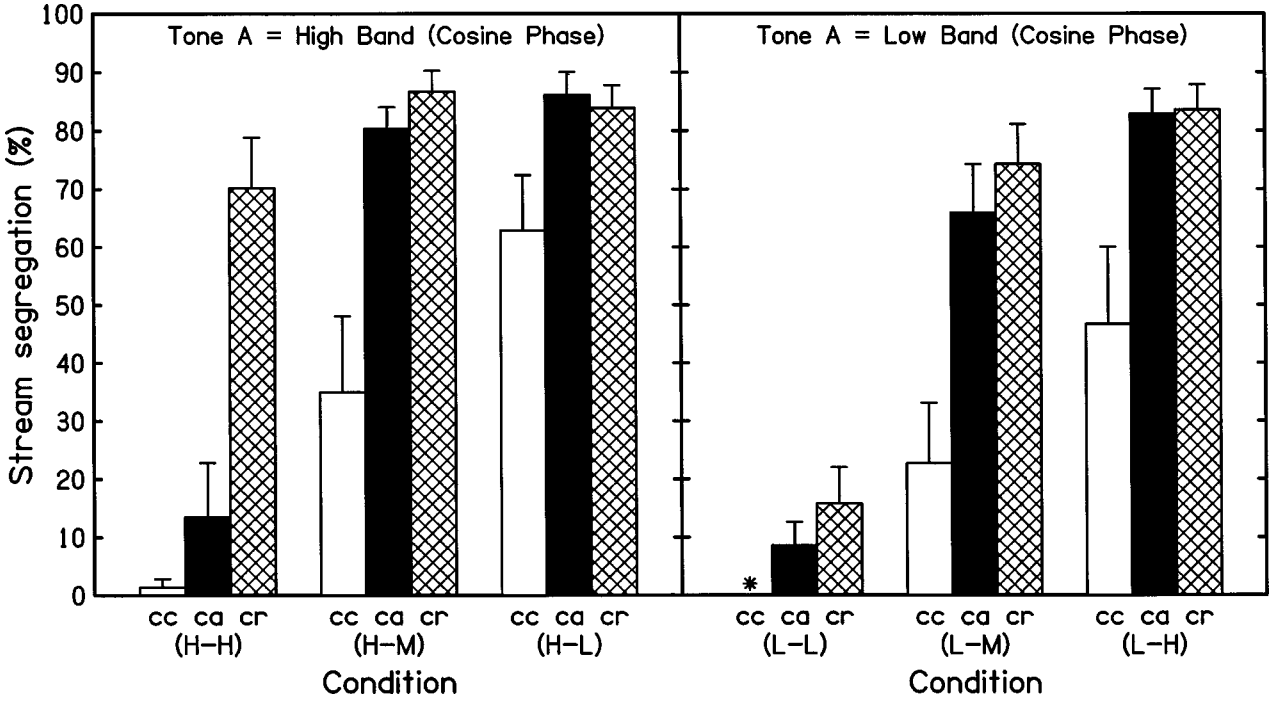


FIG. 3. Proportion of time for which the test sequences were heard as segregated (means and intersubject standard errors for six subjects). The left and right panels show the results for conditions in which tone A was high and cosine (Hc) or low and cosine (Lc), respectively. The phase of tone B (cosine, alternating, or random) is depicted using open, filled, or hatched histogram bars, respectively, in each panel. The results for band differences of 0, 1, and 2 steps are depicted by sets of histogram bars arranged from left to right in each panel. An asterisk indicates a stream segregation score of 0%.

B. Results

The data were assessed using a three-way repeated-measures ANOVA with the same factors as in experiment 1. Figure 3 shows clear effects of band difference and phase difference on the proportion of time for which the sequence was heard as segregated. As for experiment 1, the main effects of band difference (means: 0=18.3%, 1=59.7%, 2=74.4%) and phase difference (c-c=28.1%, c-a=56.3%, c-r=68.0%) were highly significant [$F(2,10)=260.24$, $p<0.001$; $F(2,10)=47.08$, $p<0.001$]. However, the main effect of the passband of tone A (H=57.8%, L=43.8%) was also highly significant [$F(1,5)=1059.01$, $p<0.001$]. As for experiment 1, the interaction of band difference and phase difference was highly significant [$F(4,60)=5.94$, $p<0.001$] but, in contrast, so was the three-way interaction term [$F(4,60)=4.48$, $p=0.003$]. Also, the interaction of band A and phase difference was marginally significant [$F(2,60)=3.23$, $p=0.047$]. The interaction of band A and band difference was not significant. Figure 4 shows that the tendency for the stimuli to be heard as segregated increased as listeners heard progressively more cycles of the sequence. This effect was most pronounced over the first 10 s.

The increased number of significant interactions makes the results harder to interpret. To simplify matters, we have

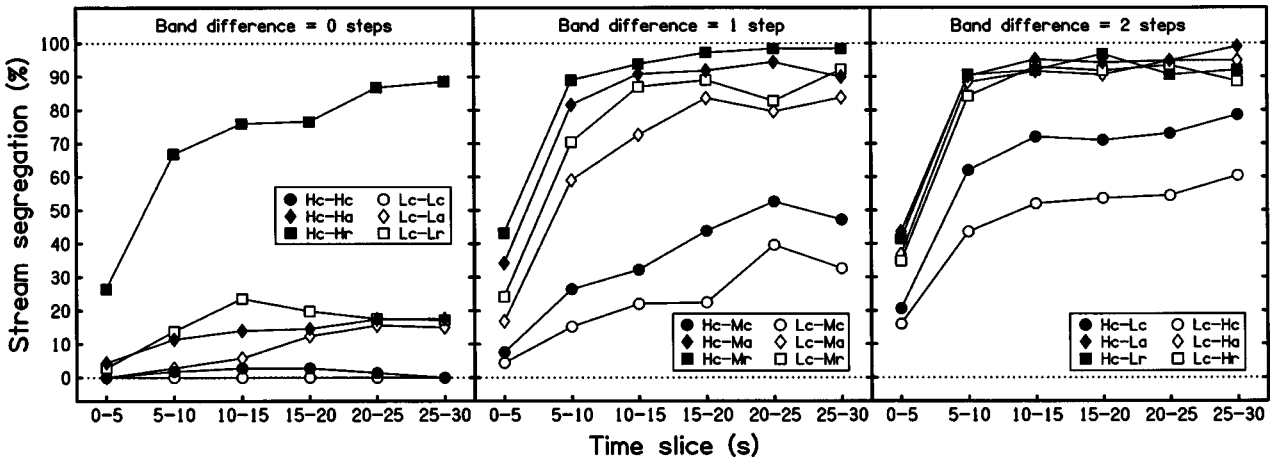


FIG. 4. Proportion of time for which the test sequences were heard as segregated across successive 5-s slices (means for six subjects). The left, middle, and right panels show the results for conditions in which the band difference was 0, 1, and 2 steps. The results for conditions in which tone A was Hc or Lc are depicted by filled and open symbols, respectively. The results for conditions in which the phase of tone B was cosine, alternating, or random are depicted by circles, diamonds, and squares, respectively. Each panel contains an inset identifying the conditions depicted.

restricted pairwise comparisons to those factors that were tested in experiment 1 (band difference, phase difference, and their interaction). All pairwise comparisons within the band-difference factor showed significant effects on stream segregation [(0 vs 1 steps: $t(10)=16.24$, $p<0.001$); (0 vs 2 steps: $t(10)=22.00$, $p<0.001$); (1 vs 2 steps: $t(10)=5.76$, $p<0.001$)]. Pairwise comparisons within the phase-difference factor showed that changing tone B from cosine phase to either alternating or random phase increased stream segregation significantly [$t(10)=6.68$, $p<0.001$; $t(10)=9.45$, $p<0.001$]. The effect of which of these phases was chosen for tone B was only marginally significant [$t(10)=2.76$, $p<0.05$]. Pairwise comparisons of discrimination thresholds in experiment 1 showed a similar overall pattern for these factors.

Once again, the effect of phase was dependent on the band difference between tones A and B, but the pattern of results was somewhat different from that seen in experiment 1. In the absence of a difference in band, changing tone B from cosine to alternating phase did not increase segregation significantly [$t(60)=1.55$, ns], but changing tone B to random phase did [$t(60)=6.37$, $p<0.001$]. This difference between alternating and random phase was significant [$t(60)=4.82$, $p<0.001$]. When the band difference was two steps, phase differences still led to increased segregation, but the difference between alternating and random phase disappeared [(c-c vs c-a: $t(60)=4.37$, $p<0.001$); (c-c vs c-r: $t(60)=4.47$, $p<0.001$); (c-a vs c-r: $t(60)=0.11$, ns)]. The same pattern was seen for a band difference of one step. Once a band difference was present, changing tone B to either alternating or random phase was similarly effective at increasing segregation.

C. Discussion

The dependence of stream segregation on differences in band and phase between tones A and B is broadly similar to the pattern observed for discrimination thresholds in experiment 1. In particular, differences in passband have a large effect on segregation and a substantial increase can be produced by a phase difference even when A and B have the same power spectrum. Spearman's ranked-order correlation coefficient for scores in the two experiments was 0.69 ($p<0.001$). However, there are differences between the data sets that merit consideration.

(1) Why were phase differences able to increase stream segregation further when there was a two-step difference in passband, whereas phase differences did not increase thresholds in experiment 1 under these conditions? First, the thresholds measured in experiment 1 were subject to ceiling effects (maximum delay was 40 ms). Second, the task in experiment 1 could perhaps be done on the basis of the tone-B set alone for longer delays (the increase in offset-to-onset time between consecutive tones B in the transitional portion equals a quarter of the total delay). Hence, stream segregation was probably not the factor governing performance for large differences between tones A and B. Neither of these other factors applied in experiment 2. Figure 3 shows that a two-step band difference produced a mean segregation of around 50%–60% in the absence of phase differences, allowing scope for the additional effects of phase differences.

(2) Why was substantial segregation obtained in the absence of spectral cues only for the Hc-Hr combination? The results of experiment 1 also show a trend towards that combination being most effective at increasing thresholds (see Fig. 2), but this effect was not significant. The greater stream segregation associated with the Hc-Hr combination suggests that the perceptual difference between these tones was greater than for the other combinations. An account can be offered in terms of how the physical and perceptual properties of cosine-, alternating-, and random-phase stimuli depend on passband. According to Ritsma (1962, 1963), harmonics below 5 kHz and below the 20th are required for periodicity pitch to be heard. Our H-band stimuli did not meet the latter criterion, but were able to evoke a kind of ''rattle'' pitch that was dependent on component phase. This is probably because they contained many more harmonics than Ritsma's stimuli. However, the change from L-band to H-band would have had the greatest effect for our random-phase stimuli. These stimuli resemble filtered noises, for which the number of envelope maxima per period is directly proportional to the bandwidth (see Rice, 1954). In practice, the number of envelope fluctuations per period would have been limited by auditory filtering, but the H-band stimuli had a higher center frequency than the L-band stimuli, and so the number of envelope peaks per period would still have been greater for the Hr stimulus than for the Lr stimulus. Therefore, if the cue that induces stream segregation is based on average envelope fluctuation rate, the difference between Hc and Hr would be greater than between Lc and Lr. The greater number of fluctuations per period for Hr vs Lr would not affect pitch value, because only the F0 period would be common across channels. Informal listening by the authors suggests that Hr and Lr did not differ greatly in pitch strength, but rather that the difference in *timbre* between Hc and Hr was greater than between Lc and Lr. This difference, which presumably relates to differences in envelope fluctuations, is best described in terms of ''harshness.'' The apparent discrepancy between the effects of differences in component phase for tones A and B in the two experiments may reflect a nonlinear relationship between temporal discrimination thresholds and streaming judgments. Perhaps there is a transitional region around the TCB in which small changes in threshold correspond to large changes in stream segregation.

(3) Why was the degree of stream segregation dependent on whether tone A was high band or low band? The effect of passband for tone A may reflect the sequence structure used in experiment 2. Specifically, the removal of every alternate tone B resulted in sequences comprising twice as many tokens of tone A as of tone B, which introduces an asymmetry not present in experiment 1. Hence, for a given band difference, the spectral center-of-gravity of the whole sequence was higher when tone A=Hc than when tone A=Lc. However, it is not obvious why this should lead to greater stream segregation.

Finally, it is interesting to compare our data on the change in stream segregation across time (see Fig. 4) with

those of Anstis and Saida (1985) for alternating pure-tone sequences without silences between consecutive tones. They used stimuli with a fixed frequency interval between tones A and B, but with an adjustable tone duration (and hence TRT). Subjects were asked to listen to the AB sequence and to reduce the rate as soon as it began to segregate. This process continued until the sequence was terminated after 60 s. A log-log plot of rate against time was best described as two straight lines; one sloping steeply for the first 10 s or so and another sloping about half as steeply thereafter. The tendency for stream segregation to increase appeared to be incomplete even after 60 s. Although our sequences were composed of very different kinds of tone, our findings are broadly similar to those of Anstis and Saida (1985). The most rapid increase in stream segregation takes place in the first two time slices (0–10 s), and is followed by a more gradual change that appears incomplete even in the final time slice (25–30 s). This pattern confirms the advantage of using long sequences in assessing the degree of stream segregation (cf. Grimault *et al.*, 2000).

IV. GENERAL DISCUSSION

The principal finding of our experiments is that substantial stream segregation can be induced by phase differences for sequences of complex tones with identical power spectra. This effect appears to be much larger than the effect of changing *F*0 reported by Vliegen *et al.* (1999) and Grimault *et al.* (2000). The fact that the effect of phase differences occurs when the listener is required to integrate the sequence (experiment 1) indicates that this effect is mediated by a primitive grouping process, which is automatic and obligatory. This section considers further (*i*) our assumption that pitch and timbre are the percepts mediating the effect of component phase; (*ii*) the relationship between temporal discrimination and sequential grouping; and (*iii*) the peripheral-channeling hypothesis in relation to our findings and those of other recent studies.

A. Phase differences may result in loudness differences

Carlyon and Datta (1997) have argued that the fast-acting compression characteristic of auditory responses (see Moore and Oxenham, 1998) must operate on a cycle-by-cycle basis. This implies that complex tones with identical power spectra but different phase spectra may evoke different excitation patterns. In particular, a stimulus with a high peak factor (ratio of peak to rms value) will be more compressed than one with a low peak factor, and so will evoke less overall excitation. Carlyon and Datta (1997) used forward masking to compare the excitation evoked by wideband harmonic stimuli that differed only in their phase spectra. The stimuli comprised components summed either in positive or negative Schroeder phase (Schroeder, 1970); the former and latter produce waveforms on the basilar membrane with high and low peak factors, respectively. As predicted, Carlyon and Datta (1997) observed that positive-phase stimuli produced significantly less forward masking and that the components responsible for the masking were judged as quieter than their negative-phase counterparts (though they did not quantify the latter effect). This finding suggests that our observations of stream segregation for sequences of tones with identical power spectra may have arisen from differences in *loudness* rather than in pitch and timbre. However, we believe that this is unlikely to account for our results:

First, Gockel *et al.* (2002) found, using a loudness-matching task, that harmonic complex tones with components added in cosine phase (high peak factor) were heard as louder than tones whose components had random phase (low peak factor). This effect was both small (only a few dB) and contrary to the predictions of the fast-acting compression hypothesis. Moreover, our listeners did not report marked differences in loudness between stimuli with different phase spectra; rather, it was changes in pitch value, pitch strength, and timbre that were reported as salient differences. Second, although stimuli can be selected on the basis of loudness differences (van Noorden, 1975; Hartmann and Johnson, 1991), there is no clear evidence that primitive stream segregation can be induced by such differences (see Bregman, 1990, pp. 126–127).

Overall, it seems likely that differences in loudness had only a marginal effect; rather, we believe that the effects of phase on pitch (value, strength, ambiguity) and on timbre account for our findings. In any case, we can be confident that any differences in excitation pattern resulting from differences in the phase spectrum would not have affected *which* peripheral channels were excited. Hence, there would have been no differences in peripheral channeling, in the sense meant by Hartmann and Johnson (1991), between stimuli distinguished only by their phase spectra. Although recent estimates of frequency selectivity (Shera *et al.*, 2002; Moore, 2002) suggest that Vliegen *et al.* (1999) underestimated the differences in excitation pattern between their unresolved harmonic complexes as *F*0 was varied, this point also applies to their study.

B. Temporal discrimination and the temporal coherence boundary

The question of whether a discrete TCB is compatible with evidence that temporal discrimination thresholds vary continuously merits discussion. Thresholds for detecting the temporal displacement of tone B from isochrony in ABAB... and ABA-ABA-... sequences increase in a near-linear manner on a logarithmic scale of frequency interval (van Noorden, 1975; Vliegen *et al.*, 1999). Furthermore, Vliegen *et al.* pointed out that there was no discontinuity in the performance function to indicate the TCB; indeed, thresholds continued to increase for frequency intervals well beyond this boundary. They raised the possibility that temporal discrimination and stream segregation are affected *independently* by large frequency differences between successive tones—perhaps large differences in a temporal discrimination task impair performance simply by distracting the listener. However, this suggestion is not easily reconciled with the known effects of sequential grouping on gap detection and discrimination (e.g., van Noorden, 1975).

Vliegen *et al.* (1999) argued that the lack of a discontinuity in the temporal-discrimination function may instead

reflect a gradual transition between integration and segregation. These states could either lie at opposite ends of a continuous scale or be discrete percepts, for which the transition reflects a change in the probability that one (or other) will be heard. Subjective reports from our experiment 2 indicate that the latter is a better description. Subjects typically report hearing perceptual reversals between the two states of integration and segregation, in a manner akin to classic ambiguous figures like ''vase/faces'' (see Attneave, 1971). Fisher (1967) showed such a probability transition for a range of ambiguous visual figures by varying the cues favoring one or other interpretation. We suggest that the apparently discrete nature of the TCB arises from using the method of adjustment to measure it (van Noorden, 1975).

C. The peripheral-channeling hypothesis revisited

The experiments reported here have gone further than Cusack and Roberts (2000) by showing that differences in peripheral channeling are not even a minimum requirement for primitive stream segregation. Moreover, unlike those of Vliegen *et al.* (1999) and Grimault *et al.* (2000), our findings have shown that unresolved harmonic complexes in the same passband can undergo *substantial* stream segregation when stimulus parameters are optimized. Grimault *et al.* (2002) have come to a similar conclusion using stimuli considerably different from ours. They controlled for spectral and temporal fine-structure cues by using repeating sequences of ABA- triplets in which sounds A and B were amplitude-modulated bursts of broadband noise with a common passband. Sounds of this kind evoke a weak but discernible pitch (Burns and Viemeister, 1976). Grimault *et al.* (2002) showed using subjective report that sequences were generally heard as integrated for differences in AM rate between A and B below 0.75 oct and as segregated for differences above 1.03 oct. They also presented data indicating that streaming behavior did not relate to listeners' AM-detection thresholds, and that differences in perceived AM depth probably did not mediate the observed effect of differences in AM rate.

The popularity of the peripheral-channeling hypothesis owes much to its simplicity and its ability to explain a wide range of findings from the sequential grouping literature. However, its relative success may simply reflect the major contribution to timbre that is made by spectral characteristics like center-of-gravity and compactness/dispersion (e.g., von Bismarck, 1974; Grey, 1977). Similarly, differences in pitch between pure tones or resolved harmonic complexes are closely related to differences in excitation pattern. The relation between measures of the perceptual distance between stimuli and their likelihood of undergoing stream segregation has received little attention, but the research of Iverson (1995) is an exception. His work has suggested that, when other factors are controlled, similarity of timbre is a good predictor of the extent to which sequences of sounds will segregate. Moore and Gockel (2002) propose that there is a direct relation between the degree of perceptual difference between sounds and the extent to which those sounds will undergo stream segregation, irrespective of the cause of that perceptual difference. This proposal is compatible with our finding that, to the extent that two sounds can be made different in timbre or pitch by varying their temporal characteristics, stream segregation can occur in the absence of differences in the power spectrum. However, it is not clear that all salient perceptual changes can influence stream segregation. For example, Grimault *et al.* (2002) found that the TCB for noise bursts differing in AM rate occurred at a similar interval to that for unresolved harmonic complexes differing in *F*0 (Grimault *et al.*, 2000). Given that the former stimuli evoke a much weaker pitch than the latter, this suggests that pitch *strength* plays a negligible role in stream segregation (beyond the point at which a pitch *value* can be perceived). Further work is needed to evaluate fully the notion that the Gestalt principle of similarity provides an adequate description of auditory stream segregation. For now, the extent to which the *perceived similarity* of stimuli is ultimately the key factor governing their sequential grouping remains an open question.

D. Concluding remarks

Differences between consecutive sounds on either a spectral or a temporal dimension can affect their perceptual organization. The experiments presented here have shown that differences in auditory representation produced by phase differences between unresolved harmonics (most probably differences in timbre and pitch) can lead to substantial primitive stream segregation in the absence of significant spectral cues. Although the importance of spectral differences as a basis for stream segregation is well established, it is now clear that a comprehensive account of auditory sequential grouping cannot be provided by the peripheral-channeling hypothesis.

ACKNOWLEDGMENTS

The Medical Research Council (UK) provided support for Brian Moore and Brian Glasberg, and for the equipment used in this study. Our thanks go to Mark Georgeson and Mike Harris in relation to the literature on perceptual reversal in visual figures. We are also grateful to Valter Ciocca and Nicolas Grimault for their comments on an earlier version of this manuscript.

[1]Pressnitzer and Patterson (2001) reported psychophysical measurements of the difference tones produced by a set of upper harmonics (15–25) of 100 Hz, added in cosine phase. The distortion spectrum was composed of reintroduced lower harmonics. The most intense difference tone was the first, corresponding to the missing *F*0. Its level was in the range 10–15 dB below that of the primaries, which were each presented at 54 dB SPL. The level of the second difference tone was 25–30 dB below that of the primaries, and higher difference tones progressively declined in level. Our study used complex tones composed of unresolved harmonics of 100 Hz presented at a similar level. Therefore, we chose a level of pink noise sufficient to mask the first difference tone (at *F*0) in our cosine-phase stimuli. This was to guard against the possibility that changing the phase relations of our complexes from cosine phase to alternating or random phase might reduce the level of the associated distortion spectrum. Without a masking noise, the possibility that variations in the level of the distortion spectrum introduced secondary peripheral-channeling cues to affect streaming could not be ruled out. Taking the M-band as our reference, about 18 components fall in the passband (1768–3536 Hz). For a total power of 70 dB SPL, the level per component is 57.4 dB. Assuming a worst case of a first difference tone only 10 dB lower in level than the primaries, and a signal-to-masker ratio at threshold of 5 dB (Moore *et al.*, 1997, p. 230), the noise level required to

mask this difference tone is 42.4 dB/ERB. The ERB is about 35 Hz at 100 Hz, and so the required spectrum level of the pink noise at 100 Hz is 27 dB. Pink noise has a spectral roll-off of −3 dB/oct, which is sufficient to mask all higher difference tones.

[2]This procedure is known as the restricted LSD test (Snedecor and Cochran, 1967). It was used because tests like Tukey's HSD lack power when a large number of pairwise comparisons are possible. It should be noted that almost all comparisons quoted as significant in experiments 1 and 2 are associated with low p values. Therefore, it is reasonable to regard them as reliable.

Anstis, S., and Saida, S. (**1985**). ''Adaptation to auditory streaming of frequency-modulated tones,'' J. Exp. Psychol. Hum. Percept. Perform. **11**, 257–271.

Attneave, F. (**1971**). ''Multistability in perception,'' Sci. Am. **225**, 62–71.

Beauvois, M. W., and Meddis, R. (**1996**). ''Computer simulation of auditory stream segregation in alternating-tone sequences,'' J. Acoust. Soc. Am. **99**, 2270–2280.

Bilsen, F. A. (**1973**). ''On the influence of the number and phase of harmonics on the perceptibility of the pitch of complex signals,'' Acustica **28**, 60–65.

Bregman, A. S. (**1978**). ''Auditory streaming is cumulative,'' J. Exp. Psychol. Hum. Percept. Perform. **4**, 380–387.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).

Bregman, A. S., and Campbell, J. (**1971**). ''Primary auditory stream segregation and perception of order in rapid sequences of tones,'' J. Exp. Psychol. **89**, 244–249.

Bregman, A. S., Liao, C., and Levitan, R. (**1990**). ''Auditory grouping based on fundamental frequency and formant peak frequency,'' Can. J. Psychol. **44**, 400–413.

Broadbent, D. E., and Ladefoged, P. (**1959**). ''Auditory perception of temporal order,'' J. Acoust. Soc. Am. **31**, 1539.

Burns, E. M., and Viemeister, N. F. (**1976**). ''Nonspectral pitch,'' J. Acoust. Soc. Am. **60**, 863–869.

Carlyon, R. P., and Datta, A. J. (**1997**). ''Excitation produced by Schroeder-phase complexes: Evidence for fast-acting compression in the auditory system,'' J. Acoust. Soc. Am. **101**, 3636–3647.

Carlyon, R. P., Cusack, R., Foxton, J. M., and Robertson, I. H. (**2001**). ''Effects of attention and unilateral neglect on auditory stream segregation,'' J. Exp. Psychol. Hum. Percept. Perform. **27**, 115–127.

Cusack, R., and Roberts, B. (**1999**). ''Effects of similarity in bandwidth on the auditory sequential streaming of two-tone complexes,'' Perception **28**, 1281–1289.

Cusack, R., and Roberts, B. (**2000**). ''Effects of differences in timbre on sequential grouping,'' Percept. Psychophys. **62**, 1112–1120.

Dowling, W. J. (**1973**). ''The perception of interleaved melodies,'' Cogn. Psychol. **5**, 322–337.

Fisher, G. H. (**1967**). ''Measuring ambiguity,'' Am. J. Psychol. **80**, 541–557.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.

Gockel, H., Moore, B. C. J., and Patterson, R. D. (**2002**). ''Influence of component phase on the loudness of complex tones,'' Acta Acust. Acust. **88**, 369–377.

Grey, J. M. (**1977**). ''Multidimensional perceptual scaling of musical timbres,'' J. Acoust. Soc. Am. **61**, 1270–1277.

Grimault, N., Bacon, S. P., and Micheyl, C. (**2002**). ''Auditory stream segregation on the basis of amplitude-modulation rate,'' J. Acoust. Soc. Am. **111**, 1340–1348.

Grimault, N., Micheyl, C., Carlyon, R. P., Arthaud, P., and Collet, L. (**2000**). ''Influence of peripheral resolvability on the perceptual segregation of harmonic complex tones differing in fundamental frequency,'' J. Acoust. Soc. Am. **108**, 263–271.

Grose, J. H., Hall, J. W., Buss, E., and Hatch, D. (**2001**). ''Gap detection for similar and dissimilar gap markers,'' J. Acoust. Soc. Am. **109**, 1587–1595.

Hartmann, W. M., and Johnson, D. (**1991**). ''Stream segregation and peripheral channeling,'' Music Percept. **9**, 155–183.

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.

Iverson, P. (**1995**). ''Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes,'' J. Exp. Psychol. Hum. Percept. Perform. **21**, 751–763.

Kohlrausch, A., and Houtsma, A. J. M. (**1992**). ''Pitch related to spectral edges of broadband signals,'' in *Processing of Complex Sounds by the Auditory System*, edited by R. P. Carlyon, C. J. Darwin, and I. J. Russell (Oxford U.P., Oxford, UK), pp. 81–88.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

McCabe, S. L., and Denham, M. J. (**1997**). ''A model of auditory streaming,'' J. Acoust. Soc. Am. **101**, 1611–1621.

Miller, G. A., and Heise, G. A. (**1950**). ''The trill threshold,'' J. Acoust. Soc. Am. **22**, 637–638.

Moore, B. C. J. (**2002**). ''Frequency resolution,'' in *Genetics and the Function of the Auditory System*, edited by L. Tranebjærg, V. Christensen-Dalsgaard, T. Andersen, and T. Poulsen (Holmens Trykkeri, Copenhagen, Denmark), pp. 227–271.

Moore, B. C. J., and Gockel, H. (**2002**). ''Factors influencing sequential stream segregation,'' Acta Acust. Acust. **88**, 320–333.

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., and Rosen, S. M. (**1979**). ''Tune recognition with reduced pitch and interval information,'' Q. J. Exp. Psychol. **31**, 229–240.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). ''A model for the prediction of thresholds, loudness, and partial loudness,'' J. Audio Eng. Soc. **45**, 224–240.

Ortmann, O. (**1926**). ''On the melodic relativity of tones,'' Psychol. Monogr. **35**, 1–47 (Whole No. 162).

Pressnitzer, D., and Patterson, R. D. (**2001**). ''Distortion products and the pitch of harmonic complex tones,'' in *Physiological and Psychophysical Bases of Auditory Function*, edited by D. J. Breebaart, A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs, and R. Schoonhoven (Shaker, Maastrict, The Netherlands), pp. 84–91.

Rice, S. O. (**1954**). ''Mathematical analysis of random noise,'' in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York), pp. 133–294.

Ritsma, R. J. (**1962**). ''Existence region of the tonal residue. I,'' J. Acoust. Soc. Am. **34**, 1224–1229.

Ritsma, R. J. (**1963**). ''Existence region of the tonal residue. II,'' J. Acoust. Soc. Am. **35**, 1241–1245.

Schroeder, 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.

Shera, C. A., Guinan, J. J., and Oxenham, A. J. (**2002**). ''Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements,'' Proc. Natl. Acad. Sci. U.S.A. **99**, 3318–3323.

Singh, P. G. (**1987**). ''Perceptual organization of complex-tone sequences: A tradeoff between pitch and timbre?'' J. Acoust. Soc. Am. **82**, 886–899.

Singh, P. G., and Bregman, A. S. (**1997**). ''The influence of different timbre attributes on the perceptual segregation of complex-tone sequences,'' J. Acoust. Soc. Am. **102**, 1943–1952.

Smoorenburg, G. F. (**1972**). ''Audibility region of combination tones,'' J. Acoust. Soc. Am. **52**, 603–614.

Snedecor, G. W., and Cochran, W. G. (**1967**). *Statistical Methods*, 6th ed. (Iowa U.P., Ames, Iowa).

van Noorden, L. P. A. S. (**1975**). ''Temporal coherence in the perception of tone sequences,'' doctoral thesis, Eindhoven University of Technology, The Netherlands.

Vliegen, J., and Oxenham, A. J. (**1999**). ''Sequential stream segregation in the absence of spectral cues,'' J. Acoust. Soc. Am. **105**, 339–346.

Vliegen, J., Moore, B. C. J., and Oxenham, A. J. (**1999**). ''The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task,'' J. Acoust. Soc. Am. **106**, 938–945.

von Bismarck, G. (**1974**). ''Timbre of steady sounds: A factorial investigation of its verbal attributes,'' Acustica **30**, 146–159.

Warren, R. M., Obusek, C. J., Farmer, R. M., and Warren, R. P. (**1969**). ''Auditory sequence: Confusion of patterns other than speech and music,'' Science **164**, 586–587.

Wessel, D. L. (**1979**). ''Timbre space as a musical control structure,'' Comput. Music J. **3**, 45–52.

The effect of spatial separation on informational and energetic masking of speech[a)]

Tanya L. Arbogast,[b)] Christine R. Mason, and Gerald Kidd, Jr.
Department of Communication Disorders and The Hearing Research Center, Boston University, 635 Commonwealth Avenue, Boston, Massachusetts 02215

(Received 28 February 2002; revised 13 July 2002; accepted 6 August 2002)

The effect of spatial separation of sources on the masking of a speech signal was investigated for three types of maskers, ranging from energetic to informational. Normal-hearing listeners performed a closed-set speech identification task in the presence of a masker at various signal-to-noise ratios. Stimuli were presented in a quiet sound field. The signal was played from 0° azimuth and a masker was played either from the same location or from 90° to the right. Signals and maskers were derived from sentences that were preprocessed by a modified cochlear-implant simulation program that filtered each sentence into 15 frequency bands, extracted the envelopes from each band, and used these envelopes to modulate pure tones at the center frequencies of the bands. In each trial, the signal was generated by summing together eight randomly selected frequency bands from the preprocessed signal sentence. Three maskers were derived from the preprocessed masker sentences: (1) different-band sentence, which was generated by summing together six randomly selected frequency bands out of the seven bands not present in the signal (resulting in primarily informational masking); (2) different-band noise, which was generated by convolving the different-band sentence with Gaussian noise; and (3) same-band noise, which was generated by summing the same eight bands from the preprocessed masker sentence that were used in the signal sentence and convolving the result with Gaussian noise (resulting in primarily energetic masking). Results revealed that in the different-band sentence masker, the effect of spatial separation averaged 18 dB (at 51% correct), while in the different-band and same-band noise maskers the effect was less than 10 dB. These results suggest that, in these conditions, the advantage due to spatial separation of sources is greater for informational masking than for energetic masking. [DOI: 10.1121/1.1510141]

PACS numbers: 43.66.Dc, 43.66.Lj, 43.66.Pn [MRL]

I. INTRODUCTION

The ''cocktail party'' (Cherry, 1953) environment contains multiple speech sources and often involves the problem of understanding one talker while ignoring one or more speaking at the same time. This environment is acoustically complex. There are multiple paths of direct sound, one from each source, and in a nonanechoic room, there are also several additional paths of reverberant sound associated with each sound source. The ability to separate simultaneous spectral components into individual sound sources and attend only to the source of interest is both an impressive and essential task for successful communication in the cocktail party environment.

When speech is embedded in background speech, as in the cocktail party, two different types of masking occur—*energetic* and *informational* (Freyman *et al.*, 1999, 2001; Brungart, 2001a; Brungart *et al.*, 2001). Traditional masking theory posits the existence of a bank of overlapping auditory filters through which sound is filtered. Energetic or peripheral masking occurs because the masker energy, or portion thereof, falls within the same auditory filter as the signal energy, resulting in a degraded representation of the signal in the auditory system from the cochlea and beyond. Informational masking occurs in the absence of, or in addition to, overlapping spectra between the signal and masker; there is good representation of the signal in the neural firing patterns beyond the cochlea. Traditionally, informational masking is believed to stem from uncertainty in the stimulus (e.g., Pollack, 1975; Watson *et al.*, 1976; Watson, 1987; Neff and Green, 1987; Lutfi, 1989), although qualitative similarity between the signal and masker may also play a role (e.g., Leek *et al.*, 1991; Neff, 1995; Brungart, 2001a; Kidd *et al.*, 2002). The interpretation is that the uncertainty in the stimulus makes it difficult to detect the signal, or the signal fuses with the masker due to similarity, rendering it difficult to extract information about any one frequency component. Informational masking is thought to operate at a higher level in the auditory system than the periphery. Properties of informational masking have been well-studied using nonspeech stimuli, which affords good control over the type of masking produced. Similar research using speech stimuli is difficult to find, although several studies have attributed certain outcomes to the presence of informational masking in concurrent speech tasks (e.g., Brungart, 2001a; Brungart *et al.*, 2001; Freyman *et al.*, 1999, 2001).

Cherry (1953) identified spatial separation between mul-

[a)]Portions of this research were presented at the Midwinter Meeting of the Association for Research in Otolaryngology, St. Petersburg Beach, FL, January, 2002. This work has also been submitted by the first author in partial fulfillment of the requirements for the ScD degree.

[b)]Electronic mail: rohtla@bu.edu

tiple sound sources as one of several potential means by which listeners might solve the cocktail party problem. A large body of research exists on the benefits of spatial separation between speech and noise, and to a lesser extent between two speech sources. For a normal-hearing listener, a 90° spatial separation between signal speech and the masker (speech or noise) improves speech recognition performance, often by 5 to 10 dB depending on room characteristics, speech materials, etc. (Duquesnoy, 1983; Gelfand *et al.*, 1988; Bronkhorst and Plomp, 1990; Nilsson *et al.*, 1992; Bronkhorst and Plomp, 1992; Peissig and Kollmeier, 1997; Hawley, 2000). Spatial separation between a signal and masker gives rise to several cues that can account for the improved speech recognition: better ear advantage due to headshadow, binaural interaction, and auditory object formation. The latter cue is a function of auditory scene analysis (cf. Bregman, 1990), the organization of simultaneous sound components into auditory objects, which once formed, can be segregated (perceived by the listener as separate objects). A large proportion of the improvement in speech recognition performance reported due to spatial separation is often attributed to headshadow and binaural interaction cues only (Zurek, 1993).

However, a handful of studies have found spatial advantages that cannot be completely accounted for by a combination of headshadow and binaural interaction effects. In a nonanechoic room, Kidd *et al.* (1998) found spatial advantages of 5 to 40 dB for the identification of nonspeech puretone frequency patterns in the presence of a multitone, informational masker. By comparison, the spatial advantage in Gaussian noise was 10 dB or less, the majority of which was accounted for by headshadow. In a virtual sound field, Ericson and McKinley (1997) studied the spatial advantage of separating approximately equal-level signal and masker talkers by 90° in the presence of a diotic pink noise of varying levels. The size of the separation advantage was not directly reported, but was estimated from their graphs of diotic and 90° separation psychometric functions at a performance level of 50% correct. The advantage was approximately 14 dB for opposite-sex signal and masker talkers, and about 21 dB for two male talkers. When the two talkers were female, the advantage appeared to be at least as large as for the two male talkers, but because the slopes of the functions were vastly different and the diotic functions did not exceed 45% correct, it was difficult to determine. These advantages, specified in terms of the level of the pink noise, and not the signal-to-masker ratio of the two talkers, are difficult to compare with other research which used a noise and/or a speech masker separately.

When both the signal and masker are speech stimuli, the *perception* of spatial separation between the signal and masker can be sufficient for a significant speech recognition advantage to occur, consistent with auditory object formation. Freyman *et al.* (1999) used localization dominance in the precedence effect to demonstrate this. When listeners were asked to identify nonsense sentences spoken by a female talker in the presence of speech-shaped noise, they showed an average 8-dB spatial release from masking with a 60° separation. In the presence of another female talker, the release was nearly 14 dB. This larger release was attributed to the presence of informational masking in the speech masker, which allowed the listener to use the perceived separation of sources as a cue to source segregation, leading to a larger spatial release than possible with headshadow and binaural interaction cues alone. When the precedence effect was used to create a perceived spatial separation of 60° between the signal and speech-shaped noise, there was no release from masking. However, when the masker was another talker, a significant release of 4 to 9 dB was obtained, even though both stimuli were presented at both locations. Spatial release from the speech masker was 5 to 8 dB greater than that predicted by detection thresholds in the speech-shaped noise for the same location/perceived location configurations. In a subsequent study of perceived spatial separation, Freyman *et al.* (2001) reported a similar spatial release (in the precedence condition) from a two-talker, same-sex masker in English, a foreign language, and a smaller release in reversed English. But, similar to a flat-envelope speech-shaped noise, essentially no spatial advantage was found for speech masked by wideband speech envelope modulated noise, or by 8-band speech envelope modulated noise, both derived from the two-talker masker.

The above studies indicate that spatial release from a speech masker is not necessarily subject to the same rules and limitations as release from a noise masker. Informational masking in the speech masker may be the key to this difference between the maskers. Many previous studies on the spatial release from a speech masker have generally found moderate effects not much larger than that for a noise masker. This may be due to the presence of a significant amount of energetic masking in the task. There may be a larger spatial release from an informational masker as a result of perceptual segregation of signal and masker into separate auditory objects. It is difficult, however, to isolate the informational portion of the masking using speech stimuli since speech is broadband. Two speech sources will have considerable spectral overlap and therefore portions of each speech source will fall within the same auditory filters, resulting in peripheral masking. However, the advantages of using speech stimuli include a less abstract task for the listener and a more direct application to the problem of the cocktail party.

In order to study informational masking using speech stimuli, peripheral masking ideally should be eliminated to remove its influence from the results. Therefore, spectral overlap between the signal and competing speech should be minimal. One approach to this problem is use of cochlear implant simulation processing on the speech stimuli. Cochlear implantees receive a spectrally degraded version of the speech stimulus through their implants. The processing strategy utilized in many implant processors filters speech into several adjacent frequency channels or bandpass filters, extracts the temporal amplitude envelope within each channel, and uses this envelope to modulate the series of electrical pulses sent to each channel. Therefore, the spectral information in speech is binned into a number of channels and is not fully represented. Simulations of cochlear implant pro-

cessing in normal-hearing listeners have revealed that speech recognition performance is very good (90% correct or better) with as few as 4–6 channels (Shannon *et al.*, 1995; Dorman *et al.*, 1997). However, in order to satisfy the requirement that spectral overlap between signal and masker be minimal, several modifications to the standard method of processing were implemented and are described in detail in the Methods section.

This study was designed to determine: (1) if informational masking can be isolated, at least partially, in speech stimuli and (2) if the size of the spatial release from masking is a function of the type of masking produced and more specifically, if the magnitude of release is different when informational, rather than energetic masking is the primary type of interference present.

II. METHODS

A. Listeners

Listeners were four normal-hearing college students aged 19–23. All had pure-tone thresholds less than or equal to 15 dB HL at octave frequencies from 250–8000 Hz bilaterally. One additional normal-hearing, college-age listener was employed for a brief study of the intelligibility of each of two noise maskers alone, but did not participate in the main study.

B. Stimuli

Speech materials were the CD-recorded sentences of the coordinate response measure (CRM) corpus (Bolia *et al.*, 2000). The corpus contains sentences spoken by eight different talkers—four male and four female. All sentences in the corpus are the same except for three words, which are the callsign (e.g., ''Baron,'' ''Charlie,'' ''Ringo''), a color, and a number. Each sentence has the following structure: ''Ready [*callsign*] go to [*color*] [*number*] now.'' Combinations of all eight callsigns, four colors (blue, red, green, and white), and eight numbers (1 through 8) result in 256 different sentences for each talker. The four male talkers of the corpus were used in this study.

The CRM corpus was chosen for several features that make it ideal for this study. One feature is the multiple-choice format that provides a somewhat easier task than open-set materials. The form of the sentences provides three important advantages: (1) linguistic context is not a factor; (2) the same sentence can be presented more than once without concern regarding listener memory for a specific sentence; and (3) the callsign can be used as a marker for the signal sentence. This corpus has been used in several recent studies of speech intelligibility, particularly in relation to multiple talkers, spatial separation, and energetic versus informational masking (Ericson and McKinley, 1997; Brungart, 2001a; Brungart and Simpson, 2001; Brungart *et al.*, 2001; Simpson *et al.*, 1999; Bolia *et al.*, 1999). Some of its basic characteristics and its relation to the articulation index have been reported by Brungart (2001b). Sentences are aligned at the onset of the word ''Ready,'' but speaking rates vary slightly from talker to talker and within talker. Differences in intelligibility of 5 to 20 percentage points have been reported across individual number, color, and talker (Brungart, 2001b, 2001a). However, in the current study these effects were evenly distributed across all conditions and test levels and therefore were not anticipated to have significant impact on the results, except to possibly increase the variability of each individual data point.

Signal sentences were drawn from those with the callsign ''Baron,'' of which 128 were available from the four male talkers. All 128 sentences were used 7 times each to create a pool of 896 processed signal sentences. Each signal sentence was paired with a randomly chosen masker sentence from among those with a different callsign, color, number, and talker than the signal sentence. A different color and number were required to allow for more robust analysis of error/confusion patterns between signal and masker sentences. If color and/or number were the same in the signal and masker, then a correct response could imply that the listener identified either the signal or the masker.

Each sentence was digitally preprocessed in order to produce cochlear implant simulated speech (Shannon *et al.*, 1995), with a total of 15 channels or bands. Software for this processing was provided by House Ear Institute and was modified to use pure-tone rather than narrow-band noise carriers and logarithmically spaced bandpass filters. Dorman *et al.* (1997) showed that speech recognition performance did not differ between narrow-band noise and pure-tone carriers. All sentences were down-sampled from 40 to 20 kHz prior to processing. The preprocessing algorithm included a high-pass Butterworth filter at 1200 Hz with a 6-dB/octave slope, followed by a bank of 15 bandpass filters. The bandpass filters were 1/3 octave, fourth-order Butterworth filters. The center frequencies of the filters were evenly spaced on a log scale from 215 to 4891 Hz with successive center frequencies at a ratio of 1.25. The output of each bandpass filter was half-wave rectified and low-pass filtered at 50 Hz by a fourth-order Butterworth filter to extract the amplitude envelope within each band. The envelope was then used to modulate a pure tone having a starting phase of zero and frequency equal to the center frequency of the corresponding band. The output of preprocessing was a set of 15 speech-envelope-modulated pure tones for each sentence.

The processed signal sentences were generated by summing a set of eight randomly chosen (for each sentence) bands of the 15 available from the signal sentences at the output of the preprocessing. Each signal sentence was paired with one masker sentence, but three different maskers were created from each masker sentence (as described below). Therefore, on each trial the signal sentence was played with one of the three maskers derived from the masker sentence to which it was paired.

1. Different-band sentence masker

The different-band sentence masker was generated by summing six bands of the masker sentence at the output of preprocessing. The six bands were randomly chosen (for each sentence) from the seven bands that were not chosen for the signal sentence with which it was paired. Therefore, the bands comprising a signal/different-band sentence masker pair were mutually exclusive. The top panels of Fig. 1 show

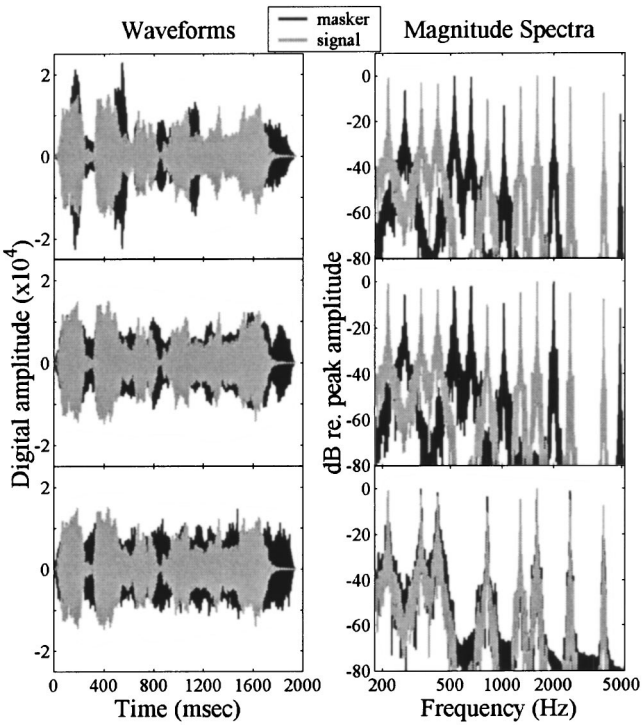


FIG. 1. Time waveforms (left column) and magnitude spectra (right column) for a sample signal sentence and the three types of maskers. The (same) signal sentence is plotted in gray in each panel. The black indicates the masker sentence for the different-band sentence masker (top panels), the different-band noise masker (middle panels), and the same-band noise masker (bottom panels).

the time waveforms and magnitude spectra of a sample pair of signal sentence (gray) and different-band sentence masker (black). The different-band sentence masker was designed to produce primarily informational masking. The magnitude spectra show that the signal and masker peaks did not overlap, minimizing spectral interaction and thus theoretically minimizing energetic masking. Informational masking was theoretically maximized because both signal and masker sentences were intelligible as speech and fundamental frequency information was minimal, making the talkers relatively difficult to discriminate from each other. The harmonic structure (and thus fundamental frequency) of voiced speech was essentially eliminated because the spectrum was represented by sparse, narrow-band, logarithmically spaced frequency bands.

2. Different-band noise masker

The different-band noise masker was derived from the different-band sentence masker described above. The long-term, complex spectrum of each different-band sentence masker was multiplied by the long-term, complex spectrum of a broadband Gaussian noise, and the result was inverse fast Fourier transformed (equivalent to convolving the different-band sentence with Gaussian noise). The length of the resultant waveform was truncated to equal the longer of either the different-band sentence masker it was derived from, or the signal sentence with which it was paired. A rise–fall time of 20 ms was imposed with a cosine^2 ramp. The spectra of this masker and the different-band sentence from which it was derived were nearly identical, therefore, they contained approximately the same (minimal) amount of energetic masking. Actual amounts of energetic masking likely varied slightly because the sentence masker contained more amplitude modulation than the different-band noise masker. The different-band noise had no intelligibility as speech and was qualitatively different from the signal sentence. Subjectively, this masker sounded like an amplitude-modulated multitone complex. Therefore, the proportion of informational masking was theoretically low relative to the different-band sentence masker, and energetic masking was approximately the same as for the different-band sentence masker. This masker was included as a control for energetic masking in the different-band sentence masker. The time waveforms and magnitude spectra of a sample different-band noise and signal sentence pair (same signal sentence as in the top panels) are shown in the middle panels of Fig. 1.

3. Same-band noise masker

The same-band noise masker was generated by summing eight bands of the preprocessed masker sentence. The eight bands were the same eight bands that comprised the signal sentence. The long-term, complex spectrum of the result was multiplied by the complex spectrum of a broadband Gaussian noise and inverse fast Fourier transformed (equivalent to convolving the eight-band masker sentence with Gaussian noise). The length of the waveform was truncated to equal the longer of either the processed signal sentence to which it was paired or the different-band masker sentence. A rise–fall time of 20 ms was added with a cosine^2 ramp. The time waveforms and magnitude spectra of a sample same-band noise and signal sentence pair (same signal sentence as in the top/middle panels) are shown in the lower panels of Fig. 1. The magnitude spectra almost completely overlapped, thus maximizing energetic masking. In addition, informational masking was theoretically minimized because the masker had no speech intelligibility and, as with the different-band noise, was not qualitatively similar to the signal sentence. This masker condition was included to contrast the effect of spatial separation for a predominantly informational (different-band sentence) and predominantly energetic (same-band noise) masker.

C. Stimulus presentation

All stimuli were presented in a sound field, located within a single-walled IAC sound booth with the typical metal, perforated surfaces except for the carpeted floor. Two Acoustic Research 215PS speakers were used to present the stimuli, one at 0° azimuth and the other at 90° azimuth to the right of the listener in the horizontal plane. The distance between the listener and the speakers was 5 ft. and the speaker height was approximately ear level for a seated listener. Stimuli were played through Tucker-Davis Technologies (TDT) hardware. The signal and masker were played out through separate TDT system channels. Stimuli were converted at a rate of 50 kHz via a 16-bit, 8-channel D/A converter (DA3-8), low-pass filtered with a 20-kHz cutoff (FT-6), and attenuated via programmable attenuators (PA-4). The

signal was sent through a switch (SS-1) to route it to the desired signal speaker. The signal and masker were sent to mixers where they were summed if they were both presented to the 0° speaker location. The stimuli were passed through power amplifiers (Tascam) and then to one or two speakers in the sound field, depending on the experimental condition. The masker level was fixed at 60 dB SPL.

Inverse filters were created for each speaker to compensate for the frequency response. The signal sentences were convolved with the inverse filter of the speaker at 0°. The maskers were convolved with the inverse filter of each speaker (0° and 90°), resulting in two versions of each masker, one appropriate to play from each speaker depending on the spatial separation condition. All processed stimuli were subsequently up-sampled to 50 kHz, scaled to equal rms amplitude, and saved to disk for later use. The difference in overall speaker output was equalized by adjusting the amplifier gain for each channel.

D. Procedures

Identification thresholds were estimated adaptively for the signal alone. Psychometric functions were measured for the signal-plus-masker conditions. Both signal-only and signal-plus-masker conditions were single interval forced-choice tasks with an identification response consisting of a color (four choices) and a number (eight choices). Each response was considered correct only if both the color and number of the signal sentence were identified accurately. Each trial proceeded as follows: warning, stimulus, response interval (color followed by number), and feedback. The listener indicated color response by use of four keys labeled with the written word in a background of the color assigned to it. The number response was by use of separate number keys (1–8). At the end of each block, the listener received threshold or percent-correct feedback for the block just completed.

The tracking rule used to measure signal-only identification thresholds caused a one-step decrease in signal level following a correct response, and a one-step increase in signal level following an incorrect response. This tracking rule targets the 50% correct performance level (Levitt, 1971). The initial 6-dB step size was reduced to 4 dB after the first three reversals. Threshold was taken as the average of a minimum of six reversals, not including the first four or five depending on the total number of reversals. Each block contained 36 trials. The signal sentence was played at the 0° location at 60 dB SPL. The threshold estimates from two blocks of trials were averaged to obtain signal-only identification threshold.

Signal-plus-masker psychometric functions were measured using four to eight signal levels, depending on the condition and the listener's early results. The levels were chosen individually to span percent-correct performance from near chance to near perfect. Signal levels were mixed randomly with replacement within every 50-trial block. Each of the three types of maskers was tested in two spatial separation conditions, 0° (signal and masker at 0°) and 90° (signal at 0° and masker at +90°). Each block of trials contained only one spatial separation condition and one masker type. Blocks of each spatial separation condition were alternated to disperse any learning effects equally across the two separation conditions. A minimum of 100 trials per signal level was obtained for each spatial separation and masker type condition.

Two blocks of signal-only identification trials were collected first. For three of the four listeners, data for the different-band sentence masker were collected next, followed by the different-band and same-band noise maskers intermixed within the same session. The fourth listener generated data with each of the three maskers mixed within all the sessions. Data were collected within six sessions of 2 h each, including several rest breaks. The listeners' heads were not physically restrained, but they were instructed to keep their heads straight, facing the 0° location throughout stimulus presentation.

Additional data were briefly collected to assess the intelligibility of each masker in isolation. The listener's task was to identify the color and number of the masker sentence. One of the four listeners was tested with one 50-trial block of the different-band sentence masker at 60 dB SPL from the 0° location. One additional normal-hearing listener (not one of the original four) performed the identification task for four 60-trial blocks, two for the different-band noise masker and two for the same-band noise masker, presented at 60 dB SPL at the 0° location.

Listener training was minimal. Each listener initially received a 50-trial block of identification practice. The signal sentences were presented without a masker at a fixed level of 60 dB SPL from the 0° speaker. If performance was at least 90% correct, the listener was given four 50-trial practice blocks with the signal presented in the different-band sentence masker, two blocks for each separation condition. One fixed signal level was used in each block. The signal-to-masker ratio was +15 dB for the first block, but was decreased to +10 dB for the second practice block. These data were not included in any analyses or plots.

E. Data analyses

Psychometric functions were constructed showing percent-correct identifications as a function of signal level. The psychometric functions were fit by the Nelder–Mead simplex method (FMINS function in MATLAB) with a logistic of the form

$$p(c) = \alpha + ((1.0 - \alpha)/(1.0 + e^{-k(x-m)})), \quad (1)$$

where $\alpha = 0.03$ (chance performance for 1 out of 32), k is the slope of the function, m is the level of the midpoint (about 51.5% correct), and x is the signal sentence level.

All statistical analyses were performed with repeated-measures analysis of variance (ANOVA) with listener as the error term. *Post hoc* testing was performed with Tukey's Studentized Range (HSD) Test, which controls for type I errors, at the $\alpha = 0.05$ level of significance. Two difference measures were calculated: Δm and Δk. These values were calculated from the midpoint and slope parameters of the fitted logistic functions described above, and represent the difference in spatial separation condition (0°–90°), not masker type.

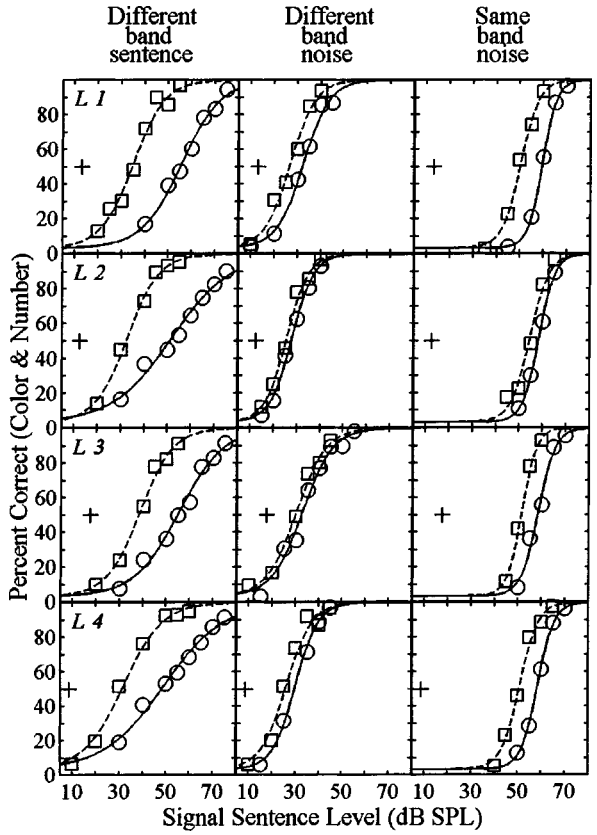


FIG. 2. Psychometric functions for each listener (rows L1 to L4) and each masker type (columns). The plus symbol marks signal-only adaptive identification threshold for 50% correct (plotted three times for each listener as a reference). The circles are the data points and the solid line is the logistic fit for 0° separation, and the squares are the data points and the dashed line is the logistic fit for 90° separation.

III. RESULTS

The signal sentences by themselves (without a masker) were very intelligible at a normal conversational level (60 dB SPL). All listeners achieved 96% to 98% correct for the initial 50-trial practice block, and all reported that the task was easy to perform. Signal-only adaptive identification thresholds ranged from 8.7 to 17.7 dB SPL. The intelligibility of the different-band sentence masker by itself was also very high (98% correct), while the different-band and same-band noise maskers were each identified near chance (3%)—4.2% for the different-band noise and 3.3% for the same-band noise.

The signal-plus-masker data and fitted logistic functions are plotted in Fig. 2. Each panel contains the data for one listener and one masker type. Percent correct for identifying both color and number correctly in the same trial is plotted as a function of signal sentence level. Zero dB signal-to-masker ratio corresponds to a signal level of 60 dB SPL. Each row displays the data for one listener (denoted L1 to L4). The different-band sentence masker results are located in the first column, the different-band noise in the middle column, and the same-band noise in the last column. The circles represent 0° and squares represent 90° separation. The plus symbol indicates the signal-only 50%-correct mean threshold from the adaptive procedure and is plotted three times as a reference, once for each masker type and listener. The fitted logistic functions are indicated by the solid line for 0° and by the dashed line for 90°.

The first important feature of the data is the consistent monotonicity of the psychometric functions across the four listeners and three maskers used in the experiment. In all three masker types, the function for 90° lies to the left of that for 0° along the abscissa, and more importantly, the magnitude of the shift changes as a function of masking condition. The location and slope of the psychometric functions also change with masker type independent of spatial separation. Table I contains m (midpoint) and k (slope) values and an estimate of the goodness of fit (proportion of variance accounted for) from the fitted psychometric functions, as well as the difference measures Δm and Δk, for each listener and condition. The midpoint parameter was used as a summary measure of the effect each masker had on performance and for comparison between spatial conditions. The midpoint was chosen because it is not influenced by ceiling or floor effects, and is readily available from the logistic fits. However, it should be kept in mind that the midpoint may not fully describe the effects, particularly when the function slopes are not parallel, in which case the size of the effect depends on the performance level.

Figure 3 plots the mean m (top panel) and k (bottom panel) values from Table I as a function of masker type with spatial separation as the parameter. The circles are for 0° and the squares for 90° separation. The dashed line at the bottom of the top panel indicates the mean signal-only identification threshold in dB SPL.

A. Effect of masker type

For 0° spatial separation, the mean m value was highest for the same-band noise at 58.7 dB, followed closely by 52.9 dB for the different-band sentence, while the different-band noise resulted in a lower value of 30.8 dB. An ANOVA confirmed a significant main effect of masker type [$F(2,6) = 311.6$, $p < 0.0001$], and *post hoc* testing revealed that the same-band noise value was significantly higher than the different-band sentence masker, and both were higher than the different-band noise midpoint value. Midpoint values in the 90° separation condition followed a similar pattern with a mean of 51.8 dB for the same-band noise, 34.5 dB for the different-band sentence, and 27.2 dB for the different-band noise. The ANOVA on the 90° midpoint values revealed a significant main effect of masker [$F(2,6) = 256.7$, $p < 0.0001$] and *post hoc* testing showed that each midpoint value was significantly different from the others.

The mean slope values for both the 0° and 90° separation conditions increased from approximately 0.1 for the different-band sentence masker to nearly 0.3 for the same-band noise, with the different-band noise values intermediate. An ANOVA on the slope values revealed a significant main effect of masker type in both the 0° [$F(2,6) = 67.03$, $p < 0.0001$] and 90° conditions [$F(2,6) = 39.03$, $p < 0.001$]. *Post hoc* testing revealed that in the 0° condition, all three slopes were significantly different from each other. In the 90° condition, the slope for the same-band noise was significantly steeper than for the other two maskers, but the

TABLE I. Results of the logistic fits applied to the data for each masker (subtables A, B, and C). The *m*- and *k*-values and amount of variance accounted for (var) by the fit (in percent) for each listener, as well as means and standard deviations across listener are displayed in each subtable for each spatial separation condition. The final two columns of each table give the Δm and Δk difference values between the 0° and 90° separation conditions, in addition to the means and standard deviations across listener.

	0° *separation*			90° *separation*			0°–90°	
Listener	*m*	*k*	Var	*m*	*k*	Var	Δm	Δk
				(A) Different-band sentence masker				
1	55.58	0.1195	98.9	34.91	0.1462	98.0	20.67	−0.027
2	51.72	0.0799	97.9	32.70	0.1487	99.4	19.02	−0.069
3	55.31	0.1064	98.6	38.54	0.1436	99.3	16.77	−0.037
4	48.98	0.0784	98.4	31.71	0.1203	99.3	17.27	−0.042
Average	**52.90**	**0.0961**	···	**34.47**	**0.1397**	···	**18.43**	**−0.044**
(s.d.)	**(3.15)**	**(0.02)**		**(3.03)**	**(0.01)**		**(1.78)**	**(0.02)**
				(B) Different-band noise masker				
1	32.42	0.1772	99.3	26.83	0.1849	98.5	5.59	−0.008
2	27.93	0.2150	99.4	25.68	0.2151	99.3	2.25	−0.0001
3	32.81	0.1550	98.3	30.24	0.1586	99.2	2.57	−0.004
4	30.01	0.2044	99.7	26.12	0.1824	97.3	3.89	0.022
Average	**30.79**	**0.188**	···	**27.22**	**0.185**	···	**3.58**	**0.003**
(s.d.)	**(2.28)**	**(0.03)**		**(2.07)**	**(0.02)**		**(1.52)**	**(0.01)**
				(C) Same-band noise masker				
1	59.53	0.3279	100	50.30	0.2560	99.3	9.23	0.072
2	58.35	0.2967	99.8	54.53	0.2479	97.8	3.82	0.049
3	58.49	0.2777	98.3	51.52	0.3223	99.7	6.97	−0.045
4	58.39	0.2839	99.9	50.82	0.2533	99.3	7.57	0.031
Average	**58.69**	**0.2966**	···	**51.79**	**0.2699**	···	**6.90**	**0.027**
(s.d.)	**(0.56)**	**(0.02)**		**(1.90)**	**(0.04)**		**(2.26)**	**(0.05)**

different-band sentence slope at 0.14 was not different from the 0.19 slope for the different-band noise.

B. Effect of spatial separation

For the different-band sentence masker, the mean difference in *m* value (0°–90°), or Δm, is much larger at 18.4 dB than for the other maskers. The different-band noise masker produced a mean Δm of 3.6 dB and the same-band noise masker resulted in a slightly larger Δm of 6.9 dB. All three mean Δm's are statistically larger than zero on paired-sample *t*-tests [$t(3)=20.8$, $p<0.001$ for different-band sentence; $t(3)=4.7$, $p<0.05$ for different-band noise; and $t(3)=6.1$, $p<0.01$ for same-band noise]. An ANOVA revealed a significant main effect of masker type on Δm [$F(2,6)=143.1$, $p<0.001$]. *Post hoc* testing revealed that the mean Δm for each masker was significantly different from all the other maskers.

The effect of masker type on the difference in slope, Δk (0°–90°), was less obvious because it was very small. The different-band sentence masker Δk was consistently negative across listener and had a mean of about −0.04 (the 0° functions were shallower than the 90° functions). This slope change was significantly larger than zero on a paired-sample *t*-test [$t(3)=-4.9$, $p<0.05$]. The average Δk for the different-band noise masker was essentially zero [$t(3)=0.4$, $p=0.716$] and for the same-band noise masker it was less than 0.03 and not significant [$t(3)=1.1$, $p=0.368$], both on paired-sample *t*-tests. An ANOVA on the Δk's with masker type as the factor approached significance [$F(2,6)=4.85$, $p=0.056$].

C. Error analysis

Examination of the identification errors revealed some striking patterns across masker and spatial separation conditions. Figure 4 plots the mean percentage of incorrect responses that matched the color or number from the masker sentence as a function of condition.

Percentages were calculated separately for color and number by dividing the number of responses corresponding to the masker, by the total number of incorrect responses for a limited subset of signal levels. In order to obtain a clearer picture of error patterns, trials presented with signal levels significantly above or below the level required for 50% correct were eliminated. Few errors would occur at levels significantly above 50% correct. For signal levels significantly below that for 50% correct in the different-band sentence masker, the listener clearly heard the masker sentence but the signal sentence was barely intelligible, if at all. Listeners were aware that the signal sentence could not contain the same color and number as the masker sentence, and they reported that at very low signal-to-masker ratios they purposely chose a random color and number other than that heard clearly from the masker sentence. Therefore, it was primarily the signal-to-masker ratios near that required for 50% correct performance that were of most interest for examination of confusions between signal and masker. For each masker/spatial separation condition, trials presented with signal levels equal to or within $\pm j$ dB of the midpoint of the fitted psychometric function, *m*, were included in the error calculations, with $j=10$ for the different-band sentence

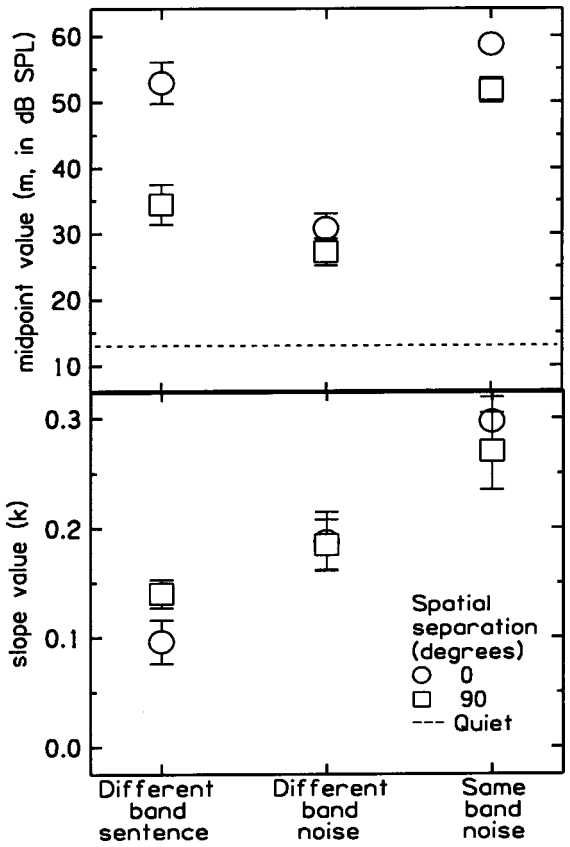


FIG. 3. Midpoint (top panel) and slope (lower panel) values from the fitted logistic curves as a function of masker type, with spatial separation as the parameter. Data are plotted as the means across four listeners and the errors bars are ±1 standard deviation. The circles represent 0° separation and the squares are for 90° separation. The dashed line at the bottom of the top panel indicates the mean signal-only sentence identification threshold.

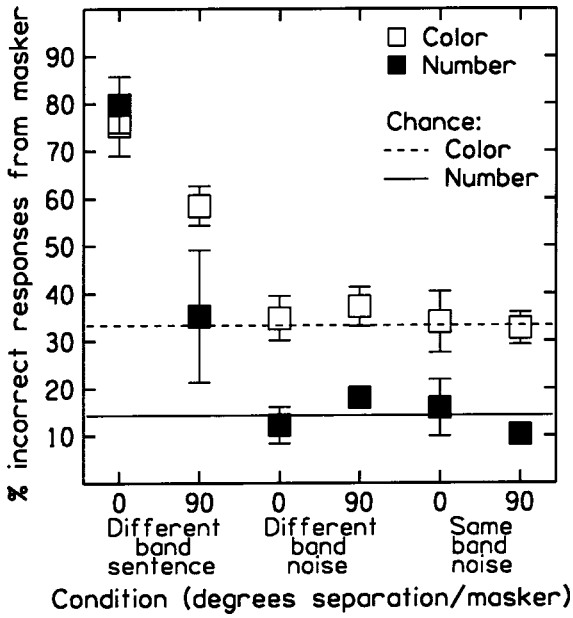


FIG. 4. Percent of incorrect color and number responses that matched color and number in the masker sentence. Results are for a limited range of signal sentence levels near the 50%-correct point on the psychometric functions (see the text for more detailed explanation), and are plotted as a function of spatial separation and masker type. The data are means across the four listeners and the errors bars are ±1 standard deviation. The open squares represent color response errors and the closed squares are number response errors. Chance is indicated for color by the dashed line (33.3%) and for number by the solid line (14.3%).

masker and $j=5$ for the two noise maskers due to the steeper functions.

Each point in Fig. 4 represents the mean of the four listeners and includes at least 360 incorrect trials. The solid squares are number errors and the solid line indicates expected performance for random guessing or a random distribution of errors for number (14.3%). The open squares are for color errors and the dashed line represents expected performance for a random distribution of color errors (33.3%). Errors for the different-band sentence masker were greater than that expected for randomly distributed responses. For 0° separation, 75% (color) to 80% (number) of errors were from the masker sentence. In the 90° configuration, these numbers dropped to 35% for number and 58% for color, but both remained above their respective random distribution levels. ANOVAs were performed separately for color and number errors in the different-band sentence masker, with spatial separation as the factor. The results confirmed a significant main effect of spatial separation on both color [$F(1,3)=163.0$, $p<0.001$] and number errors [$F(1,3)=59.72$, $p<0.01$]. In the remaining masker/spatial separation conditions, percentages approximated that expected for a random distribution of responses.

IV. DISCUSSION

The psychometric functions obtained in all three maskers were monotonic. Psychometric functions for a speech signal masked by a single talker have been known to contain nonmonotonicities (plateaus/small dips between −10- and 0-dB signal-to-masker ratio), particularly when the characteristics of the two talkers are similar (Egan *et al.*, 1954; Dirks and Bower, 1969; Brungart, 2001a; Freyman *et al.*, 1999; Brungart *et al.*, 2001). The presumed origin of these nonmonotonicities is segregation by level. It is difficult to segregate two similar talkers when the signal-to-masker ratio is near 0 dB. However, when the signal level drops just below that of the masker, the listener can segregate by level and extract or assign information from/to the softer talker. There is no evidence in the current data to suggest nonmonotonic functions. The level cue is probably a weak cue used only when all else fails to segregate concurrent talkers. Here, the stronger cue may have been timbre/pitch differences between the two talkers arising from the particular set of bands chosen during processing for the signal and masker. However, the timbre/pitch cue varied from trial to trial because the bands were picked randomly on each trial. Therefore, although this cue was possibly more salient than level differences, it was probably not stable enough to be a very strong cue for segregating the signal and masker sentences.

A. Masker type—Relative proportions of informational and energetic masking

The influence of masker type, independent of spatial separation, is interesting to consider with a few caveats. The different-band sentence masker had greater amplitude modulations than the two noise maskers. The larger amplitude peaks in the different-band sentence masker may have contributed to greater nonsimultaneous masking than either of

the noise maskers. However, the fact that, in general, amplitude-modulated noise actually results in less masking than an equal-rms flat-amplitude noise (e.g., Festen and Plomp, 1990) suggests that forward and backward masking do not produce more total energetic masking in an amplitude-modulated noise than in a flat-amplitude noise. Based on this argument, the different-band sentence masker possibly contained slightly less energetic masking than the different-band noise. In addition, the fact that speech is dominated by slower envelope fluctuations (and therefore longer dips) relative to a noise, also suggests that the different-band sentence masker produced slightly less energetic masking than the different-band noise masker. Therefore, the different-band noise masker was a conservative control for energetic masking in the different-band sentence masker.

A second caveat is that the midpoints of the psychometric functions were used for comparison between maskers. Due to the different slopes found across masker type, the size of the difference between maskers will change depending on the performance level one looks at. However, the main comparison made below is between the different-band sentence masker and the different-band noise. The psychometric functions in these maskers were significantly separated, in the same direction, at all performance levels above chance, but it should be noted that the difference between these maskers actually increased with performance level.

With full knowledge of the above, the effect of masker type on the 0° separation psychometric functions leads to interpretations regarding the type of masking that each causes. Of particular interest is the difference in midpoint values between the different-band sentence and different-band noise masker functions—about 22 dB. The different-band noise was intended to provide an approximate measure of how effectively the construction of both different-band maskers (sentence and noise) minimized energetic masking. These two maskers contained energy that was spectrally quasi-interleaved with the signal energy to the same extent. Therefore, they should have produced about the same amount of energetic masking based on their spectra, but the design of the stimuli emphasized informational masking in the different-band sentence masker and minimized informational masking in the different-band noise masker. The different-band noise resulted in a mean of 18 dB of (primarily) energetic masking, and the different-band sentence masker likely resulted in about the same amount of energetic masking. However, the different-band sentence masker produced a mean of 40-dB overall masking. This difference is too large to be accounted for by the temporal differences between the different-band sentence and different-band noise maskers. Therefore, some other type of masking contributed to the total amount of interference present in the different-band sentence masker.

Three factors suggest that this other type of masking was informational: (1) the design of the stimuli; (2) the psychometric function slopes; and (3) the error analysis. The first factor was discussed in the Methods section. The different-band sentence masker was designed to maximize informational masking of the signal sentence. The processing of the stimuli resulted in nonoverlapping spectral peaks that reduced energetic masking while preserving the intelligibility and qualitative similarity of the two sentences. The second factor, psychometric function slopes, provides further evidence of masking type. In studies comparing energetic and informational masking, psychometric functions in noise maskers tend to be fairly steep, while informational maskers usually result in relatively shallow psychometric functions (Kidd *et al.*, 1995, 1998, 2002). Psychometric functions for speech masked by a single talker are shallower than those masked by a noise (Dirks and Bower, 1969; Festen and Plomp, 1990, Freyman *et al.*, 1999; Wilson *et al.*, 1990), which might be explained by the larger proportion of informational masking present when speech is masked by a speech masker than a noise masker (Freyman *et al.*, 1999, 2001; Brungart, 2001a; Brungart *et al.*, 2001). At 0° separation, the different-band sentence masker functions had a mean slope of 0.10 (2.3% per dB[1]), the different-band noise 0.19 (4.5% per dB) and the same-band noise 0.30 (7.1% per dB); this is a factor of 2 increase in slope from the different-band sentence to the different-band noise, and a factor of 3 increase from the different-band sentence to the same-band noise.

The reason for shallower psychometric functions in the presence of an effective informational masker is not entirely clear, although several possibilities seem reasonable to consider. Psychometric functions obtained for a predominantly energetic masker are inherently strongly dependent upon signal-to-masker ratio. However, performance in a predominantly informational masker may depend much less on signal-to-masker ratio because level difference between signal and masker is likely a weak segregation cue. Other factors such as signal-masker similarity and/or auditory scene analysis cues (other than level) might have more influence in this circumstance.

Psychometric function slopes probably cannot be used in isolation to determine masking type (informational/energetic). Shallow slopes may also be accounted for by inattention or inappropriately focused attention on some proportion of trials (e.g., Allen and Wightman, 1994). Performance at one signal-to-masker ratio might result from an average of trials for which attention was appropriately focused and trials for which it was not. For the different-band sentence masker it is likely that attention was incorrectly focused on the masker sentence during some proportion of the trials (this is supported by the error analysis discussed below). This is less likely to have occurred for the noise maskers because they were perceptually dissimilar from the signal. However, a similar averaging can be imagined without invoking inattention. When maskers or signals are not homogeneous in terms of effectiveness or recognition, the psychometric function produced by a mixture of these stimuli is shallower than the individual stimulus psychometric functions because the overall function is an average of the individual functions (e.g., Wilson and Carter, 2001). Some combinations of signal and different-band sentence masker were more easily segregated than others (due to spectral and/or temporal differences). For the different-band noise masker, variations in masker effectiveness arose be-

cause of differences in the spread of excitation for various combinations of frequency bands. The slopes of the psychometric functions for the different-band sentence masker may be shallower because they were influenced by all these factors (weak dependence on level, misfocused attention, and stimulus nonhomogeneity), whereas the different-band noise masker may have only been susceptible to the influences of nonhomogeneity. Other factors including the design of the stimulus, the error analysis (see below), and the size of the spatial release from masking (e.g., Freyman *et al.*, 1999; Kidd *et al.*, 1998) imply that informational masking was dominant for the different-band sentence masker and energetic masking was dominant for the different-band noise (although less energetic than the same-band noise masker).

The error analysis is the third factor that supports a high informational to energetic masking ratio in the different-band sentence masker. This analysis is similar to that used by Brungart (2001a) in which it was argued that if incorrect responses were due to primarily energetic masking (i.e., the signal was inaudible), the distribution of responses on incorrect trials should be random among the remaining alternatives. When the signal was inaudible, there was no particular reason (and in fact it would have been detrimental) to pick the masker color/number combination because it clearly belonged to the masker. In addition, in the current study, the listeners were aware that a masker–signal pair always contained a different color–number set. Therefore, a significant proportion of masker responses indicates difficulty correctly labeling each sentence as ''signal'' or ''masker,'' which implies that the listener confused the masker and signal and not that the masker necessarily degraded the representation of the signal in the auditory system. Brungart (2001a) reported that the percentage of errors from the masker sentence was significantly larger than expected from random error distribution for a CRM sentence masked by another CRM sentence (both unprocessed), particularly when the masker was the same talker or a same-sex talker. In the current study, listeners responded with the masker color and number no more often than was expected by chance (33.3% for color; 14.3% for number) for both of the noise maskers. In the different-band sentence masker, the proportion of responses that corresponded to the masker was much larger at 75%–80% for the 0° separation and in the 90° separation condition, 35% for number and 58% for color. This suggests that when the signal-to-masker ratio was near the 50%-correct point, the listeners often confused the signal and masker, implying the presence of informational masking.

B. Spatial separation

The different-band sentence masker resulted in the largest spatial release from masking at 18.4 dB. The different-band noise release was 3.6 dB and the same-band noise release was 6.9 dB. The nearly 7-dB release obtained for the same-band noise is consistent with that expected for 90° separation of speech and noise, based on binaural interaction and headshadow cues (Zurek, 1993), and is similar to that reported for 90° separation in other studies of speech masked by noise (e.g., Bronkhorst and Plomp, 1990; Duquesnoy, 1983; Peissig and Kollmeier, 1997; Hawley, 2000; Bronkhorst and Plomp, 1992; Nilsson *et al.*, 1992). The smaller release that occurred for the different-band noise was presumably due primarily to a release (via binaural interaction and headshadow) from the small amount of energetic masking that was present. It was smaller than the release for the same-band noise because the effective level of the noise was very low within auditory filters containing the signal, and binaural interaction is less effective at low noise levels (e.g., Hirsh, 1948; Dolan and Robinson, 1967; Dolan, 1968; McFadden, 1968). An alternative measure of spatial release is the difference in identification performance at one signal level. When calculated with the midpoint of the 0° function as the reference, the spatial advantage is 17, 37, and 43 percentage points for the different-band noise, same-band noise, and different-band sentence, respectively. The difference in spatial advantage between the different-band sentence and same-band noise becomes less striking, but the overall trend remains the same. Spatial release could have also been calculated at a different performance level, in which case the exact magnitude of the advantage in each masker would change, but again, the same trend would be apparent.

The 18.4-dB release obtained in the different-band sentence masker is much larger than can be explained by both headshadow and binaural interaction cues. The size of the release is also a bit larger than found in much of the previous speech-on-speech masking research. Freyman *et al.* (1999) found a 14-dB advantage in an anechoic chamber for a spatial separation of 60° between the signal sentence and a single, same-sex speech masker. In real and virtual rooms, Hawley (2000) found spatial advantages between 5 and 11 dB for one to two same-sex maskers separated from the signal by 60° to 90°. However, the content of the masker sentence(s) was held constant and was well-known to the listener. Therefore, stimulus uncertainty was not high and the informational portion of masking was probably relatively low. Both Duquesnoy (1983) and Peissig and Kollmeier (1997) found spatial advantages between 5 and 7 dB for a 90° separation between signal and speech masker. However, both studies employed stimuli or procedures that likely reduced stimulus uncertainty or signal–masker similarity, therefore limiting informational masking (Duquesnoy used a different-sex masker, and Peissig and Kollmeier used only two signal sentences in continuous discourse masking speech). Comparisons to the current study are also complicated by the fact that the studies above tested some unknown combination of energetic and informational masking of speech. The 18-dB advantage found here was also in the presence of both energetic and informational masking, but the ratio of informational to energetic masking was probably substantially larger than in the above studies. Ericson and McKinley (1997) found large advantages in a same-sex masker that are comparable or larger than those found for the predominantly informational masker in the current study. However, as discussed in the Introduction, direct comparison to the current results is difficult due to significant methodological and procedural differences. Nonetheless, the functions were striking in terms of the size of the spatial advantage and perhaps provide some further support for the existence of substantial advantages due to perceptual factors.

The large spatial release found in the current study may be a consequence of the greater isolation of informational masking due to signal–masker similarity and processing of the signal and masker sentences into mutually exclusive frequency bands. Informational masking may be subject to a significantly large spatial release as a result of perceptual segregation of the two sources, as suggested by Freyman *et al.* (1999) and Freyman *et al.* (2001). They demonstrated that the presence of a perceptual separation between signal and masker sentences produced a 4–10-dB advantage for 60° of separation. This may seem small compared to the current study, but it did not include binaural interaction and headshadow cues, and was the result of an informational to energetic masking ratio which was probably smaller than that created here. The 18-dB advantage found in this study is comparable to the advantages found by Kidd *et al.* (1998) using nonspeech signal patterns in a multitone informational masker. Informational masking was maximized relative to energetic masking, similar to the current study.

The effect here is believed to be mainly due to perceptual factors, not acoustic. The small contribution due to acoustic factors can be estimated by the 3.6-dB advantage found for the different-band noise masker; the remainder of the effect is presumably due to perceptual segregation by spatial location. When the signal is played with a predominantly informational masker, this perceptual effect can improve signal recognition. For 0° separation, the necessary information about the signal is present, it is audible, but the listener cannot separate it from the informational portion of the masker because there are not many salient differences between the signal and masker. When the masker is moved away from the signal, the perceptual effect of spatial separation enables the listener to disentangle the signal and masker, or to correctly assign ''signal'' and ''masker'' to the appropriate sentences. The perception of the signal and masker at two separate locations effectively reduces the similarity and confusion between signal and masker. A primarily energetic masker does not benefit as much from this same effect because it is mostly affected by acoustic, not perceptual factors. The signal and masker sentences were perceptually segregated spatially, but probably by other factors as well, such as small differences in speaking rate, continuity of envelope variations over time, and timbre/pitch differences created by using mutually exclusive frequency bands within each signal–masker pair. These additional factors were also present in the 0° separation condition, but the spatial separation may have helped make these additional cues more salient.

Masker type influenced the change in psychometric function slope with spatial separation. For the different-band and same-band noises, the slopes of the functions did not change significantly with spatial separation. However, for the different-band sentence masker, the slope was steeper for the 90° function (0.14 or 3.3% per dB^{1}) than for the 0° function (0.10 or 2.3% per dB); a factor of 1.4 steeper. The difference in slope between the two functions resulted in diverging functions as signal level increased and therefore, the advantage of spatial separation increased (below 90% correct) with signal level. This result is not consistent with the masking level difference literature in which the psychometric function slopes are parallel or converge for N_oS_o and N_oS_{II} for suprathreshold tasks such as loudness matching, frequency discrimination, or intensity discrimination (see Yost, 1997 for a review) and for nonspeech pattern identification (Kidd *et al.*, 1995). This difference is important because it illustrates the difference in mechanisms causing the release from masking. The measurement of masking level differences is dominated by energetic masking and the release from masking occurs due to a decrease in energetic masking as a result of binaural interaction. In contrast, the different-band sentence is dominated by informational masking for the 0° separation condition, and the release from masking for the 90° condition occurs due to a decrease in informational masking as a result of perceptual factors. The masking therefore becomes proportionally more energetic, resulting in a steeper psychometric function slope. The release from informational masking causes the divergence of the functions.

An interesting comparison is across both masker type and spatial separation at the same time, specifically the different-band noise masker functions (either the 0° or 90°) to the different-band sentence at 90°. The 90° separation between the signal and different-band sentence masker shifted the midpoint close to that for the different-band noise functions. The midpoint difference between the different-band sentence and different-band noise maskers shrank from 22 dB at 0° to 4–7 dB at 90°, and the slope of the different-band sentence masker function increased to 0.14 at 90° separation, toward the value of 0.19 for the different-band noise functions. Although the function for the different-band sentence masker at 90° was distinguishable from the different-band noise functions, a substantial amount of the informational masking active in the 0° separation condition was apparently eliminated by moving the different-band sentence masker 90° to the right.

The results of the error analysis for the different-band sentence masker were also affected by spatial separation. Listeners confused the signal and different-band sentence masker more often when they originated from the same location (75%–80% of errors) than when the masker was moved 90° to the right (35%–58% of errors). This suggests, along with the psychometric function slopes and magnitude of the spatial advantage, that the spatial separation between signal and masker reduced informational masking in the different-band sentence masker.

V. CONCLUSIONS

(1) The results suggest that the ratio of informational to energetic masking can be maximized in speech stimuli using a modification of cochlear implant simulation processing. This allows for the examination of informational masking in the cocktail party problem more directly by reducing the potentially confounding effects of energetic masking.

(2) For 0° spatial separation, the masking caused by the different-band sentence masker was primarily informational, as evidenced by: (a) approximately 22 dB more masking than for the different-band noise masker con-

trol; (b) shallow psychometric function slopes; and (c) high percentage of confusions with the masker sentence.

(3) A robust spatial separation effect was obtained with all types of maskers tested, but the size of the spatial release from masking differed significantly with masker type. The different-band sentence masker (mainly informational) produced the largest spatial advantage (18 dB). The same-band noise (mainly energetic) produced a significantly smaller release (7 dB), in line with previous research of speech masked by noise. The different-band noise (minimally energetic and informational) produced the smallest spatial release (4 dB).

(4) A relatively large ratio of informational to energetic masking (due to signal–masker similarity and/or stimulus uncertainty) is likely the key to producing such a large spatial release from masking.

(5) The 90° spatial separation condition reduced informational masking caused by the different-band sentence masker, as evidenced by: (a) 14 dB more spatial release from masking than for the different-band noise masker control; (b) steeper psychometric function slopes for the 90° separation condition than for 0° separation; and (c) significant reduction in the percentage of confusions with the masker sentence when the masker was moved to 90°.

(6) The results of this study provide support for the notion that spatial release from informational masking and energetic masking are produced by different mechanisms. The spatial release from an informational masker is mainly produced by a perceptual effect, whereas the spatial release from an energetic masker is primarily mediated by binaural interaction and headshadow.

ACKNOWLEDGMENTS

This work was supported by NIH/NIDCD Grant Nos. DC04545, DC04663, and DC0100, and by the Boston University Hearing Research Center. The authors are grateful to Andrew Brughera for critical technical assistance, Robert Shannon and Qian-Jie Fu of House Ear Institute for providing the cochlear implant simulation algorithm, and Richard Freyman for early discussions regarding this project. We are also grateful to Douglas Brungart and an anonymous reviewer for providing helpful comments on an earlier draft of the paper.

[1]Slope in percent per dB is from a straight-line fit between 40% and 60% correct performance.

Allen, P., and Wightman, F. (**1994**). ''Psychometric functions for children's detection of tones in noise,'' J. Speech Hear. Res. **37**, 205–215.

Bolia, R. S., Ericson, M. A., Nelson, W. T., McKinley, R. L., and Simpson, B. D. (**1999**). ''A cocktail party effect in the median plane?'' J. Acoust. Soc. Am. **105**, 1390–1391.

Bolia, R. S., Nelson, W. T., Ericson, M. A., and Simpson, B. D. (**2000**). ''A speech corpus for multitalker communications research,'' J. Acoust. Soc. Am. **107**, 1065–1066.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (M.I.T. Press, Cambridge, MA).

Bronkhorst, A. W., and Plomp, R. (**1990**). ''A clinical test for the assessment of binaural speech perception in noise,'' Audiology **29**, 275–285.

Bronkhorst, A. W., and Plomp, R. (**1992**). ''Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing,'' J. Acoust. Soc. Am. **92**, 3132–3139.

Brungart, D. S. (**2001a**). ''Informational and energetic masking effects in the perception of two simultaneous talkers,'' J. Acoust. Soc. Am. **109**, 1101–1109.

Brungart, D. S. (**2001b**). ''Evaluation of speech intelligibility with the coordinate response measure,'' J. Acoust. Soc. Am. **109**, 2276–2279.

Brungart, D. S., and Simpson, B. D. (**2001**). ''Contralateral masking effects in dichotic listening with two competing talkers in the target ear,'' J. Acoust. Soc. Am. **109**, 2486.

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.

Cherry, E. C. (**1953**). ''Some experiments on the recognition of speech with one and two ears,'' J. Acoust. Soc. Am. **25**, 975–979.

Dirks, D. D., and Bower, D. R. (**1969**). ''Masking effects of speech competing messages,'' J. Speech Hear. Res. **12**, 229–245.

Dolan, T. R. (**1968**). ''Effect of masker spectrum level on masking-level differences at low signal frequencies,'' J. Acoust. Soc. Am. **44**, 1507–1512.

Dolan, T. R., and Robinson, D. E. (**1967**). ''An explanation of masking level differences that result from interaural intensive disparities of noise,'' J. Acoust. Soc. Am. **42**, 977–981.

Dorman, M. F., Loizou, P. C., and Rainey, D. (**1997**). ''Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs,'' J. Acoust. Soc. Am. **102**, 2403–2411.

Duquesnoy, A. J. (**1983**). ''Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons,'' J. Acoust. Soc. Am. **74**, 739–743.

Egan, J. P., Carterette, E. C., and Thwing, E. J. (**1954**). ''Some factors affecting multi-channel listening,'' J. Acoust. Soc. Am. **26**, 774–782.

Ericson, M. A., and McKinley, R. L. (**1997**). ''The intelligibility of multiple talkers separated spatially in noise,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Erlbaum, Mahwah, NJ), pp. 701–724.

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.

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.

Gelfand, S. A., Ross, L., and Miller, S. (**1988**). ''Sentence reception in noise from one versus two sources: Effects of aging and hearing loss,'' J. Acoust. Soc. Am. **83**, 248–256.

Hawley, M. L. (**2000**). ''Speech intelligibility, localization and binaural hearing in listeners with normal and impaired hearing,'' Ph.D. dissertation, Biomedical Engineering, Boston University, Boston, MA.

Hirsh, I. J. (**1948**). ''The influence of interaural phase on interaural summation and inhibition,'' J. Acoust. Soc. Am. **20**, 536–544.

Kidd, G., Jr., 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, G., Jr., Mason, C. R., and Rohtla, T. L. (**1995**). ''Binaural advantage for sound pattern identification,'' J. Acoust. Soc. Am. **98**, 1977–1986.

Kidd, G., Jr., 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.

Leek, M. R., Brown, M. E., and Dorman, M. F. (**1991**). ''Informational masking and auditory attention,'' Percept. Psychophys. **50**, 205–214.

Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Lutfi, R. A. (**1989**). ''Informational processing of complex sound. I. Intensity discrimination,'' J. Acoust. Soc. Am. **86**, 934–944.

McFadden, D. (**1968**). ''Masking-level differences determined with and without interaural disparities in masker intensity,'' J. Acoust. Soc. Am. **44**, 212–223.

Neff, D. L. (**1995**). ''Signal properties that reduce masking by simultaneous, random-frequency maskers,'' J. Acoust. Soc. Am. **98**, 1909–1920.

Neff, D. L., and Green, D. M. (**1987**). ''Masking produced by spectral uncertainty with multi-component maskers,'' Percept. Psychophys. **41**, 409–415.

Nilsson, M., Gelnett, D., Sullivan, J., Soli, S. D., and Goldberg, R. L. (**1992**). ''Norms for the hearing in noise test: The influence of spatial separation, hearing loss, and English language experience on speech reception thresholds,'' J. Acoust. Soc. Am. **92**, 2385.

Peissig, J., and Kollmeier, B. (**1997**). ''Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners,'' J. Acoust. Soc. Am. **101**, 1660–1670.

Pollack, I. (**1975**). ''Auditory informational masking,'' J. Acoust. Soc. Am. Suppl. 1 **57**, S5.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). ''Speech recognition with primarily temporal cues,'' Science **270**, 303–304.

Simpson, B. D., Bolia, R. S., Ericson, M. A., and McKinley, R. L. (**1999**). ''The effect of sentence onset asynchrony on call sign detection and message intelligibility in a simulated 'cocktail party,' '' J. Acoust. Soc. Am. **105**, 1024.

Watson, C. S. (**1987**). ''Uncertainty, informational masking and the capacity of immediate auditory memory,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 267–277.

Watson, C. S., Kelly, W. J., and Wroton, H. W. (**1976**). ''Factors in the discrimination of tonal patterns: II. Selective attention and learning under various levels of stimulus uncertainty,'' J. Acoust. Soc. Am. **60**, 1175–1181.

Wilson, R. H., and Carter, A. S. (**2001**). ''Relation between slopes of word recognition psychometric functions and homogeneity of the stimulus materials,'' J. Am. Acad. Audiol **12**, 7–14.

Wilson, R. H., Zizz, C. A., Shanks, J. E., and Causey, G. D. (**1990**). ''Normative data in quiet, broadband noise, and competing message for Northwestern University Auditory Test No. 6 by a female speaker,'' J. Speech Hear Disord. **55**, 771–778.

Yost, W. A. (**1997**). ''The cocktail party problem: Forty years later,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. A. Gilkey and T. R. Anderson (Erlbaum, Hillsdale, NJ), pp. 329–348.

Zurek, P. M. (**1993**). ''Binaural advantages and directional effects in speech intelligibility,'' in *Acoustical Factors Affecting Hearing Aid Performance*, edited by G. A. Studebaker and I. Hochberg (Allyn and Bacon, Boston, MA), pp. 255–276.

Tests of a within-channel account of comodulation detection differences

Brian C. J. Moore[a] and Stephen J. Borrill
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, United Kingdom

(Received 17 January 2002; revised 3 May 2002; accepted 23 July 2002)

The threshold for detecting a narrow-band noise signal in one or more masking noise bands is higher when the signal and masker bands have the same envelope (correlated condition) than when they have independent envelopes (uncorrelated condition). This comodulation detection difference (CDD) might be caused by perceptual grouping of the signal and masker bands when they are correlated. Alternatively, CDD may occur because, in the uncorrelated condition, the signal can be detected in the dips of the masker. A previous paper [S. J. Borrill and B. C. J. Moore, J. Acoust. Soc. Am. **111**, 309–319 (2002)] described results and a model supporting a dip-listening explanation. The model predicted steeper psychometric functions for the correlated than for the uncorrelated condition, a prediction confirmed by experiment 1. In experiment 2, the width of the signal and masker bands was varied. The dip-listening model predicts a small decrease in CDD with increasing bandwidth, while an account based on perceptual grouping predicts a substantial decrease, as across-channel sensitivity to envelope disparity decreases with increasing envelope modulation rate. The CDD was independent of bandwidth. Experiment 3 showed no effect of masker–signal onset asynchrony on CDD, even though asynchrony should reduce perceptual grouping. An explanation of CDD is proposed based on the suppression that has been observed in cochlear mechanics and in the auditory nerve. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1508793]

PACS numbers: 43.66.Dc, 43.66.Mk [MRL]

I. INTRODUCTION

Several researchers have studied the ability to detect a noise-band signal, centered at frequency f_s, in the presence of one or more noise bands centered away from f_s. These bands have been referred to as masking bands, cue bands, or flanking bands (FBs), and their envelopes have either been correlated (comodulated) with that of the signal band, or uncorrelated with it. In many such experiments, the envelopes of the masking bands were correlated with each other, but not with that of the signal band envelope; the maskers in this case are referred to as ''co-uncorrelated.'' Typically, detection of the signal is better in the co-uncorrelated condition than when all bands, for both signal and masker, have the same envelope (correlated condition) (Cohen and Schubert, 1987a; McFadden, 1987). The difference in threshold is often referred to as a ''comodulation detection difference'' (CDD).

One explanation for CDD is that it depends partly on perceptual grouping processes. Common modulation of noise bands in different frequency regions may cause the noise bands to be perceptually fused, while independent modulation may cause them to be perceptually segregated (Bregman *et al.*, 1990; Darwin and Carlyon, 1995; Moore, 1997). It may be easier to detect a narrow-band noise when it is perceptually segregated from the FBs (the co-uncorrelated condition) than when it is grouped together with the FBs (the correlated condition) (McFadden, 1987; Wright, 1990).

Perceptual grouping may be regarded as a process that involves comparison of the outputs of auditory filters tuned to different center frequencies, i.e., an cross-channel process, and CDD has often been discussed in terms of an across-channel process (McFadden, 1987; Wright, 1990). However, CDD might arise partly or exclusively from within-channel processes, i.e., from processes dependent on the output of a single auditory filter centered close to the signal frequency (Cohen and Schubert, 1987a, 1987b; Fantini and Moore, 1994). Subjects may be able to listen for the signal in the ''dips'' of the uncorrelated masker, using the output of an auditory filter centered close to the signal frequency (Zwicker and Schutte, 1973; Buus, 1985; Cohen and Schubert, 1987a, 1987b; Moore and Glasberg, 1987; Borrill and Moore, 2002). However, in the correlated condition dips in the masker coincide with dips in the signal, and there is no selective advantage to be obtained from dip listening. This could account for CDD.

A review of previous research on the mechanisms underlying CDD was given by Borrill and Moore (2002). They concluded that there was more evidence for a within-channel dip-listening explanation than for an explanation based on perceptual grouping. However, the evidence was not conclusive. Therefore, they conducted experiments to test the two explanations. A within-channel explanation leads to the prediction that masking in the correlated condition should only occur when the masker produces significant excitation at the signal ''place.'' Thus, the amount of masking in the correlated condition should approach zero at about the same signal–masker frequency separation as for the co-uncorrelated condition. On the other hand, there is no obvious reason why perceptual grouping should not occur for widely separated frequency bands. If higher thresholds in the correlated condition are due to perceptual grouping of the

[a]Electronic mail: bcjm@cus.cam.ac.uk

signal and masker, some masking in that condition would be expected to occur even when the masker produces no excitation at the signal place. Thus, masking would be expected for the correlated condition at masker–signal frequency separations where there is no masking for the co-uncorrelated condition. In their first experiment, Borrill and Moore found that, contrary to the prediction based on the perceptual-grouping explanation, the amount of masking in the correlated condition approached zero at about the same signal–masker frequency separation as for the co-uncorrelated condition; CDD did not occur when the signal and masker bands were widely separated in frequency.

In their second experiment, Borrill and Moore (2002) used two masker bands, one above and one below the signal frequency, and included conditions where only the upper or the lower masker band was comodulated with the signal band. Also, the overall level of the masker was systematically varied. The results indicated that the magnitude of CDD was determined by the comodulation of the signal band with the masker band producing the most masking. These results support an explanation based on the spread of excitation and dip listening, rather than an explanation based on perceptual grouping.

One apparent difficulty for the dip-listening explanation comes from consideration of results for the ''all-uncorrelated condition,'' in which all bands (both masker and signal) are independent. One might expect results for this condition to be intermediate between those for the correlated and uncorrelated conditions, as some dip listening would be possible in the all-uncorrelated condition, but it would be less effective than for the co-uncorrelated condition. McFadden and Wright (1990) and Wright (1990), using 100-Hz-wide bands, reported that thresholds were similar for the correlated and all-uncorrelated conditions, which is inconsistent with this expectation. However, Fantini and Moore (1994) and Borrill (1997), using narrower bands, found that thresholds in the all-uncorrelated condition were intermediate between those for the correlated and co-uncorrelated conditions, as expected from the dip-listening hypothesis. Possibly, dip listening becomes less effective when the noise bands are broader, and hence have relatively fast fluctuations. We will return to this point later.

Borrill and Moore (2002) presented a within-channel model of CDD based on the concept of dip listening. In the model, a sliding temporal window was used to select a number of sequential sections of the masker and signal. The output of the window was fed to a simulated auditory filter centered on the signal frequency. The level dependence of the filter was controlled by the instantaneous masker level in the chosen section. The signal-to-masker ratio at the output of filter was calculated and compared to a criterion ratio. If the signal-to-masker ratio was above the criterion, then that section was counted as providing a detection opportunity or ''look'' in which it was possible for the signal to be detected (Viemeister and Wakefield, 1991). The sliding window was advanced throughout the sample and the total number of detection opportunities was calculated, from which the ''cumulative detection time'' (CDT) was calculated. It was assumed that performance was monotonically related to the CDT. The

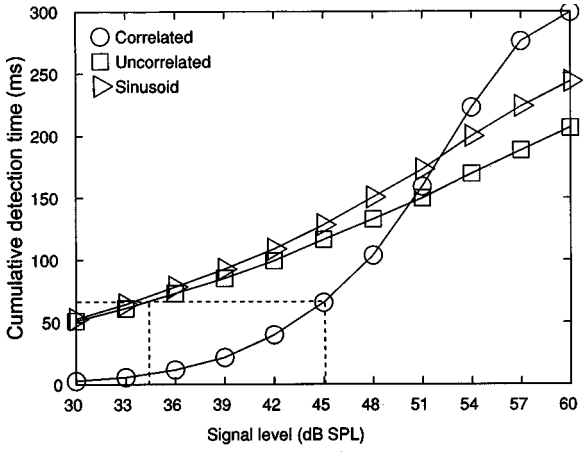


FIG. 1. Cumulative detection time (CDT) plotted as a function of signal level for the three conditions tested by Borrill and Moore (2002); correlated (circles), co-uncorrelated (squares), and sinusoid (triangles). The horizontal dashed line is plotted at the CDT that is assumed in the model to be required for threshold. Vertical dashed lines indicate the signal thresholds for the co-uncorrelated and correlated conditions. The figure is reproduced from Borrill and Moore (2002).

general pattern of the results obtained in the first experiment of Borrill and Moore (2002) could be accounted for rather well if the criterion signal-to-masker ratio was set to 2.17 dB and the threshold was assumed to correspond to a CDT of 59.5 ms. This model is clearly oversimplified. For example, it may not be the case that all of the looks are treated equally; an optimal decision rule would require that a look with a higher to signal-to-masker ratio contributes more to detection than a look that only just meets the criterion signal-to-masker ratio (Viemeister and Wakefield, 1991). However, it appears that human listeners do not always optimally combine information from independent looks (Buus *et al.*, 1986; Moore and Sek, 1994). We believe that the current model is adequate for predicting the general form of the results to be expected from a dip-listening mechanism.

Figure 1, reproduced from Borrill and Moore (2002), shows the CDT plotted as a function of signal level for the three conditions tested by Borrill and Moore, correlated, co-uncorrelated, and sinusoidal (where the signal was a sinusoid rather than a band of noise), for $\Delta f = 600$ Hz, taking the criterion signal-to-masker ratio as 2.17 dB. At the criterion CDT of 59.5 ms, the signal levels in the correlated and co-uncorrelated conditions correspond to the mean obtained thresholds (threshold was defined as the signal level needed for 79% correct in a three-alternative forced-choice task). This is as it should be, since the criterion CDT was selected to achieve this condition. More importantly, for the sinusoidal condition, the signal level at the criterion CDT corresponded very well with the mean obtained threshold; the predicted threshold was very slightly below that for the co-uncorrelated condition, which corresponds exactly to the obtained result.

In this paper we present three experiments designed to provide further tests of the explanations of CDD described above. The results are not consistent with an explanation based on perceptual grouping, but they are also not fully consistent with the dip-listening model described above. In

the final discussion section, we propose a new explanation for CDD based on suppression. This explanation is consistent with all of the results, and it has the advantage that it does not require a mechanism to select the brief time intervals over which the signal-to-noise ratio is high.

In the work described in this paper, we used either the correlated condition or the co-uncorrelated condition, in which the envelopes of the masker bands were comodulated with each other but not with that of the signal band. For brevity, in the remainder of this paper the latter will simply be referred to as the uncorrelated condition.

II. EXPERIMENT 1

A. Rationale

In Fig. 1, the slope of the function relating the CDT to signal level is steeper for the correlated than for the uncorrelated condition. Indeed, the model predicts that at sufficiently high signal levels, a correlated signal should be more detectable than an uncorrelated signal. The reason for this is that at high signal levels in the correlated condition the criterion signal-to-masker ratio will always be exceeded. This is not true for the uncorrelated condition; even if the signal level is high enough for the signal to be highly detectable, there will be brief epochs when the signal-to-masker ratio is below the criterion value.

If it is assumed that detectability is monotonically related to the CDT, then the functions in Fig. 1 should be monotonically related to psychometric functions. The model therefore predicts that psychometric functions should be steeper for the correlated than for the uncorrelated condition. The first experiment tested this prediction.

B. Method

1. Stimuli and conditions

The task was to detect a 20-Hz-wide band of noise centered at 1500 Hz. Two masking bands (also 20 Hz wide) were centered at 900 and 2100 Hz. The masking bands were always correlated with each other. The signal could be either correlated or uncorrelated with the masking bands. The noise-band signal and maskers were generated digitally using either a Masscomp 5400 or a Silicon Graphics Indy computer. Two ''tracks'' were generated. The signal was put on one track and the maskers on the other. Each 20-Hz-wide band of noise was produced by adding together 201 sinusoids spaced 0.1 Hz apart, each with a random phase and amplitude (being drawn from a rectangular and Rayleigh distribution, respectively). Each band of noise was therefore a harmonic series with a 0.1-Hz fundamental. One full period (10 seconds) of each track was calculated, after which the track could be recycled, to create a continuous noise.

To make two bands correlated, the phase and amplitude of each component in one band were the same as for the analogous component in the other band. For example, the first components in the two bands had the same phase and amplitude, even though they had different frequencies. The signal and masker were played out simultaneously through twin 16-bit digital-to-analog converters (Masscomp DA04H) at a sampling rate of 16 000 Hz. They were low-pass filtered at 7 kHz using Kemo VBF8 filters (96-dB/oct slope) and then recorded onto separate channels of a digital audiotape (DAT—Sony 55ES). The tape was then played continuously for each condition.

The signal was gated with 50-ms raised-cosine ramps and with a steady-state duration of 300 ms. The signal was presented at random in one of three observation intervals which were marked by lights on the response box. The intervals were separated by 200 ms. The masker was gated simultaneously with the three intervals in which the signal might appear. The start of each trial was initiated by the subject's response on the previous trial. As the subjects did not respond after a fixed time, the noise samples on a specific trial were, in effect, randomly chosen from the continuously recycling signal and masker tracks.

Signal timing was controlled by a Texas Instruments 990/4 computer. Two analog multipliers (AD534L) connected in series were used as a gate. The gating voltage was derived from a 12-bit digital-to-analog converter. The signal level was controlled by a Charybdis model D programmable attenuator. The signal and masker were mixed using an active adder and then passed through a manual attenuator before being delivered to the left earpiece of a Sennheiser HD 414 headset. The masker bands each had a spectrum level of 65 dB (*re*: 20 μPa) and an overall level of 78 dB SPL.

A continuous low-pass noise was added to the stimuli to mask possible combination bands (Greenwood, 1971, 1972). The noise was produced by filtering a white noise (Hewlett Packard 3722A) using four 48-dB/octave low-pass filters (Kemo VBF/8) in series. The cutoff frequency (-3 dB) of each filter was set to 400 Hz. The level of the low-pass noise was adjusted so that it had a spectrum level of 38 dB at 300 Hz (the frequency at which a cubic-difference band may have been present).

2. Procedure

A three-interval, three-alternative forced-choice procedure was used. Subjects were required to indicate the interval containing the signal by pressing the appropriate button on a response box. Five signal levels were presented in a cyclic order decreasing in level, as described by Moore and Sek (1992). The highest level was chosen to make the signal relatively highly detectable so that the subject heard a reminder ''easy'' signal every five trials. This helps to prevent subjects from ''losing'' the appropriate detection cues (Taylor *et al.*, 1983). The signal levels were equally spaced in dB and were chosen separately for each subject on the basis of pilot trials so as to cover the region of the psychometric function between chance performance and near-ceiling performance. In each run, every signal level was presented ten times (50 trials per run). For each subject, 20 complete runs were obtained so that there were 200 trials per signal level.

3. Subjects

Three subjects participated, all with absolute thresholds less than 10 dB HL at all audiometric frequencies. All subjects had extensive practice in similar tasks.

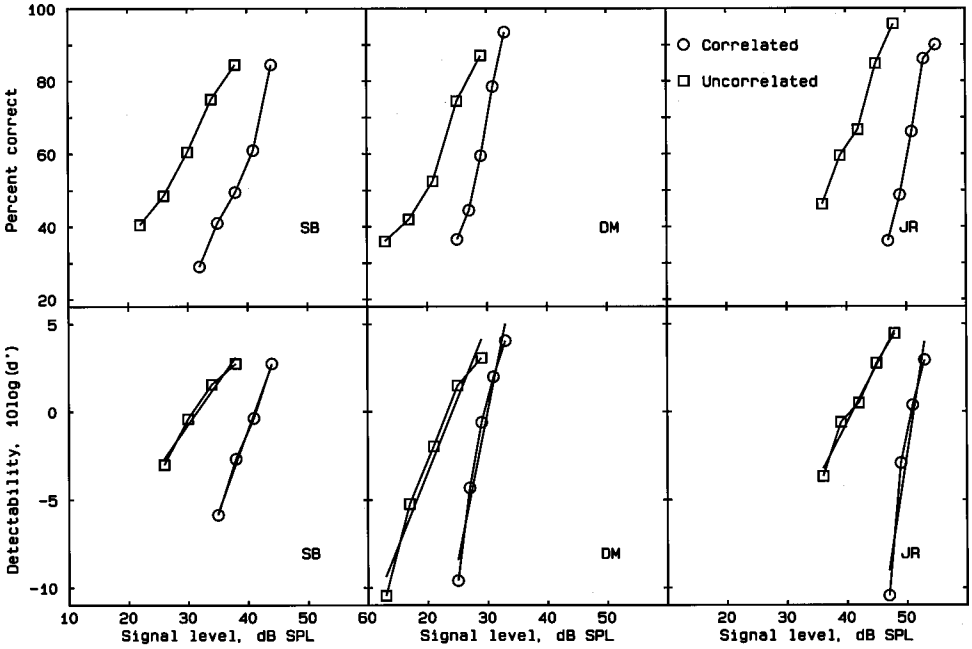


FIG. 2. Results of experiment 1, showing psychometric functions for detection of the signal band for the correlated and uncorrelated conditions. Each column represents a different subject. The top row shows the percentage-correct scores. The bottom rows shows the scores converted to d' values. Linear regression lines fitted to the d' values are shown.

C. Results and discussion

Standard tables (Hacker and Ratcliff, 1979) were used to convert the percentage-correct values to d', the measure of detectability. Figure 2 shows the results for each subject, plotted both as percent correct (top) and as $\log(d')$ (bottom). In the lower panels, the data points were not plotted when d' was negative. Linear best-fitting lines to the $\log(d')$ values were calculated. The slopes of these lines are shown in Table I. The slopes are clearly very different for the two conditions. To assess the significance of the slope differences, one-tailed *t*-tests were performed (Snedecor and Cochran, 1967, Chap. 6; Howell, 1997, Chap. 9); we used one-tailed tests as we predicted that the slopes would be steeper for the correlated than for the uncorrelated condition. The *t*-tests are based on the standard errors of the slopes (estimated by the deviations of the actual values from the fitted values), with degrees of freedom equal to $N1+N2-4$, where $N1$ is the number of data points contributing to the first slope estimate and *N2* is the number of data points contributing to the second slope estimate ($N1=N2=4$ or 5 for our data). The *t* values were significant for all subjects: for SB, $t(4)=6.68$, $p<0.01$; for DM, $t(6)=4.0$, $p<0.01$; for JR, $t(5)=3.66$, $p<0.02$. The mean slope of the psychometric function for the correlated condition, 1.6, is over twice that for the uncorrelated condition, 0.66.

TABLE I. Slopes of the best fitting lines to the psychometric function [$\log(d')$ versus signal level] for each subject, for the correlated (C) and uncorrelated (U) conditions.

Subject	SB		DM		JR	
Condition	C	U	C	U	C	U
Slope	0.936	0.480	1.68	0.85	2.17	0.65

The observed difference in slopes is consistent with the within-channel dip-listening model proposed by Borrill and Moore (2002) and outlined in the Introduction. We can think of no obvious way of accounting for the slope difference with an across-channel mechanism based on perceptual grouping.

III. EXPERIMENT 2. EFFECTS OF MASKER AND SIGNAL BANDWIDTH

A. Rationale

It seems reasonable to assume that, if CDD is caused by perceptual grouping mechanisms, then it should only occur under conditions where subjects are able to discriminate across-frequency differences in envelope correlation (Richards, 1987). One factor that affects the ability to detect such across-frequency differences is the rapidity of the envelope fluctuations, and this can be manipulated by varying the noise bandwidth. As the bandwidth is increased, the envelope fluctuations become more rapid. For a rectangular band of noise, the mean number of envelope maxima per second is 0.641 times the bandwidth (Rice, 1954). For example, for a 20-Hz bandwidth the mean number of maxima per second is 12.8, while for a 160-Hz bandwidth it is 102.6. It appears that, for noise bandwidths up to 100 Hz, subjects are able to distinguish two widely spaced bands with correlated envelopes from two widely spaced bands with uncorrelated envelopes, although performance is not perfect; d' is typically about 1 for bands centered at high frequencies, and somewhat less for bands centered at low frequencies (Moore and Emmerich, 1990; Hall and Grose, 1993).

One would expect that, for bandwidths above 100 Hz, envelope correlation discrimination would worsen, although there appears to be no data directly testing this expectation.

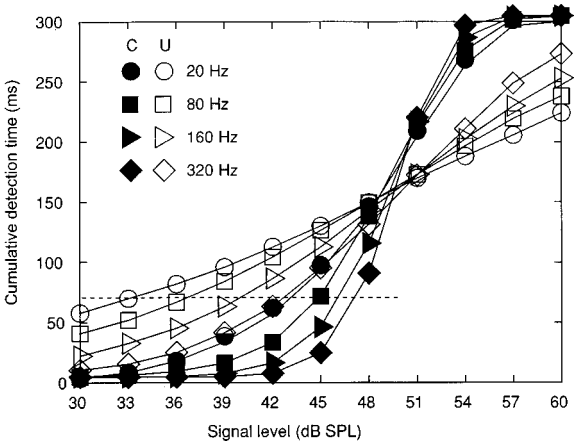


FIG. 3. Values of the cumulative detection time (CDT) produced by the model for various masker–signal bandwidths (20 to 320 Hz), plotted as a function of signal level, for the correlated (circles) and co-uncorrelated (squares) conditions. The horizontal dashed line is plotted at the value of the CDT required for threshold to be reached. The signal levels at which this line intersects the curves represent the predicted thresholds.

However, the ability to compare the phase of amplitude modulation applied to two widely separated sinusoidal carrier frequencies decreases markedly for modulation rates above 64 Hz, when precautions are taken to prevent within-channel cues from being used (Wakefield and Edwards, 1987; Strickland *et al.*, 1989; Yost and Sheft, 1989). If CDD depends on across-channel perceptual grouping processes based on envelope similarity, one would expect it to decrease markedly with increasing bandwidth (and increasing average fluctuation rate), simply because subjects are less able to discriminate across-channel envelope disparities when the envelope fluctuations are rapid. However, a mechanism based on within-channel dip-listening will not necessarily be substantially affected by the bandwidth of the signal and masker.

Figure 3 shows values of the CDT produced by the model for various masker–signal bandwidths (20 to 320 Hz), plotted as a function of signal level. Signal and masker frequencies and the masker level were the same as for experiment 1. The horizontal dashed line is plotted at the value of the CDT (59.5 ms) required for threshold to be reached (corresponding to 79% correct in a 3AFC task), as determined by Borrill and Moore (2002). The signal levels at which this line intersects the curves represent the predicted thresholds. The model predicts that thresholds should increase with increasing bandwidth for both the correlated (C) and uncorrelated (U) conditions, by about 5 dB for the former and 10 dB for the latter, over the range 20 to 320 Hz. The magnitude of CDD (the difference between thresholds in the C and U conditions), is predicted to decrease by about 5 dB over the same range of bandwidths. The predicted decrease in CDD is linked to the duration of the sliding temporal window used in the model, which was set to 12.5 ms. In our earlier work (Borrill and Moore, 2002) and in experiment 1 of the present paper, the exact duration of the window was not critical, as only a single noise bandwidth of 20 Hz was used, which gave a mean time between envelope peaks of about 78 ms. If a smaller temporal window duration were used (Plack and Moore, 1990; Oxenham and Moore, 1994), it would not affect the predictions for our earlier results, but the predicted effect of bandwidth on CDD would be smaller. For example, with a window duration of 8 ms, the threshold in the uncorrelated condition would be predicted to increase by only 6–7 dB when the bandwidth was increased from 20 to 320 Hz, as compared to about 10 dB with the 12.5-ms window, and CDD would be predicted to decrease by only about 3 dB as compared to 5 dB with the 12.5-ms window.

In summary, an across-channel mechanism based on perceptual grouping leads to the prediction that CDD should decrease markedly with increasing bandwidth. In contrast, a mechanism based on within-channel dip listening leads to the prediction that CDD will not necessarily be substantially affected by the bandwidth of the signal and masker. In experiment 2 we explored the effect of systematically varying the bandwidth of the signal and masker bands.

B. Method

1. Stimuli and conditions

We wanted to be able to use a reasonably wide range of bandwidths while keeping the bandwidth well below the bandwidth of the auditory filter, to prevent the envelope fluctuations from being affected by auditory filtering. Therefore, we decided to use a relatively high signal center frequency of 4000 Hz. For this center frequency, the auditory filter bandwidth (ERB) is about 480 Hz (Glasberg and Moore, 1990). Bandwidths of the signal and masker were 20, 40, 80, or 160 Hz.

The signal was presented with a masker consisting of two bands of noise of the same bandwidth. The masking bands were linearly spaced in frequency on each side of the signal band at 2400 and 5600 Hz. For a center frequency of 2400 Hz, the auditory filter bandwidth is about 305 Hz, which is still markedly greater than the largest noise bandwidth used. Each masker noise band was presented at a nominal overall level of 68 dB SPL. However, the response of the Sennheiser HD414 earphone at the eardrum has a broad peak around 3.5 kHz, so the effective level of each noise band was closer to 78 dB SPL, the same as the overall level at the eardrum of the 20-Hz bandwidth noise used in experiment 1. A low-pass noise with a 3-dB point at 900 Hz and a nominal spectrum level of 28 dB was presented continuously to mask combination bands, in particular the band centered at 800 Hz. At this center frequency, the level per ERB of the low-pass noise was about 58 dB. This should have been sufficient to mask the combination band, given that the frequency ratio of the two primary bands was 1.67 (Greenwood, 1972). The timing of the stimuli was the same as for experiment 1, and the stimuli were generated in the same way as for experiment 1.

2. Procedure

Masked thresholds were estimated using a three-alternative forced-choice method with an adaptive three-down, one-up procedure estimating the 79.4% point on the psychometric function (Levitt, 1971). The signal level at the start of a run was set about 10 dB above the threshold value

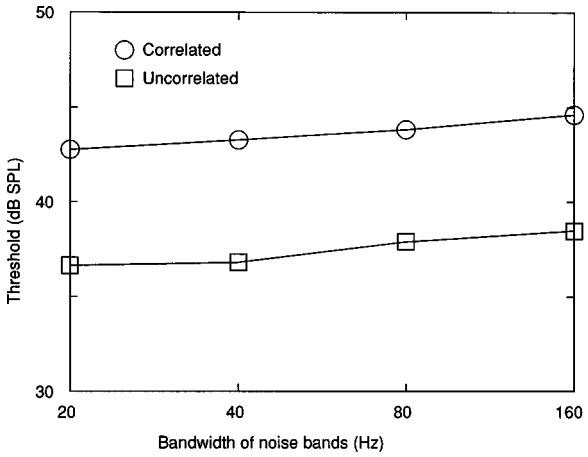


FIG. 4. Mean results across the three subjects for experiment 2, showing the signal threshold in the correlated and uncorrelated conditions, plotted as a function of masker/signal bandwidth.

estimated in pilot runs. The initial step size was 5 dB. After three reversals, the step size was decreased to 2 dB. A run was terminated after 12 reversals. The threshold was defined as the mean of the levels at the last eight reversals. Three threshold estimates were obtained for each condition, and a fourth was obtained if the range of the first three exceeded 3 dB. This happened for about 20% of the conditions. The final threshold value was taken as the mean of these three (or four) estimates. Subjects were tested individually in a double-walled sound-attenuating booth. Feedback was provided on the response box.

3. Subjects

Three subjects participated, all with absolute thresholds less than 10 dB HL at all audiometric frequencies. All subjects had previously completed several earlier experiments measuring CDD with similar stimuli.

C. Results and discussion

The pattern of results was very similar across subjects, although one subject, DM, gave thresholds that were about 10 dB lower than for the other two subjects. The individual differences in overall threshold values presumably reflect differences in frequency selectivity across subjects (Moore, 1987). Figure 4 shows mean results across the three subjects. A clear CDD was present for all bandwidths, but its magnitude did not change with bandwidth. The standard deviation of the CDD across subjects was 1.7, 0.9, 1.6, and 0.6 dB, for bandwidths of 20, 40, 80, and 160 Hz, respectively. Note that these standard deviations slightly overestimate the inherent variability of the CDD values, since subject GM consistently gave slightly lower CDDs than the other two subjects (by 1–2 dB). A within-subjects analysis of variance was conducted on the thresholds with factors condition (correlated versus uncorrelated) and bandwidth. The main effect of condition was significant, confirming that a CDD occurred; $F(1,2)=117.9$, $p=0.008$. However, the main effect of bandwidth was not significant; $F(3,6)=0.74$, $p=0.56$. Also, the interaction of condition and bandwidth was not significant; $F(3,6)=0.1$, $p=0.96$. This indicates that the magnitude of CDD did not vary with bandwidth. Fantini and Moore (1994) also found that the magnitude of CDD hardly varied with bandwidth, although the maximum bandwidth that they used was 64 Hz (the range was from 4 to 64 Hz).

Thresholds for both the correlated and the uncorrelated conditions tended to increase with increasing bandwidth. However, the increase over the range tested was only 1.5 to 2 dB, and was not statistically significant. Fantini and Moore (1994) found a slight decrease in thresholds for the correlated and uncorrelated conditions as the noise bandwidth was increased from 4 to 64 Hz. The model predicts an increase of about 7 dB for the uncorrelated condition and about 3 dB for the correlated condition over the range of bandwidths tested here. Also, the CDD remained constant at about 7 dB with increasing bandwidth, while the model predicts a decrease from about 9.5 to about 6 dB. As noted earlier, the effects of bandwidth predicted by the model are linked to the duration of the temporal window assumed in the model. A smaller window duration would lead to smaller predicted effects of bandwidth, which would be more in line with the data. However, for a noise bandwidth of 160 Hz, the average number of envelope maxima per second is about 103, so the average time between maxima is only 9.8 ms. To account for the finding that CDD did not change with bandwidth for bandwidths up to 160 Hz, and that thresholds *per se* changed only slightly with bandwidth, the temporal window duration in the model would have to be decreased to about 2 ms, which would be inconsistent with previous estimates of the equivalent rectangular duration of the temporal window of the auditory system (Moore *et al.*, 1988; Plack and Moore, 1990; Oxenham and Moore, 1994). Therefore, the lack of effect of masker bandwidth on CDD casts some doubt on the validity of the dip-listening model proposed by Borrill and Moore (2002).

The results are not consistent with the predictions derived from an across-channel mechanism based on perceptual grouping: since sensitivity to across-channel envelope disparities decreases markedly for envelope modulation rates above 64 Hz, such a mechanism leads to the prediction that there should have been a marked decrease in CDD for the widest bandwidth used. This was not the case.

IV. EFFECTS OF MASKER–SIGNAL ONSET ASYNCHRONY

A. Rationale

The maskers and signals used in experiments 1 and 2 and in the experiments of Borrill and Moore (2002) all had simultaneous onsets and offsets. This would have promoted perceptual grouping of the signal and masker bands (Rasch, 1978; Darwin, 1984; Moore *et al.*, 1986). Onset asynchrony has been reported as having powerful effects in promoting perceptual segregation (Bregman, 1978; Rasch, 1978; Darwin, 1984; Bregman, 1990; Roberts and Moore, 1991; Darwin and Carlyon, 1995). If fact, onset asynchrony can produce perceptual segregation even when other cues would promote perceptual fusion (Bregman, 1990; Darwin and Carlyon, 1995). Therefore, if perceptual grouping of the masker and signal bands in the correlated condition plays a role in

producing CDD, then introducing a sufficiently large onset asynchrony between the signal and masker bands should reduce threshold in the correlated condition, and reduce the magnitude of CDD. That prediction was tested in experiment 3. The effect of onset asynchrony on CDD has previously been explored by McFadden and Wright (1990), and our results will be compared to theirs later on in this paper.

B. Method

1. Stimuli and listening conditions

The task of the subjects was to detect a 20-Hz-wide band of noise centered at 1500 Hz. The signal was presented with a masker consisting of either two or six 20-Hz-wide bands of noise. The masker bands were always correlated with each other. The signal band could be either correlated or uncorrelated with the masker bands. When two masker bands were used, these were centered at 1071 and 2100 Hz. A low-pass noise with a cutoff frequency of 700 Hz and a spectrum level within its passband of 38 dB was added to mask possible combination bands. If six masker bands were used, these were centered at 547, 765, 1071, 2100, 2940, and 4116 Hz. These frequencies were chosen so that the bands were spaced from each other (and from the signal frequency) by roughly 2.5 ERBs (Glasberg and Moore, 1990). No low-pass noise was used in this case, as the masker bands remote from the signal frequency would effectively mask any combination bands produced by interaction of the signal and the adjacent masker bands. The overall level of each masker band was 76 dB SPL.

The onset asynchrony between the masker and signal was either 0 or 200 ms. The signal and masker offsets were simultaneous. The signal always had a steady-state duration of 300 ms with 50-ms raised cosine ramps. The masker was therefore presented for either 300 ms (synchronous condition) or 500 ms (asynchronous condition). The interstimulus interval was 200 ms. Stimuli were generated in the same way as for experiment 1.

2. Procedure and subjects

The procedure and subjects were the same as for experiment 2.

C. Results and discussion

The pattern of the results was similar across subjects, although again subject DM showed lower overall thresholds than the other subjects. The mean results are shown in Fig. 5. Independently of the number of masking bands or the correlation of the envelopes of the masker and signal, introducing a 200-ms onset asynchrony had little or no effect on the signal threshold; the largest effect of asynchrony was 1.1 dB, which occurred for the two-band masker in the correlated condition. The effect of asynchrony was also small in the data for the individual subjects. The means and standard deviations (sds) across subjects of the differences between thresholds for the synchronous and asynchronous conditions were (sds in parentheses): 1.1 (0.1) dB for the correlated two-band condition, 0.7 (1.0) dB for the correlated six-band condition, −0.1 (2.7) dB for the uncorrelated two-band condition, and 0.1 (0.8) dB for the uncorrelated six-band condition. A within-subjects analysis of variance was conducted on the thresholds, with factors condition (correlated versus uncorrelated), onset asynchrony, and number of masker bands. The main effect of condition was significant, confirming that a CDD occurred: $F(1,2)=44.1$, $p=0.022$. No other main effects or interactions were significant.

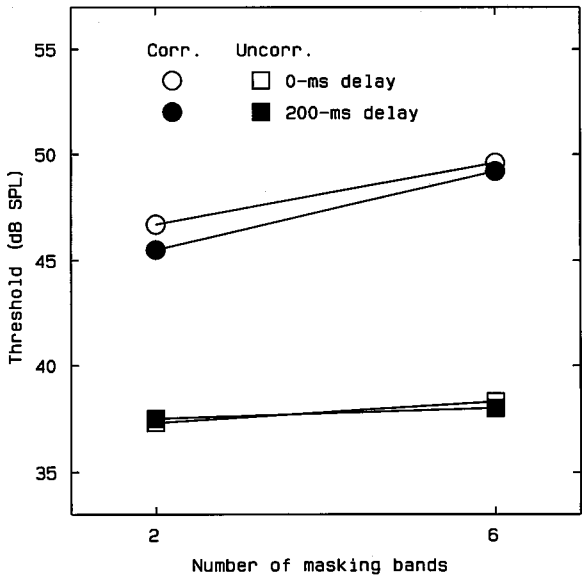


FIG. 5. Results of experiment 3, showing thresholds for detection of the signal band plotted as a function of the number of masking bands. Each panel shows results for one subject. The parameters were the masker–signal onset asynchrony (open symbols, 0 ms; filled symbols, 200 ms) and the condition (correlated, circles; uncorrelated, squares).

The lack of effect of onset asynchrony is inconsistent with a mechanism for CDD based on perceptual grouping. Signal thresholds were 1–3 dB higher for the six-band masker than for the two-band masker, but the effect of number of bands was not significant. The small effect of the number of bands is consistent with the inner bands being mainly responsible for the masking and the CDD, as assumed in our model.

McFadden and Wright (1990) also investigated the effect on CDD of introducing an onset asynchrony between the masker and signal bands. They used both long and short signal durations. For the short signal duration (steady state of 30 or 40 ms), the CDDs were very small. Their longer signal had a steady-state duration of 220 ms and 10-ms ramps, closer to the duration used in our experiments. Eight of the 13 subjects had relatively high thresholds. For these subjects, threshold declined as the asynchrony was increased. However, the decline was similar for correlated, uncorrelated, and all-uncorrelated conditions, so the magnitude of CDD was almost unaffected by the asynchrony. The other listeners tended to have lower thresholds, and for them thresholds hardly changed with asynchrony. The results of our experiment resemble those of the more sensitive group of subjects, perhaps reflecting the fact that our subjects were highly practiced. In any case, both our results and those of McFadden and Wright indicate that the magnitude of CDD is not affected by onset asynchrony of the masker and signal bands.

This is not consistent with an explanation of CDD based on perceptual grouping processes.

V. DISCUSSION

The experiments described in this paper provide further evidence to support the proposal of Borrill and Moore (2002) that CDD is not based on perceptual grouping processes, but probably depends on a process similar to ''listening in the dips'' (Zwicker and Schutte, 1973; Buus, 1985; Moore and Glasberg, 1987; Hall *et al.*, 1995). The results of experiment 1 are consistent with the prediction of the model proposed by Borrill and Moore that psychometric functions should be steeper for the correlated than for the uncorrelated condition. However, the results of experiment 2 showed smaller effects of bandwidth than predicted by the model. The discrepancy between the data and predictions could be reduced by shortening the duration of the temporal window used in the model, but to account for the complete lack of effect of bandwidth on CDD, and the small effect of bandwidth on thresholds *per se*, would require an implausibly short window duration. Thus, the results of experiment 2 cast some doubt on the validity of the model of Borrill and Moore. The results of experiments 2 and 3, together with earlier results of McFadden and Wright (1990), suggest that a perceptual-grouping mechanism based on across-channel comparisons plays little or no role in producing CDD.

The model of CDD proposed by Borrill and Moore (2002), and outlined in the Introduction, is incomplete in that it does not specify how subjects take advantage of moments when the signal-to-noise ratio is relatively high. Borrill and Moore discussed several possible ways in which this might occur.

One possibility is that the presence of the signal is indicated by a reduced effective modulation depth, as the signal band will tend to ''fill in'' the dips in the masker band (Schooneveldt and Moore, 1989). In other words, a change in envelope distribution is the detection cue, and this cue is more salient in the co-uncorrelated condition than in the correlated condition. Schooneveldt and Moore (1989) suggested that such envelope cues were partly responsible for the comodulation masking release that occurs when a noise masker is amplitude modulated by a narrow-band noise (Hall *et al.*, 1984). It has also been proposed that envelope cues are used for detection of a tone in narrow-band noise, especially when the overall masker level is roved (Kidd *et al.*, 1989, 1993; Richards, 1992; Richards and Nekrich, 1993). The ability of listeners to use these envelope cues appears to decrease as the envelope fluctuation rate increases (Schooneveldt and Moore, 1989). However, the results of experiment 2 showed almost no effect of signal and masker bandwidth on the magnitude of CDD and only a very small effect on the masked thresholds, which suggests that changes in envelope distribution *per se* were not used as a cue.

Another possibility discussed by Borrill and Moore (2002) is that subjects are able selectively to extract information at times corresponding to dips in the masker envelope. They might do this by making use of changes in the temporal fine structure of the signal (Schooneveldt and Moore, 1987; Fantini and Moore, 1994); during times when the masker is dominant, the fine structure would correspond to the masker center frequency, while during the masker dips the fine structure would partly reflect the signal frequency. If temporal fine structure cues were used, one would expect that the magnitude of CDD would decrease for high signal center frequencies, since the strength of neural phase locking to temporal fine structure decreases above about 1 kHz (Johnson, 1980; Palmer and Russell, 1986). The results of experiment 2 showed that a CDD of about 7 dB occurred at 4 kHz, which is only a little less than the CDD of 8–11 dB found at lower center frequencies in experiment 3 and in the experiments of Borrill and Moore (2002). It is not clear whether the amount of phase locking at 4 kHz would be sufficient to account for the CDD observed at that frequency.

A third possibility discussed by Borrill and Moore (2002) is that across-channel information is used to indicate the appropriate times to listen; temporal dips at the outputs of channels tuned away from the signal frequency could be used to indicate the optimum times to use the output of the filter centered at the signal frequency (Buus, 1985; Moore, 1988; Buus *et al.*, 1996). A mechanism similar to this has been proposed to account for comodulation masking release (Buus, 1985; Hall *et al.*, 1995; Buus *et al.*, 1996). For this type of explanation, information is extracted primarily from the output of a single auditory filter, but information from filters with other center frequencies indicates the optimum times at which to sample the output of the single filter.

One possibility not discussed by Borrill and Moore (2002) is that CDD depends on the operation of the suppression that has been observed in the auditory nerve (Sachs and Kiang, 1968; Arthur *et al.*, 1971; Javel *et al.*, 1983) and in cochlear mechanics (Ruggero *et al.*, 1992; Nuttall and Dolan, 1993; Rhode and Cooper, 1993; Cooper, 1996; Geisler and Nuttall, 1997). The response to a narrow-band signal at a place on the basilar membrane or in the auditory nerve tuned to the frequency of that signal can be reduced (suppressed) by a second signal with a different center frequency. It has been proposed that simultaneous masking is produced partly by suppression, especially when the masker frequency is remote from the signal frequency (Wightman *et al.*, 1977; Delgutte, 1988, 1990; Moore and Vickers, 1997; Gifford and Bacon, 2000). Suppression depends partly on the relative magnitude of the suppressor and the signal that is being suppressed, increasing as the suppressor-to-signal ratio increases. For low-frequency suppressors, it also tends to increase with the absolute level of the suppressor. In the correlated condition of a CDD experiment, dips in the masker envelope always coincide with dips in the signal envelope, so the strength of suppression would not vary greatly over time; however, because suppression tends to be stronger at higher levels of a low-frequency suppressor, signal peaks would be suppressed somewhat more than signal dips. In contrast, in the uncorrelated condition of a CDD experiment, dips in the masker envelope sometimes coincide with peaks in the signal envelope. At times when this happens, suppression will be relatively weak, and so the peaks in the signal result in relatively strong neural activity. Such a mechanism would result in a stronger overall representation of the signal in the uncorrelated condition.

Unlike the model proposed by Borrill and Moore (2002), an explanation based on suppression does not require a mechanism to select the moments when the signal-to-noise ratio is relatively high. Rather, in the uncorrelated condition, the effective signal level at such moments is automatically enhanced by a release from suppression. Suppression is a very fast-acting process (Arthur *et al.*, 1971). Thus, an explanation based on suppression is consistent with our finding that the magnitude of CDD was unaffected by the width of the signal and masker bands (and hence unaffected by the rate of envelope fluctuation). Furthermore, suppression operates at least as effectively at high frequencies as at low frequencies (Yates, 1995; Hicks and Bacon, 1999; Gifford and Bacon, 2000). This is consistent with our finding that a substantial CDD occurred for a signal frequency of 4 kHz, although our finding of a slightly smaller CDD at 4 kHz than at lower center frequencies suggests that suppression may not be the only factor involved.

An explanation of CDD based on suppression can account, at least qualitatively, for the finding of experiment 1 that psychometric functions were steeper for the correlated than for the uncorrelated condition. As the signal level increases, with the masker level held constant, the effect of suppression decreases. If suppression leads to a difference in effective signal excitation for the uncorrelated and correlated conditions, then a decrease in the magnitude of suppression will lead to a decrease in the difference between effective signal excitation for the correlated and uncorrelated conditions. This will lead to a convergence of the psychometric functions for the two conditions at high signal levels, as observed in the data.

It is of interest to consider what would be predicted for the all-uncorrelated condition described in the Introduction (masker bands uncorrelated with each other, as well as uncorrelated with the signal band), based on the effects of suppression. When multiple all-uncorrelated bands are present, the strongest suppression would occur at the few random times when the peaks in the envelopes of the bands coincided, and the greatest release from suppression would occur when the dips in the envelopes coincided. When the number of masker bands is large, the dips would coincide only rarely, so significant release from suppression would also occur rarely. Thus, signal thresholds for the all-uncorrelated condition should be higher than for the co-uncorrelated condition, but perhaps still below those for the correlated condition. As noted earlier, McFadden and Wright (1990) and Wright (1990) found that thresholds were similar for the correlated and all-uncorrelated conditions. They used eight 100-Hz-wide masker bands. Perhaps the combination of a large number of bands and a relatively wide noise bandwidth (leading to rapid fluctuations) prevents effective release from suppression. Fantini and Moore (1994) used four masker bands ranging in width from 4 to 6 Hz, while Borrill (1997) used six 20-Hz-wide masker bands. Both of these studies found that thresholds for the all-uncorrelated condition were in between those for the co-uncorrelated and correlated conditions.

It should be noted that the effects of suppression are quite similar to those associated with listening in the dips. For example, in describing a form of dip listening, Buus (1985) wrote "If the weights are set high when the masker energy is low (as indicated by low activity in channels tuned to frequencies away from the signal) and set low when the masker energy is high, quite efficient detection of the signal could result." Suppression would have precisely this type of effect, although the setting of weights would be done automatically by peripheral mechanisms, rather than requiring any form of cognitive processing.

VI. CONCLUSIONS

Several aspects of the results of earlier experiments and the present experiments are inconsistent with an explanation of CDD based on perceptual grouping of the signal and masker bands in the correlated condition. These aspects include

(1) In the correlated condition, no masking occurs when the signal and masker bands are widely separated in frequency (Borrill and Moore, 2002). With increasing masker–signal separation, the amount of masking becomes zero at about the same separation in the correlated and uncorrelated conditions.
(2) The ability to compare envelope fluctuations across frequency bands decreases with increasing envelope fluctuation rate, but the magnitude of CDD does not decrease as the bandwidth of the stimuli (and hence the rate of envelope fluctuations) increases (experiment 2).
(3) The magnitude of CDD is unaffected by an onset asynchrony between the masker and the signal, even though this would be expected to have a strong influence on perceptual grouping (McFadden and Wright, 1990, and experiment 3).

The results of experiment 2, showing that the magnitude of CDD was not affected by signal and masker bandwidth over the range 20–160 Hz, cast some doubt on the dip-listening model proposed by Borrill and Moore (2002). We have proposed an alternative model based on the concept of suppression. This model does not require an explicit mechanism to determine times when the signal-to-masker ratio is high. It can account for the main features of the results, including

(1) Psychometric functions are steeper in the correlated than in the uncorrelated condition.
(2) CDD is unaffected by masker/signal bandwidth.
(3) CDD is unaffected by masker-signal asynchrony.

ACKNOWLEDGMENTS

This work was supported by the MRC (UK). We thank Geoffrey Moore and Thomas Stainsby for helpful comments on an earlier version of this paper and Brian Glasberg for help with statistical analyses. We also thank Søren Buus and an anonymous reviewer for helpful comments.

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.
Borrill, S. J. (**1997**). ''Within- and across-channel processes contributing to comodulation detection differences,'' Ph.D. thesis, University of Cambridge.
Borrill, S. J., and Moore, B. C. J. (**2002**). ''Evidence that comodulation detection differences depend on within-channel mechanisms,'' J. Acoust. Soc. Am. **111**, 309–319.
Bregman, A. S. (**1978**). ''The formation of auditory streams,'' in *Attention and Performance VII*, edited by J. Requin (Erlbaum, Hillsdale, NJ).
Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (Bradford Books, MIT Press, Cambridge, MA).
Bregman, A. S., Liao, C., and Levitan, R. (**1990**). ''Auditory grouping based on fundamental frequency and formant peak frequency,'' Can. J. Psychol. **44**, 400–413.
Buus, S. (**1985**). ''Release from masking caused by envelope fluctuations,'' J. Acoust. Soc. Am. **78**, 1958–1965.
Buus, S., Schorer, E., Florentine, M., and Zwicker, E. (**1986**). ''Decision rules in detection of simple and complex tones,'' J. Acoust. Soc. Am. **80**, 1646–1657.
Buus, S., Zhang, L., and Florentine, M. (**1996**). ''Stimulus-driven, time-varying weights for comodulation masking release,'' J. Acoust. Soc. Am. **99**, 2288–2297.
Cohen, M. F., and Schubert, E. D. (**1987a**). ''The effect of cross-spectrum correlation on the detectability of a noise band,'' J. Acoust. Soc. Am. **81**, 721–723.
Cohen, M. F., and Schubert, E. D. (**1987b**). ''Influence of place synchrony on detection of a sinusoid,'' J. Acoust. Soc. Am. **81**, 452–458.
Cooper, N. P. (**1996**). ''Two-tone suppression in cochlear mechanics,'' J. Acoust. Soc. Am. **99**, 3087–3098.
Darwin, C. J. (**1984**). ''Perceiving vowels in the presence of another sound: Constraints on formant perception,'' J. Acoust. Soc. Am. **76**, 1636–1647.
Darwin, C. J., and Carlyon, R. P. (**1995**). ''Auditory grouping,'' in *Hearing*, edited by B. C. J. Moore (Academic, San Diego).
Delgutte, B. (**1988**). ''Physiological mechanisms of masking,'' in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London).
Delgutte, B. (**1990**). ''Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers,'' J. Acoust. Soc. Am. **87**, 791–809.
Fantini, D. A., and Moore, B. C. J. (**1994**). ''Profile analysis and comodulation detection differences using narrow bands of noise and their relation to comodulation masking release,'' J. Acoust. Soc. Am. **95**, 2180–2191.
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.
Gifford, R. H., and Bacon, S. P. (**2000**). ''Contributions of suppression and excitation to simultaneous masking: Effects of signal frequency and masker–signal frequency relation,'' J. Acoust. Soc. Am. **107**, 2188–2200.
Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.
Greenwood, D. D. (**1971**). ''Aural combination tones and auditory masking,'' J. Acoust. Soc. Am. **50**, 502–543.
Greenwood, D. D. (**1972**). ''Masking by combination bands: Estimation of the levels of the combination bands $(n+1)f_1-nf_h$,'' J. Acoust. Soc. Am. **52**, 1144–1154.
Hacker, M. J., and Ratcliff, R. (**1979**). ''A revised table of d' for M-alternative forced choice,'' Percept. Psychophys. **26**, 168–170.
Hall, J. W., and Grose, J. H. (**1993**). ''Monaural envelope correlation perception in listeners with normal hearing and cochlear impairment,'' J. Speech Hear. Res. **36**, 1306–1314.
Hall, J. W., Grose, J. H., and Mendoza, L. (**1995**). ''Across-channel processes in masking,'' in *Hearing*, edited by B. C. J. Moore (Academic, San Diego).
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.
Hicks, M. L., and Bacon, S. P. (**1999**). ''Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing,'' J. Acoust. Soc. Am. **105**, 326–338.
Howell, D. C. (**1997**). *Statistical Methods for Psychology*, 4th ed. (Duxbury, Belmont, CA).
Javel, E., McGee, J., Walsh, E. J., Farley, G. R., and Gorga, M. P. (**1983**). ''Suppression of auditory-nerve responses. II. Suppression threshold and growth, iso-suppression contours,'' J. Acoust. Soc. Am. **74**, 801–813.
Johnson, D. H. (**1980**). ''The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones,'' J. Acoust. Soc. Am. **68**, 1115–1122.
Kidd, G., Mason, C. R., Brantley, M. A., and Owen, G. A. (**1989**). ''Roving level tone-in-noise detection,'' J. Acoust. Soc. Am. **86**, 1310–1317.
Kidd, G., Uchanski, R. M., Mason, C. R., and Deliwala, P. S. (**1993**). ''Discriminability of narrow-band sounds in the absence of level cues,'' J. Acoust. Soc. Am. **93**, 1028–1037.
Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
McFadden, D., and Wright, B. A. (**1990**). ''Temporal decline of masking and comodulation detection differences,'' J. Acoust. Soc. Am. **88**, 711–724.
McFadden, D. M. (**1987**). ''Comodulation detection differences using noise-band signals,'' J. Acoust. Soc. Am. **81**, 1519–1527.
Moore, B. C. J. (**1987**). ''Distribution of auditory-filter bandwidths at 2 kHz in young normal listeners,'' J. Acoust. Soc. Am. **81**, 1633–1635.
Moore, B. C. J. (**1988**). ''Dynamic aspects of auditory masking,'' in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. Edelman, W. Gall, and W. Cowan (Wiley, New York).
Moore, B. C. J. (**1997**). *An Introduction to the Psychology of Hearing*, 4th ed. (Academic, San Diego).
Moore, B. C. J., and Emmerich, D. S. (**1990**). ''Monaural envelope correlation perception, revisited: Effects of bandwidth, frequency separation, duration, and relative level of the noise bands,'' J. Acoust. Soc. Am. **87**, 2628–2633.
Moore, B. C. J., and Glasberg, B. R. (**1987**). ''Factors affecting thresholds for sinusoidal signals in a narrow-band maskers with fluctuating envelopes,'' J. Acoust. Soc. Am. **82**, 69–79.
Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1986**). ''Thresholds for hearing mistuned partials as separate tones in harmonic complexes,'' J. Acoust. Soc. Am. **80**, 479–483.
Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (**1988**). ''The shape of the ear's temporal window,'' J. Acoust. Soc. Am. **83**, 1102–1116.
Moore, B. C. J., and Sek, A. (**1992**). ''Detection of combined frequency and amplitude modulation,'' J. Acoust. Soc. Am. **92**, 3119–3131.
Moore, B. C. J., and Sek, A. (**1994**). ''Effects of carrier frequency and background noise on the detection of mixed modulation,'' J. Acoust. Soc. Am. **96**, 741–751.
Moore, B. C. J., and Vickers, D. A. (**1997**). ''The role of spread of excitation and suppression in simultaneous masking,'' J. Acoust. Soc. Am. **102**, 2284–2290.
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 Moore, B. C. J. (**1994**). ''Modeling the additivity of nonsimultaneous masking,'' Hear. Res. **80**, 105–118.
Palmer, A. R., and Russell, I. J. (**1986**). ''Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells,'' Hear. Res. **24**, 1–15.
Plack, C. J., and Moore, B. C. J. (**1990**). ''Temporal window shape as a function of frequency and level,'' J. Acoust. Soc. Am. **87**, 2178–2187.
Rasch, R. A. (**1978**). ''The perception of simultaneous notes such as in polyphonic music,'' Acustica **40**, 21–33.
Rhode, W. S., and Cooper, N. P. (**1993**). ''Two-tone suppression and distortion production on the basilar membrane in the hook region of the cat and guinea pig cochleae,'' Hear. Res. **66**, 31–45.
Rice, S. O. (**1954**). ''Mathematical analysis of random noise,'' in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York).
Richards, V. M. (**1987**). ''Monaural envelope correlation perception,'' J. Acoust. Soc. Am. **82**, 1621–1630.
Richards, V. M. (**1992**). ''The detectability of a tone added to narrow bands of equal-energy noise,'' J. Acoust. Soc. Am. **91**, 3424–3435.
Richards, V. M., and Nekrich, R. D. (**1993**). ''The incorporation of level and level-invariant cues for the detection of a tone added to noise,'' J. Acoust. Soc. Am. **94**, 2560–2574.
Roberts, B., and Moore, B. C. J. (**1991**). ''The influence of extraneous sounds on the perceptual estimation of first-formant frequency in vowels under conditions of asynchrony,'' J. Acoust. Soc. Am. **89**, 2922–2932.
Ruggero, M. A., Robles, L., and Rich, N. C. (**1992**). ''Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-

nerve rate suppression,'' J. Neurophysiol. **68**, 1087–1099.
Sachs, M. B., and Kiang, N. Y. S. (**1968**). ''Two-tone inhibition in auditory nerve fibers,'' J. Acoust. Soc. Am. **43**, 1120–1128.
Schooneveldt, G. P., and Moore, B. C. J. (**1987**). ''Comodulation masking release (CMR): Effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotonic versus dichotic presentation of the flanking band,'' J. Acoust. Soc. Am. **82**, 1944–1956.
Schooneveldt, G. P., and Moore, B. C. J. (**1989**). ''Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration,'' J. Acoust. Soc. Am. **85**, 273–281.
Snedecor, G. W., and Cochran, W. G. (**1967**). *Statistical Methods* (Iowa University Press, Ames, IA).
Strickland, E. A., Viemeister, N. F., Fantini, D. A., and Garrison, M. A. (**1989**). ''Within- versus cross-channel mechanisms in detection of envelope phase disparity,'' J. Acoust. Soc. Am. **86**, 2160–2166.
Taylor, M. M., Forbes, S. M., and Creelman, C. D. (**1983**). ''PEST reduces bias in forced-choice psychophysics,'' J. Acoust. Soc. Am. **74**, 1367–1374.
Viemeister, N. F., and Wakefield, G. H. (**1991**). ''Temporal integration and multiple looks,'' J. Acoust. Soc. Am. **90**, 858–865.
Wakefield, G. H., and Edwards, B. (**1987**). ''Discrimination of envelope phase disparity,'' J. Acoust. Soc. Am. Suppl. 1 **182**, S41.
Wightman, F. L., McGee, T., and Kramer, M. (**1977**). ''Factors influencing frequency selectivity in normal and hearing-impaired listeners,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London).
Wright, B. A. (**1990**). ''Comodulation detection differences with multiple signal bands,'' J. Acoust. Soc. Am. 1 **87**, 292–303.
Yates, G. K. (**1995**). ''Cochlear structure and function,'' in *Hearing*, edited by B. C. J. Moore (Academic, San Diego).
Yost, W. A., and Sheft, S. (**1989**). ''Across-critical-band processing of amplitude-modulated tones,'' J. Acoust. Soc. Am. **85**, 848–857.
Zwicker, E., and Schutte, H. (**1973**). ''On the time pattern of the threshold of tone impulses masked by narrow band noise,'' Acustica **29**, 343–347.

Direct-to-reverberant energy ratio sensitivity

Pavel Zahorik[a)]
Waisman Center, University of Wisconsin-Madison, Madison, Wisconsin 53705

(Received 5 June 2001; revised 18 July 2002; accepted 18 July 2002)

Although the ratio of direct-to-reverberant sound energy is known to be an important acoustic cue to sound source distance, human sensitivity to changes in this cue is largely unknown. Here, direct-to-reverberant energy discrimination thresholds were measured for six listeners using virtual sound source techniques that allow for convenient and precise control of this stimulus parameter. Four different types of source stimuli were tested: a 50 ms noise burst with abrupt onset/offset, a 300 ms duration noise burst with gradual onset/offset, a speech syllable, and an impulse. Over a range of direct-to-reverberant ratios from 0 to 20 dB, an adaptive 2AFC procedure (3-down, 1-up) was used to measure discrimination thresholds. For all stimuli, these thresholds ranged from 5 to 6 dB. A *post hoc* fitting procedure confirmed that slopes of the psychometric functions were homogeneous across stimulus conditions and listeners. These threshold results suggest that direct-to-reverberant energy ratio by itself provides only a course coding of sound source distance, because threshold values correspond to greater than 2-fold changes in physical distance for the acoustic environment under examination. [DOI: 10.1121/1.1506692]

PACS numbers: 43.66.Qp [LRB]

I. INTRODUCTION

The perception of sound source distance is thought to depend on a variety of acoustic cues, including: intensity (Gamble, 1909), frequency spectrum (Békésy, 1938; Coleman, 1968), binaural cues (Hirsch, 1968), and the ratio of direct-to-reverberant sound energy (Békésy, 1938; Bronkhorst and Houtgast, 1999; Butler *et al.*, 1980; Mershon and Bowers, 1979; Mershon and King, 1975; Nielsen, 1993). Because both the availability and reliability of these cues are dependent on the given acoustical situation, the auditory system likely combines information from multiple cues to produces stable distance percepts (Zahorik, 2002). In order to understand this cue combination process fully, it is important to determine the human auditory system's sensitivity to each of these cues in isolation. Past psychophysical results provide information on sensitivity to changes in intensity (Jesteadt *et al.*, 1977), spectral profile (Green, 1988), and interaural time and intensity differences (Blauert, 1983). Even though direct-to-reverberant energy ratio may be one of the most important distance cues for conveying a sense of environmental ''space,'' human sensitivity to this cue in isolation is largely unknown.

The direct-to-reverberant energy ratio cue, ν, results primarily from the diffuse reverberant sound-field present in environments with sound reflecting surfaces. This sound-field is a collection of, perhaps, thousands of complex reflections: degraded, delayed, and attenuated copies of the original waveform. As source distance increases, reverberant energy remains roughly constant, although direct-path energy decreases by 6 dB per doubling of source distance; hence, the direct-to-reverberant energy ratio, ν, decreases. The precise amount that ν changes with distance depends critically on the amount of reverberant energy present, which is determined by properties of the acoustic environment. For room environments, reverberant energy (as a function of time) is determined principally by the size of the room and the acoustic properties of the reflecting surfaces of the room. Many outdoor environments also produce reverberation, however, and therefore a direct-to-reverberant energy ratio cue that varies with distance (Richards and Wiley, 1980).

To date, the only experiment to explicitly examine listeners' sensitivity to direct-to-reverberant energy ratio is one conducted by Reichardt and Schmidt (1966). This work, as part of a larger goal of constructing a scale of subjective spatial impression for recorded music signals, reports that a 2 dB change in direct-to-reverberant energy ratio was just noticeable. Though pioneering, this work suffers from a number of methodological issues that may limit the ability to generalize its results to other situations, such as real sound fields or different source signals. These issues include: (1) Reverberant sound simulation: The reverberant sound was produced via an arrangement of four loudspeakers surrounding the listener in an anechoic room using a reverberation device with an initial delay of 50 ms relative to the direct-path and reverberation time of 2 s (the direct-path was simulated using two additional loudspeakers placed in a stereophonic configuration in front of the listener). It seems likely that this simulation provided a relatively poor approximation to the spatial content of reverberation in a real room environment, in which reverberation is diffuse and therefore omni-directional. (2) Direct-to-reverberant energy ratio manipulation: The ratio of direct-to-reverberant energy was nominally manipulated by varying the energy level in the direct path, although reverberant energy level was also varied concomitantly in order to keep the overall level at the listener's location constant. This method of manipulation may have allowed listeners to focus on changes in the level of either direct-path or reverberant energy alone when performing the task, and not on the ratio direct-to-reverberant energy per se. (3) Source signals: Because the experiment was only concerned with a single type of source signal (recorded music), the extent to which the results may have been

a)Electronic mail: zahorik@waisman.wisc.edu

signal-dependent is unknown. In simpler but related acoustic situations involving a sound source with a single simple reflection, or echo, detection thresholds for the echo are known to be stimulus dependent. The lowest thresholds result from brief, impulsive signals (Leshowitz, 1971) and higher thresholds result from longer duration signals with slow onsets (Cherry and Taylor, 1954; Schubert and Wernick, 1969). To the extent that the mechanisms underlying echo perception and direct-to-reverberant energy ratio perception are similar, the discriminability of direct-to-reverberant energy ratio may also differ for different types of source signals. (4) Psychophysical methods: The just-noticeable difference (jnd) for direct-to-reverberant energy ratio was defined in this experiment to be the point at which 50% of the 30 listeners reported a change in subjective spatial impression of the stimulus, using a cued yes–no procedure. This psychophysical method for determining direct-to-reverberant jnd, in addition to being susceptible to response bias, may have obscured potentially interesting and significant differences in individual listener performance.

This article further examines discrimination threshold for direct-to-reverberant energy ratio, with the goal of providing a more general estimate of threshold than can be inferred from previous work (i.e., Reichardt and Schmidt, 1966). Virtual sound source methods were used extensively because they offered a precise, convenient, and highly naturalistic means of controlling direct-to-reverberant energy ratio. These methods avoided the practically impossible task of physically manipulating the properties of an acoustic environment, such as size and surface reflectance, on a trial-by-trial basis throughout the course of the experiment: a task that would have been required if real sound sources were used. They also allowed for proper spatialization of both direct and reverberant sound-fields, as well as a means of avoiding potential confounds between the absolute levels of direct and reverberant energy and their relative levels (direct-to-reverberant ratio). To assess the signal-dependence of direct-to-reverberant energy ratio sensitivity, four different types of source signals were examined, varying in length and onset. All threshold estimates were determined on an individual-listener basis, using a modern adaptive two-alternative forced-choice (2AFC) procedure that minimized response bias. Because measurements of direct-to-reverberant energy ratio discrimination threshold have been reported so infrequently, it was also important to assess whether a single threshold value adequately described the underlying psychometric function that related the percentage of correct responses in the discrimination to direct-to-reverberant energy ratio differences. For this reason, a *post hoc* analysis for estimating psychometric functions from adaptive procedures was performed.

II. METHOD

A. Participants

Six listeners (five female and one male, ages 20–28) voluntarily participated in the experiment. All listeners had normal pure-tone thresholds ($\leqslant$15 dB HL from 200 to 4000 Hz), as verified by the ASHA standard audiometric screening procedure using a Grason–Stadler GSI-10 diagnostic audiometer, calibrated to ANSI (1989) specifications. Five of the listeners (identification codes SQW, SSC, SSH, SSJ, and SSK) had participated in other auditory distance perception experiments and were paid for their participation. The author (SLO) participated as the sixth listener.

B. Stimuli

Virtual sound sources delivered via headphones were used as stimuli throughout the experiment. These virtual sources were designed to accurately recreate a loudspeaker sound source at a distance of 1.22 m in a semi-reverberant auditorium, and were synthesized by convolution with acoustical impulse-response measurements from the ears of each participant while seated in a real auditorium facing a real loudspeaker source. This technique has been shown to produce source distance percepts that are always externalized and comparable to those observed in other studies using real sound sources (Zahorik, 2002). Direct-to-reverberant energy ratio was easily manipulated by digitally scaling the reverberant energy in measured binaural room impulse-responses (BRIRs). Because both the methods used to measure BRIRs and to construct virtual source stimuli were identical to those described in greater detail elsewhere (Zahorik, 2002; Zahorik and Wightman, 2001), only brief description of these methods is provided here.

1. Virtual sound source technique

The technique of constructing virtual sound source stimuli used here requires three sources of information: BRIR measurements, headphone equalization, and the signals to be spatialized. It is fundamentally similar to the techniques used to simulate direction using head-related impulse-responses (which are equivalent to head-related transfer functions, or HRTFs) measured with a blocked-meatus (Møller *et al.*, 1995), except that the measurement period for the BRIRs is much longer in order to fully capture all of the acoustical effects of the measurement environment in addition to those effects related to the head and ears.

BRIR measurements were collected from five of the six listeners in a semi-reverberant auditorium (broadband reverberation time, $T60 \approx 0.7$ s). Twelve logarithmically spaced distances ranging from 0.3 to 13.8 m directly in front of the listener's position (0° azimuth) were measured, using a small full-range loudspeaker (Realistic Minimus 3.5) as the sound source, and Sennheiser microphones (KE4-211-2) in a blocked-meatus configuration. A maximum-length sequence (MLS) technique (Rife and Vanderkooy, 1989) was used to measure all impulse-responses, and was implemented using Tucker-Davis Technologies (TDT) hardware for signal manipulation and D/A–A/D conversion (50 kHz sampling rate, 16-bit quantization). Although only the 1.22 m measurement distance was used to construct the virtual sound source stimuli used in the experiment, measurements from 11 other distances were collected in order to quantify the natural variation of direct-to-reverberant energy ratio with source distance.

During the same measurement procedure, impulse-responses of Beyerdynamic DT 990 Pro headphones when

coupled to the head (with a blocked-meatus) were measured for five of the six listeners. These measurements were used to derive individualized headphone equalization filters used to generate virtual sound source stimuli. Participants were blindfolded during the entire measurement procedure, which lasted approximately 30 min. Given that listener SLO did not participate in the measurement procedure, this listener was presented with stimuli constructed from BRIRs and headphone equalization filters using SSJ's ears during the experiment.

Four different types of source signals were used in the experiment: a short duration noise burst, a long duration noise burst, a speech syllable, and an impulse. Both noise bursts were Gaussian with cosine-squared gating functions. The short-duration burst lasted 50 ms with 2.5 ms of onset/offset time. The long-duration burst lasted 300 ms with 150 ms of onset/offset time. The speech sample was the syllable /da/, recorded digitally from a male speaker in an anechoic room. It had a duration of approximately 300 ms. For the impulse stimulus, listeners were simply presented the manipulated BRIR, because this is formally equivalent to convolving the BRIRs with an ideal impulse.

Virtual sound source stimuli were constructed by convolving a given individualized BRIR, an individualized headphone equalization filter impulse-response, and the desired source stimulus. Stimuli were presented using Beyerdynamic (DT 990 Pro) headphones and a Tucker-Davis Technologies PD1 for D/A conversion (50 kHz, 16-bit) and headphone equalization via real-time DSP. All other stimulus manipulation (direct-to-reverberant energy scaling, BRIR convolution with source stimuli) was implemented off-line, stored to computer files, and then accessed at the appropriate times during the experimental procedure.

2. Direct-to-reverberant energy ratio manipulation

Direct-to-reverberant energy ratio, ν, was manipulated digitally by setting the energy in the direct-path portion of the 1.22 m BRIRs to a constant value (the nominal measured values for right and left ears, respectively) and scaling the reverberant energy portion of the BRIRs. The direct-path portion of the BRIRs was distinguished from the reverberant portion by a simple time-windowing procedure, which defined the direct-path to be 2.5 ms in duration (the approximate duration of anechoically measured head-related impulse-responses (Møller *et al.*, 1995), starting from the first portion of the BRIR that was 20 dB above the noise-floor in the impulse-response. All portions of the BRIR beyond this direct-path time point where considered reverberant, which included the first reflection to reach the listener (the floor reflection) with delay of approximately 4.5 ms relative to the direct-path. Figure 1 displays a representative BRIR (left ear only) modified such that ν was either 0, 10, or 20 dB. When $\nu=0$ dB, direct and reverberant energy are equal. Positive values of ν indicate that direct energy is greater than reverberant energy. Figure 2 displays the natural variation of ν with changes in source distance, r, for the environment in which the BRIRs were measured. Mean ν values based on BRIRs from the left (open symbols) and right (solid symbols) ears of five participants are shown. The

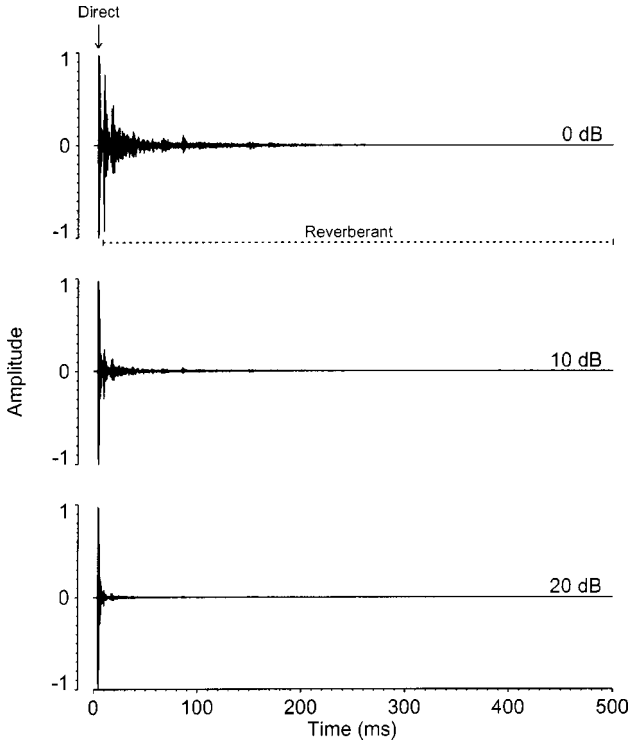


FIG. 1. Binaural room impulse-responses (BRIRs) for three different manipulated direct-to-reverberant energy rations (0, 10, 20 dB) from listener SSK (left ear) at a sound source distance of 1.22 m. Regions of direct and reverberant energy are indicated.

1.22 m BRIR, which was used to generate the stimuli used in the experiment, had a natural direct-to-reverberant energy ratio of approximately +9 dB. The data are well approximated by a single linear function of the form: $\nu(\mathrm{dB}) = -3.64\log_2(r)+10.76$. This implies that the referent levels of ν (0, 10, or 20 dB) would correspond to changes in distance by approximately a factor of 7.

The method of varying ν by reverberant energy scaling results in changes to the overall level of the stimulus. Figure 3 displays the average changes in overall level as a function

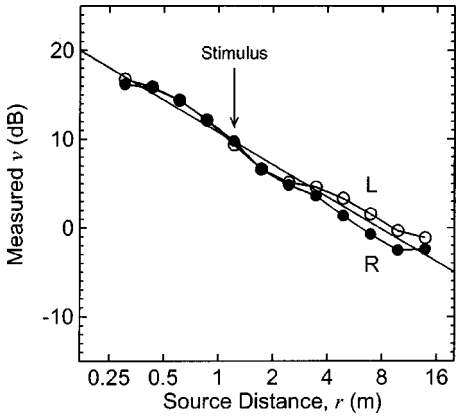


FIG. 2. Natural direct-to-reverberant energy ratio, ν, as a function of source distance, r, in the measurement environment. Solid symbols show mean results for the right ears (R), and open symbols results for the left ears (L) based on BRIRs from five participants. Standard errors of the mean ν values were all <0.3 dB. The data are well approximated by a single linear function of the form: $\nu(\mathrm{dB}) = -3.64\log_2(r)+10.76$. The 1.22 m source distance was used to construct all virtual sound source stimuli for the experiment.

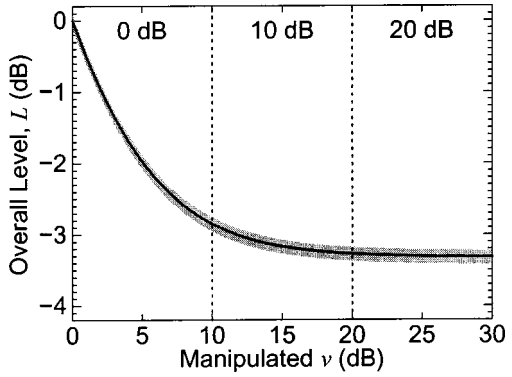


FIG. 3. Mean overall level ($n=5$) computed from the manipulated BRIRs (see Fig. 1) as a function of manipulated direct-to-reverberant energy ratio, ν. The shaded region displays a 95% confidence interval. Ranges of ν variation for each of the three standard ν levels (0, 10, and 20 dB) are also indicated. The greatest variation in level occurs within the range of the 0 dB standard, approximately 2.8 dB.

of manipulated ν value based on the five sets of BRIR measurements from individual listeners. The shaded regions of the curve display the 95% confidence limits. To minimize the effects of this confound, a 20 dB random rove (± 10 dB uniform probability-density distribution) in overall level (nominally 68 dB SPL) was applied to every stimulus presentation.

C. Procedure

An adaptive two-alternative forced choice (2AFC) procedure was used to measure ν discrimination thresholds at three different levels of ν (0, 10, and 20 dB) for each of the four stimulus types. The inter-stimulus interval (ISI) was 500 ms. Listeners were instructed to choose the interval with greater reverberation, and were provided with feedback as to a response's correctness on every trial via an auditory signal. Responses were made while seated in an IAC double-walled sound attenuating chamber using a computer-linked button box with eyes closed. Each listener completed 6 experimental runs per level/stimulus combination. Odd numbered runs started with a level of $\Delta\nu=10$ dB (i.e., 10 dB greater than the given referent level, ν, of either 0, 10, or 20 dB). Even numbered runs started with $\Delta\nu=1$ dB, which was the minimum value of $\Delta\nu$. The adaptive procedure (3-down, 1-up) decremented $\Delta\nu$ by 1 dB after three correct responses and incremented $\Delta\nu$ by 1 dB after one incorrect response, which yields threshold estimates at the 79.4% correct level (Levitt, 1970). The step size of 1 dB was fixed for all trials. Runs were terminated after 20 reversals in the adaptive procedure. Four 2-h sessions were required for each listener to complete the experiment, one session for each stimulus type.

III. RESULTS

Thresholds were determined by averaging the $\Delta\nu$ presented during the final five adaptive procedure reversals. Figure 4 displays this value of $\Delta\nu$ in decibels averaged across listeners ($n=6$) as a function of referent ν value for each of the four stimulus types. The fact that $\Delta\nu$ expressed in decibels is roughly constant (5–6 dB) across varying levels of ν in decibels implies that direct-to-reverberant energy ratio discrimination obeys Weber's law. It should be noticed that threshold values are also approximately constant (5–6 dB) across stimulus type. A greater degree of threshold variability across listeners is observed for the 300 ms noise burst than for the other three stimulus types, however.

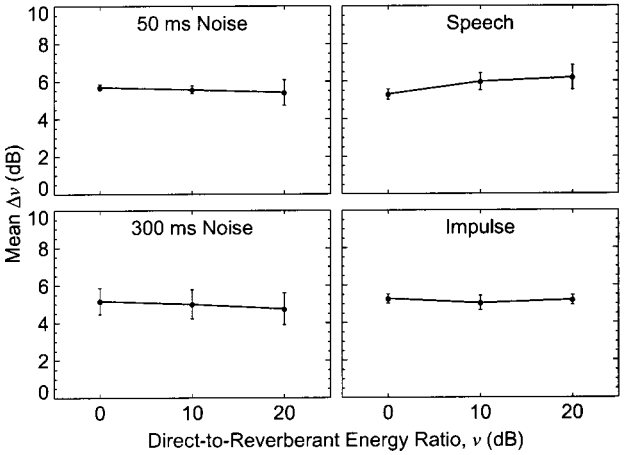


FIG. 4. Mean ($n=6$) direct-to-reverberant energy ratio discrimination thresholds, $\Delta\nu$, as a function of direct-to-reverberant energy ratio, ν, for four difference stimulus types (50 ms Noise, Speech, 300 ms Noise, and Impulse). Bars represent one standard error of the mean.

To evaluate the effects of the 20 dB random rove in overall level on estimates of $\Delta\nu$, these estimates were transformed to their corresponding changes in overall level (in the absence of the random level rove) for each value of ν, based on the relation shown in Fig. 3. The results of this transformation are displayed in Fig. 5. The theoretical lower-bound for thresholds based only on level information is 7.15 dB, given the 20 dB rove in level and the 3-down, 1-up adaptive procedure that was implemented (Green, 1988). Because all threshold estimates shown in Fig. 5 are significantly below this value, $t(71)=59.64$, $p<0.001$, one may conclude that overall level did not mediate performance in this task.

Thresholds were also nearly constant over the course of experiment, as shown in Fig. 6 by the similar threshold values obtained for each experimental run. Results from Fig. 6

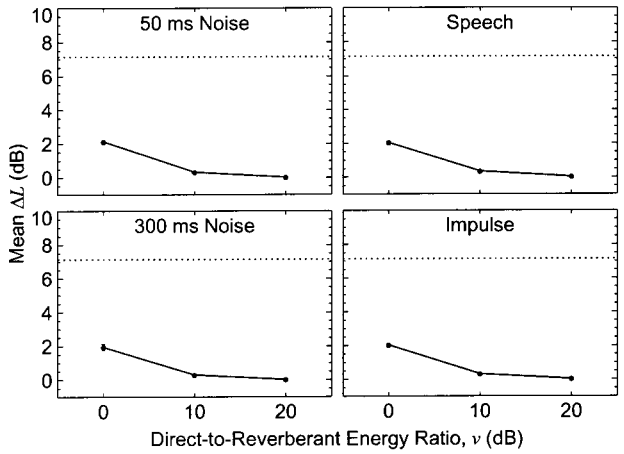


FIG. 5. Mean ($n=6$) direct-to-reverberant energy ratio discrimination thresholds expressed in terms of overall level change (see Fig. 3), as a function of direct-to-reverberant energy ratio, ν, for four difference stimulus types (50 ms Noise, Speech, 300 ms Noise, and Impulse). Bars represent one standard error of the mean (all <0.2 dB).

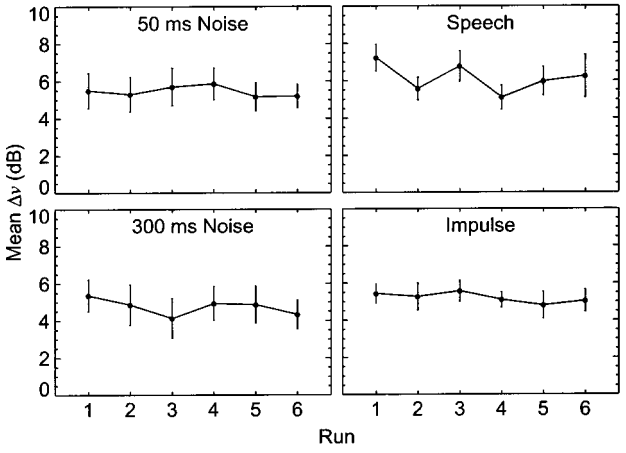


FIG. 6. Mean ($n=6$) direct-to-reverberant energy ratio discrimination thresholds, $\Delta\nu$, as a function of experimental run, for four difference stimulus types (50 ms Noise, Speech, 300 ms Noise, and Impulse). Bars represent one standard error of the mean.

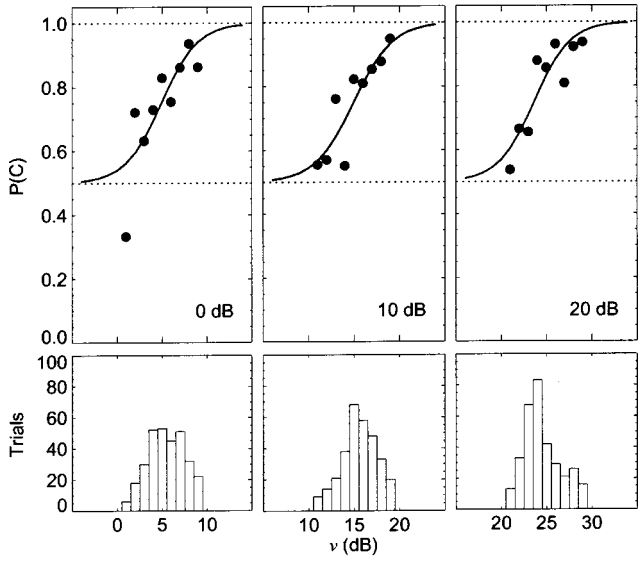


FIG. 7. Estimated psychometric functions (logistic fits) for a single representative listener (SSH) and the 50 ms Noise stimulus. Separate functions are shown for each referent level of ν (0, 10 and 20 dB). The lower panels display the number of response trials that compose each point on the psychometric functions.

also suggest that the stimulus starting levels, which alternated between $\Delta\nu=1$ dB and $\Delta\nu=10$ dB for even and odd numbered runs, did not systematically affect threshold estimates in and adaptive procedure. Further, the adaptive tracks for even and odd numbered runs were found to converge on threshold estimates in a similar manner. The mean number of trials to convergence (20 reversals) was 65.4 ($\sigma=10.3$) for even-number runs, and 50.6 ($\sigma=8.7$) for odd-numbered runs.

Given that the adaptive tracks were alternately started at stimulus levels near the 50% and 100% correct performance marks, a sufficient amount of data was available to estimate psychometric functions for individual listeners. The minimum number of trials per stimulus level was approximately ten on average, and guaranteed to be at least three, given the alternating track start points. The majority of trials were centered about the 79.4% correct point given the 3-down, 1-up adaptation scheme, with 300 to 400 total trials available for each listener at every combination of stimulus type and referent level. Leek *et al.* (1992) claim that unbiased estimates of psychometric functions from adaptive track data may be obtained with as few as 200 trials for the 3-down, 1-up scheme.

To estimate psychometric functions, the data were fit with logistic functions via a nonlinear regression procedure. This function has the following form:

$$\text{Logistic}(x)=\frac{0.5}{1+e^{-2((x-A)/B)}}+0.5,$$

where A and B are intercept and slope parameters. The intercept parameter, A, represents the value of ν in decibels that corresponds to 75% correct. The slope parameter, B, which is analogous to σ^2, is a rate of change over a fixed linear portion of the curve, also expressed in decibels.

Each fit was evaluated using a likelihood-ratio Chi-square test (Agresti, 1990), testing the null hypothesis that observed and fitted data were drawn from the same population. Computation of the likelihood-ratio Chi-square test statistic, G^2, was accomplished with the following formula:

$$G^2=2\sum\left(O\log_e\left(\frac{O}{P}\right)\right),$$

where O is a vector of observed percent correct values, and P is a vector of fitted percent correct values. Because it was desirable to reject any poor fits to the data, very liberal tests were performed, with α set at the 0.1 level. No rejections were encountered. Therefore, all logistic functions used are considered good fits to the data, and as a result good psychometric function estimates. Figure 7 displays an example of the estimated psychometric functions for the 50 ms noise stimulus (separate functions for each of the three referent values of ν) for a single representative listener, along with histograms showing the number of trials presented at each ν value. Similar fits were computed for all remaining listeners and stimulus types.

Of particular interest is the fitted function slope parameter, because the consistency of this parameter across different stimulus conditions can be taken as evidence that a single threshold value adequately describes the underlying psychometric function that relates percentage of correct responses in the discrimination to direct-to-reverberant energy ratio differences. Figure 8 displays mean slopes values for the fitted psychometric functions as a function of both stimulus type and referent ν. The slope values are relatively constant across both stimulus type and ν level. Individual variability in slope estimates is also quite low, with standard errors generally less than 1.5 dB. Taken together, these results suggest that the psychometric functions may be accurately characterized by their threshold values alone.

The threshold values for listener SLO, who was presented virtual sound sources generated from nonindividualized measurements (BRIRs and headphone equalization filters using SSJ's ears), were not significantly greater than those of the other remaining listeners, $t(70)=-3.17$, p

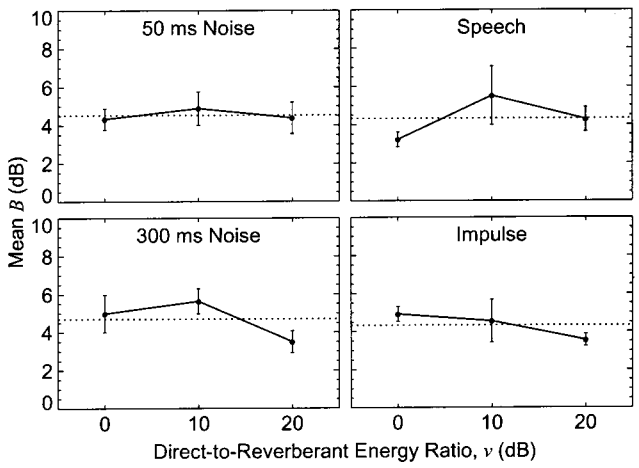


FIG. 8. Mean ($n=6$) slope value, B, of the fitted psychometric functions as a function of direct-to-reverberant energy ratio, ν, for four difference stimulus types (50 ms Noise, Speech, 300 ms Noise, and Impulse). Bars represent one standard error of the mean. The dashed line displays the mean slope for each stimulus type.

>0.99. This suggests that the use of nonindividualized measurements for virtual source synthesis does not decrease performance in this situation, as in directional localization tasks where accuracy has be found to decrease markedly under similar manipulations (Wenzel *et al.*, 1993). Estimated psychometric function slopes for listener SLO also did not differ significantly from the remaining listeners, $t(70)=0.81$, $p=0.39$.

Given the relation displayed in Fig. 2 for the acoustic environment under test, it is possible to map just-noticeable changes in direct-to-reverberant energy ratio to estimates of just-noticeable changes in source distance. Hence, if listeners use only ν to discriminate source distance, a just-noticeable change in ν of 5 dB corresponds to a change in source distance by a factor of 2.59. This suggests that direct-to-reverberant energy provides only a coarse coding of relative sound source distance.

IV. DISCUSSION

Although the experiment reported in this article has examined only a single acoustic environment, it has shown that discrimination thresholds for direct-to-reverberant energy ratio, ν, within this environment are 5–6 dB over the range of energy ratios examined (0–20 dB), and that similar thresholds were obtained for all types of source signals tested. These results warrant discussion of a number of issues, including their relation to past results, implications for distance perception, and potential insights regarding mechanisms of processing.

A. Relation to past results

Reichardt and Schmidt (1966) report discrimination thresholds for direct-to-reverberant energy ratio that are substantially smaller than those observed in the present experiments (2 dB versus 5–6 dB). It is argued that these past threshold estimates may have been spuriously low as a result of the methods used to manipulate the ratio of direct-to-reverberant energy. By holding overall stimulus level constant in that experiment, the changes in the ratio of direct-to-reverberant energy were correlated with changes in the absolute levels of both direct and reverberant sound fields. Hence, listeners could likely have performed the task by comparing only the level of the direct sound or only the level of the reverberation in the two stimulus presentations, and not the ratio of direct and reverberant energy levels per se. The relatively long reverberation time of 2 s may have facilitated listening strategies that focused specifically on the level of the ending portion of the stimulus, a portion that contained only reverberant energy. The ability to segregate this ending reverberant portion of the stimulus from the beginning portions that also contained direct-path energy would have been well within known temporal acuity limits of the auditory system (Viemeister and Plack, 1993), and may have been further facilitated by the spatial separation of the reverberant energy from the direct-path energy resulting from the arrangement of loudspeakers in the experiment. The fact that the reported thresholds were roughly similar to known level discrimination thresholds (Jesteadt *et al.*, 1977) lends further support for a listening strategy in which overall level of either direct-path energy, or more likely reverberant energy was compared in the two presentation intervals. This stimulus situation stands in contrast to that of the current experiments, where the absolute levels of either direct or reverberant sound fields were uncorrelated with their energy ratio by virtue of the randomization of overall stimulus level.

B. Distance perception implications

A variety of past results suggests that direct-to-reverberant energy ratio can be a salient cue to sound source distance (Békésy, 1938; Bronkhorst and Houtgast, 1999; Butler *et al.*, 1980; Mershon and Bowers, 1979; Mershon and King, 1975; Nielsen, 1993; Zahorik, 2001, 2002). At first glance, the discrimination thresholds reported in this article might be interpreted as undermining these past results, because they imply that listeners can only detect distance changes greater than a factor of 2 using the ν cue in the acoustic environment under examination. How could such an imprecise distance cue be considered salient? It is argued that imprecision is one of the hallmarks of auditory distance perception, and that this imprecision is reflected in both judgment variability of past works and in the relatively large thresholds for discriminating direct-to-reverberant energy ratio observed here.

To examine this issue more closely, a systematic comparison between the current discrimination threshold results and the results from past studies is essential. This comparison is not entirely straightforward, however, as none of the past studies reported distance discrimination thresholds directly. Most were concerned primarily with determining scales of perceived distance. As a result, distance discrimination thresholds must be estimated from these data, a procedure that requires information about both the mean apparent distance judgment for a given source distance and measures of judgment variability within individual listeners. Unfortunately, the only study to provide sufficient information regarding these two parameters is that of Zahorik

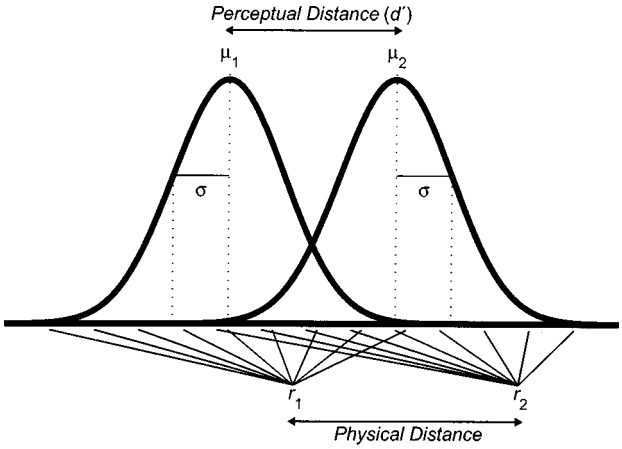


FIG. 9. A conceptual framework in which physical distances r_1 and r_2 create distributions (Gaussian with means μ_1, μ_2, and standard deviation σ) of activation along a theoretical continuum. Perceptual distance, d', is defined as $(\mu_2-\mu_1)/\sigma$.

(2002). The results from this study are particularly applicable, however, because they include data collected from the same listeners under stimulus conditions that are in many ways identical to those of the current experiments. Virtual source synthesis from the same set of binaural room impulse-responses was implemented for physical distances ranging from 0.3 to 13.8 m, which resulted in stimulus conditions in which direct-to-reverberant energy ratio was a primary distance cue, although other cues, such as intensity, were also available to listeners.

In order to estimate distance discrimination thresholds from the psychophysical scaling data of Zahorik (2002), a conceptual framework that relates physical distance to perceived distance must first be considered. Suppose that two sound sources are located at different physical distances from a listener, each of which activates a region along a hypothetical continuum of perceived distance. These regions may be described both by their mean location along the perceptual continuum, and their spread of activation. Figure 9 shows a graphical representation of this framework, assuming that the activation functions corresponding to physical distances r_1 and r_2 may be described by Gaussian distributions, each with a mean (μ_1,μ_2) and a standard deviation (σ). To the extent that apparent distance responses can be viewed as reflecting source distance percepts, analysis of the data from Zahorik (2002) validates this assumption. The distributions of the log-transformed apparent distance responses did not differ significantly from Gaussian distributions for all listeners and source distances, as verified by Kolmogorov–Smirnov tests. The average standard deviation of the responses, σ, was $0.35\log_2$ m, which was relatively stable as a function of source distance, but ranged from approximately 0.2 to $0.6\log_2$ m between listeners. The next step in the estimation procedure is to select a perceptual distance, d', that will be stipulated to correspond to the discrimination threshold. Here we select a value of 1.16 for d' because it is equivalent to 79.4% correct in a 2AFC task (Macmillan and Creelman, 1991), and will therefore facilitate comparison to the direct-to-reverberant ratio discrimination results reported in this article. Given that $d'=(\mu_2-\mu_1)/\sigma$ and a value of 0.35 is a good approximation of σ, then $1.16=(\mu_2-\mu_1)/0.35$ in this setting, which implies $\mu_2-\mu_1=0.41\log_2$ m. This difference must now be mapped to a difference in physical distance. Zahorik (2002) reports that the psychophysical function that relates perceived distance (μ) to physical source distance (r) is a compressive linear function on logarithmically spaced coordinates, with an average slope of approximately 0.4. Therefore, the estimated discrimination threshold, r_2-r_1, is $0.41\log_2$ m$/0.4=1.03\log_2$ m. Hence, this analysis suggests that listeners would likely not have been able to detect source distance changes less than approximately a factor of 2. This result is in relatively close agreement with the source distance changes associated with the direct-to-reverberant energy discrimination thresholds reported here, which were roughly a factor of 2.59. This level of agreement is perhaps especially compelling, given that the former factor of 2 estimate was based on stimulus conditions in which distance cues in addition to the direct-to-reverberant energy ratio cue were available to the listeners (Zahorik, 2002).

From this analysis, it is clear that perceptual variability is at the heart of the source distance discrimination problem where direct-to-reverberant energy ratio is a primary distance cue. Although other studies have not directly reported measures of individual response variability, many do report relatively high levels of response variability pooled across listeners (Bronkhorst and Houtgast, 1999; Mershon and Bowers, 1979; Mershon and King, 1975; Nielsen, 1993; Zahorik, 2001), which could be consistent with high perceptual variability within individuals. These results lend further support to the claim that distance percepts elicited by the direct-to-reverberant energy ratio cue are relatively imprecise, and that this imprecision is reflected both in the individual response variability in judgments of apparent source distance, as well as in the relatively large thresholds for discriminating direct-to-reverberant energy ratio. These studies also report compression in perceived distance similar to that observed by Zahorik (2002). Because of the central role of perceptual variability, the results reported here are not directly applicable to a recently proposed model of distance perception based on the ν cue (Bronkhorst and Houtgast, 1999), given that this model only predicts mean distance responses, with no provision for variability.

Additional evidence suggests that acoustic cues other than direct-to-reverberant energy ratio can provide much finer resolution of source distance changes. When intensity, for example, varies systematically with distance, just-noticeable distance changes are known to be quite small, ranging from factors of approximately 0.03 to 0.33 (Ashmead *et al.*, 1990; Simpson and Stanton, 1973; Strybel and Perrott, 1984). This suggests that when intensity or other distance cues in addition to direct-to-reverberant energy ratio are available to listeners, the processes subserving distance perception may rely on these more precise sources of information for tasks involving relative distance judgment. Direct-to-reverberant ratio may be critically important for providing absolute distance information (Mershon and King, 1975), however, by potentially supplying information about the acoustic power of the sound source (Zahorik and Wightman, 2001). This perhaps explains why listeners are known

to process information from multiple distance cues under a variety of circumstances (Zahorik, 2002).

C. Evidence against a temporal process

Direct-to-reverberant ratio discrimination thresholds were found to be effectively the same for stimuli with gradual onsets/offsets (the 300 ms noise) as for stimuli with sharp onsets/offsets (the impulse and 50 ms noise). This result may provide preliminary insight into the nature of the processes underlying direct-to-reverberant energy ratio perception. Perhaps the most basic process for direct-to-reverberant energy ratio perception involves simply computing acoustic energy at two times—at direct-path arrival and again later during the reverberant decay—and taking the ratio of the results. Although this type of process appears sufficient to explain performance for stimuli with abrupt onsets and offsets, it clearly cannot account for situations in which the stimuli are continuous or have very gradual onsets and offsets, because in such cases it is impossible to temporally separate direct-path energy from reverberant energy. The fact that direct-to-reverberant ratio discrimination thresholds were effectively the same for all stimuli, gradual onsets/offsets and sharp onsets/offsets alike, suggests that such a hypothesis, based on temporal energy, is invalid. This result is consistent with similar results obtained by Bronkhorst (2001) in a distance scaling experiment. Such a result could be explained if it were instead the corresponding frequency-domain changes, or perhaps changes in spatial qualities of the sound (Bronkhorst, 2001) caused by direct-to-reverberant energy ratio variation that were perceptually salient.

ACKNOWLEDGMENTS

I would like to thank Fred Wightman, Doris Kistler, and Adelbert Bronkhorst for their invaluable discussions and suggestions, Kathy Hoopes for her assistance in behavioral data collection, and Bob Lutfi, Rick Jenison, Jim Dannemiller, and John Brugge for providing helpful comments on an earlier version of this work. Financial support was provided by the NIH-NIDCD (P01 DC00116-23), ONR (N00014096-1-0675), NASA (Cooperative Agreement #NCC2-542), and NIH-NEI (1F32EY07010-01).

Agresti, A. (**1990**). *Categorical Data Analysis* (Wiley, New York).

ANSI (**1989**). ANSI S3.9-1989, ''American National Standard specification for audiometers'' (American National Standards Institute, New York).

Ashmead, D. H., LeRoy, D., and Odom, R. D. (**1990**). ''Perception of the relative distances of nearby sound sources,'' Percept. Psychophys. **47**, 326–331.

Békésy, G. v. (**1938**). ''Über die Entstehung der Entfernungsempfindung beim Hören (On the origin of distance perception in hearing),'' Akust. Z. **3**, 21–31.

Blauert, J. (**1983**). *Spatial Hearing* (MIT Press, Cambridge, MA).

Bronkhorst, A. W. (**2001**). ''Effects of stimulus properties on auditory distance perception in rooms,'' in *Proceedings of the 12th International Symposium on Hearing (ISH): Physiological and Psychological Bases of Auditory Function* (Shaker, Mastricht, The Netherlands), pp. 184–191.

Bronkhorst, A. W., and Houtgast, T. (**1999**). ''Auditory distance perception in rooms,'' Nature (London) **397**, 517–520.

Butler, R. A., Levy, E. T., and Neff, W. D. (**1980**). ''Apparent distance of sounds recorded in echoic and anechoic chambers,'' J. Exp. Psychol. Hum. Percept. Perform. **6**, 745–750.

Cherry, E. C., and Taylor, W. I. (**1954**). ''Some further experiments upon the recognition of speech with one and two ears,'' J. Acoust. Soc. Am. **26**, 554–559.

Coleman, P. D. (**1968**). ''Dual role of frequency spectrum in determination of auditory distance,'' J. Acoust. Soc. Am. **44**, 631–634.

Gamble, E. A. (**1909**). ''Intensity as a criterion in estimating the distance of sounds,'' Psychol. Rev. **16**, 416–426.

Green, D. M. (**1988**). *Profile Analysis. Auditory Intensity Discrimination* (Oxford University, New York), pp. 19–21.

Hirsch, H. R. (**1968**). ''Perception of the range of a sound source of unknown strength,'' J. Acoust. Soc. Am. **43**, 373–374.

Jesteadt, W., Wier, C. C., and Green, D. M. (**1977**). ''Intensity discrimination as a function of frequency and sensation level,'' J. Acoust. Soc. Am. **61**, 169–177.

Leek, M. R., Hanna, T. E., and Marshall, L. (**1992**). ''Estimation of psychometric functions from adaptive tracking procedures,'' Percept. Psychophys. **51**, 247–256.

Leshowitz, B. (**1971**). ''Measurement of the two-click threshold,'' J. Acoust. Soc. Am. **49**, 462–466.

Levitt, H. (**1970**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Macmillan, N. A., and Creelman, C. D. (**1991**). *Detection Theory: A User's Guide* (Cambridge University Press, Cambridge).

Mershon, D. H., and Bowers, J. N. (**1979**). ''Absolute and relative cues for the auditory perception of egocentric distance,'' Perception **8**, 311–322.

Mershon, D. H., and King, E. (**1975**). ''Intensity and reverberation as factors in the auditory perception of egocentric distance,'' Percept. Psychophys. **18**, 409–415.

Møller, H., Sørensen, M. F., Hammershøi, D., and Jensen, C. B. (**1995**). ''Head-related transfer functions of human subjects,'' J. Audio Eng. Soc. **43**, 300–321.

Nielsen, S. H. (**1993**). ''Auditory distance perception in different rooms,'' J. Audio Eng. Soc. **41**, 755–770.

Reichardt, v. W., and Schmidt, W. (**1966**). ''Die hörbareb Stufen des Raumeindruckes bei Musik (The audible steps of spatial impression in music performances),'' Acustica **17**, 175–179.

Richards, D. G., and Wiley, R. H. (**1980**). ''Reverberations and amplitude fluctuations in the propagations of sound in a forest: Implications for animal communication,'' Am. Nat. **115**, 381–399.

Rife, D. D., and Vanderkooy, J. (**1989**). ''Transfer-function measurement with maximum-length sequences,'' J. Audio Eng. Soc. **37**, 419–444.

Schubert, E. D., and Wernick, J. (**1969**). ''Envelope versus microstructure in the fusion of dichotic signals,'' J. Acoust. Soc. Am. **45**, 1525–1531.

Simpson, W. E., and Stanton, L. D. (**1973**). ''Head movement does not facilitate perception of the distance of a source of sound,'' Am. J. Psychol. **86**, 151–159.

Strybel, T. Z., and Perrott, D. R. (**1984**). ''Discrimination of relative distance in the auditory modality: The success and failure of the loudness discrimination hypothesis,'' J. Acoust. Soc. Am. **76**, 318–320.

Viemeister, N. F., and Plack, C. J. (**1993**). ''Time analysis,'' in *Human Psychophysics*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer, New York), pp. 116–154.

Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (**1993**). ''Localization using nonindividualized head-related transfer functions,'' J. Acoust. Soc. Am. **94**, 111–123.

Zahorik, P. (**2001**). ''Estimating sound source distance with and without vision,'' Optom. Vision Sci. **78**, 270–275.

Zahorik, P. (**2002**). ''Assessing auditory distance perception using virtual acoustics,'' J. Acoust. Soc. Am. **111**, 1832–1846.

Zahorik, P., and Wightman, F. L. (**2001**). ''Loudness constancy with varying sound source distance,'' Nat. Neurosci. **4**, 78–83.

A broadband tactile array on the fingertip

Ian R. Summers and Craig M. Chanter
Biomedical Physics Group, School of Physics, University of Exeter, Exeter EX4 4QL, United Kingdom

(Received 11 December 2001; revised 3 August 2002; accepted 6 August 2002)

A stimulator array is described which can deliver a wide range of displacement waveforms from each contactor, allowing vibratory stimuli to be targeted towards different populations of mechanoreceptors in the skin. The array has a working bandwidth of 20–400 Hz and 100 moving contactors covering an area of 1 cm^2 on the fingertip. The array was validated with two experiments on the perception of moving vibratory targets within a uniform background vibration. In the first experiment, with target and background at the same frequency, equivalent discrimination of target movement was obtained at higher values of target/background amplitude ratio for 40-Hz stimuli than for 320-Hz stimuli. In the second experiment, discrimination of target movement within a different-frequency background (320-Hz target and 40-Hz background, or vice versa) was found to be much easier than within a same-frequency background. These results suggest that tactile spatial acuity is better at 320 Hz than 40 Hz and that it is possible to target different receptor populations in the skin by using these frequencies. However, there are problems with this interpretation: on the basis of characterization of touch receptors in previous studies, spatial acuity is expected to be worse at 320 Hz than at 40 Hz. [DOI: 10.1121/1.1510140]

PACS numbers: 43.66.Wv, 43.38.Fx, 43.38.Hz, 43.40.Yq [MRL]

I. INTRODUCTION

In this paper we describe a stimulator array whose broadband capability offers new possibilities in the area of tactile psychophysics: i.e., possibilities for experiments in which different stimulation frequencies may be used to investigate the various aspects of spatiotemporal perception which are mediated by different populations of mechanoreceptors. We also describe two experiments of this type which provide novel psychophysics data and hence demonstrate that the design goals for the array have been satisfied. In a related investigation (not described here), successful operation of the present array has been demonstrated within the magnet bore of a magnetic-resonance imager, at static magnetic fields of 2.0 and 3.0 T. There are thus possibilities for using this type of stimulator array in the context of functional magnetic-resonance imaging, as a complementary line of investigation to psychophysics-based studies.

In recent years there has been increased interest in the sense of touch, largely because of the possibilities which have become apparent for novel forms of human-machine interaction (Srinivasan and Basdogan, 1997; Raj *et al.*, 1998; Lederman and Klatsky, 1999; Tan *et al.*, 2000; van Erp and van Veen, 2001). In order to evoke virtual touch sensations that mimic ''natural'' touch, a two-dimensional array of contactors may be used to produce spatiotemporal patterns of mechanical disturbance at the skin surface (e.g., over the fingertip). Spatial distribution of touch stimulation is important in relation to texture representation and in relation to defining the edges and corners of objects—this is beyond the scope of the single-contactor, force-feedback systems which form the basis of most virtual-touch devices.

Glabrous (smooth) skin contains various populations of mechanoreceptors: pacinian receptors and three types of nonpacinian receptor (Johnson *et al.*, 2000). These differ in terms of their frequency response (Bolanowski *et al.*, 1988; Gescheider *et al.*, 2001); in order to evoke a wide range of ''realistic'' touch sensations an array must operate over a significant fraction of the tactile frequency range of, say, 10 to 500 Hz. (An array with such broadband capability is not limited to sine-wave stimulation: each contactor can deliver a wide variety of stimulus waveforms within the working bandwidth of the system.) Vibratory stimulation in the upper part of this frequency range (approximately 100–500 Hz) is expected to stimulate pacinian receptors predominantly. Stimulation at lower frequencies is expected to stimulate nonpacinian receptors predominantly.

High-density tactile arrays have been available since the 1970s. However, in previous devices such as the widely used Optacon (Bliss *et al.*, 1970), the stimulus waveform from each contactor has generally been limited to a pulse train of fixed repetition rate ($\sim$250 Hz) or a sine wave of fixed frequency ($\sim$250 Hz, similarly). Stimulators with this restricted frequency range are expected to excite pacinian receptors predominantly. [Szaniszlo *et al.* (1998) describe a sophisticated array with 288 contactors over an area of 400 mm^2, with a fixed stimulation frequency of 220 Hz. This was used to stimulate skin on the face, where pacinian receptors are absent.] A further limitation of many previous devices is that their output is available at a fixed level only, with spatiotemporal patterns of stimulation produced by simple on/off keying (Craig, 1980, 2000; Craig and Xu, 1990).

Consequently, there have been only a small number of psychophysical studies involving broadband arrays which produce vibratory stimulation over a sufficiently wide frequency range to allow the targeting of both pacinian and nonpacinian receptors. Van Doren *et al.* (1987) describe measurements of detection thresholds for wave stimuli as a function of frequency and wavelength on an array, working bandwidth 1–256 Hz, which allowed spatial variation in one

dimension only (with linear contactors at 0.38-mm spacing). Both Hasser and Weisenberger (1993) and Taylor *et al.* (1998) describe arrays based on shape-memory-alloy actuators which operate over a range of stimulation frequencies. However, waveform control in such devices is difficult because of the inherent on/off nature of the actuator mechanism. In addition, operation at stimulation frequencies above 100 Hz is problematical because of the inherent time constant associated with heating/cooling of the shape-memory alloy—Hasser and Weisenberger achieved operation at frequencies up to 200 Hz by designing for only partial relaxation of the system between heating pulses.

In the mid 1990s James C. Craig (Indiana University) and Kenneth O. Johnson (Johns Hopkins University) developed a high-density array with 400 contactors over an area of 1 cm^2 and a design bandwidth of 0–500 Hz (Pawluk *et al.*, 1998). Each contactor in this array is controlled by a purpose-built linear motor (Schneider, 1988), allowing very precise specification of the displacement output. When performing to its design specification, this device offers considerable scope for novel psychophysics and neurophysiology investigations, with independent targeting of both pacinian and nonpacinian receptors. However, as yet there are no published papers with data obtained from this device, although preliminary results have been obtained (Craig, personal communication, 2002) and some work has been completed with a cut-down development version of the array (Vega-Bermudez and Johnson, 1999).

Because of the limitations in stimulus delivery outlined above, previous studies have left many unresolved issues relating to the production of virtual touch sensations. In particular, the role of different receptor populations in the skin is not clear in this context, although significant progress has been made in understanding their corresponding role in the context of ''natural'' touch (Johnson *et al.*, 2000). There is a requirement for psychophysics investigations to address these issues and hence a need for improved stimulator hardware to facilitate such investigations. The stimulator array described below is intended to meet this need, and the two experiments described below demonstrate the array's potential for obtaining psychophysics data in the research area of interest. [This research area can be considered as either (i) an extension of established single-contactor, variable-frequency techniques into the multiple-contactor context; (ii) an extension of previous studies using arrays such as the Optacon into investigations involving more flexible control of stimulus waveform; or (iii) an extension of previous neurophysiological studies on spatiotemporal perception into the area of psychophysics.] Both experiments involve simultaneous masking techniques, in contrast to previous array-based studies of tactile masking (for example, Craig and Xu, 1990; Cholewiak and Collins, 1995; Horner, 1997, 2000; Craig, 2000) which have concentrated on sequential presentation of target and masker because of the limitations of previous stimulator hardware. [For a study involving simultaneous masking, see Weisenberger and Craig (1982).] In addition, both experiments involve discrimination of direction of movement on the array. This topic has been the subject of several previous investigations (Essick *et al.*, 1992; Gardner *et al.*, 1992; Gardner and Sklar, 1994; Szaniszlo *et al.*, 1998; Cholewiak and Collins, 2000; Tan *et al.*, 2000), although it is not the main focus of the present study.

II. THE STIMULATOR ARRAY

A. Description of the array

When using a stimulator array to create ''lifelike'' spatiotemporal patterns of tactile stimulation on the skin, the intention is not to reproduce the topology of ''real'' surfaces—rather, it is to reproduce an appropriate excitation pattern over the various populations of mechanoreceptors in the skin. The spatial resolution required for stimulus presentation is thus related to the density of receptors, and is effectively determined by the spatial acuity for tactile perception—around 1 mm on the fingertip for ''real'' stimuli such as gratings (Johnson *et al.*, 2000). The array built for the experiments described in this paper has 100 piezoelectric-bimorph drive elements linked to 100 contactors arranged on a square matrix (1-mm$\times$1-mm spacing) over an area of 1 cm^2. The plane of the stimulating surface is horizontal in normal use, with motion of the contactors normal to this surface. The 10×10 square of moving contactors is surrounded by a ring of 44 fixed contactors, giving a 12×12 array overall, designed to cover the fingertip. The fixed contactors are included to ensure that each of the moving contactors is surrounded on all sides by other contactors and is thus subject to a similar mechanical load from the skin.

Further details of the mechanical arrangement are shown in Fig. 1(a). Each bimorph is clamped at one end and at the other carries an L-shaped wire link (brass, 0.6-mm diameter) whose free end acts as a circular contactor on the skin. Drive voltages to the bimorphs cause them to bend, and this results in movement of the contactors. The wire links are electrically insulated from the bimorphs to avoid any possibility of electrotactile stimulation. To ensure accurate spacing of the contactors, the links run through a low-friction (PTFE) plate containing a matrix of 0.8-mm holes at 1-mm$\times$1-mm spacing. The surface formed by the contactors lies 1 mm above the upper surface of the PTFE plate. In normal use the contactors maintain contact with the skin, but the PTFE plate is clear of the skin. (The fact that each contactor is ringed by neighboring contactors serves to limit mechanical spread of the vibratory stimulation across the skin, in the same way that a fixed surround may be used to localize the vibration from single-contactor stimulators.)

The piezoelectric-bimorph drive elements are arranged in five tiers, each tier consisting of a ring of 20 bimorphs [see Fig. 1(b)]. This arrangement accommodates bimorphs whose width is greater than the 1-mm pitch of the array. This is a key design feature, since narrow bimorphs produce less output force and typically necessitate operation at a system resonance in order to achieve appropriate levels of stimulation. In contrast, the greater output force produced by wider bimorphs allows broadband operation since sufficient output is available at frequencies away from system resonances. The parallel-connection bimorphs selected for the array are based on lead zirconate titanate [piezoelectric material type PZT-5H (Morgan Matroc, Inc.)] and have the following di-

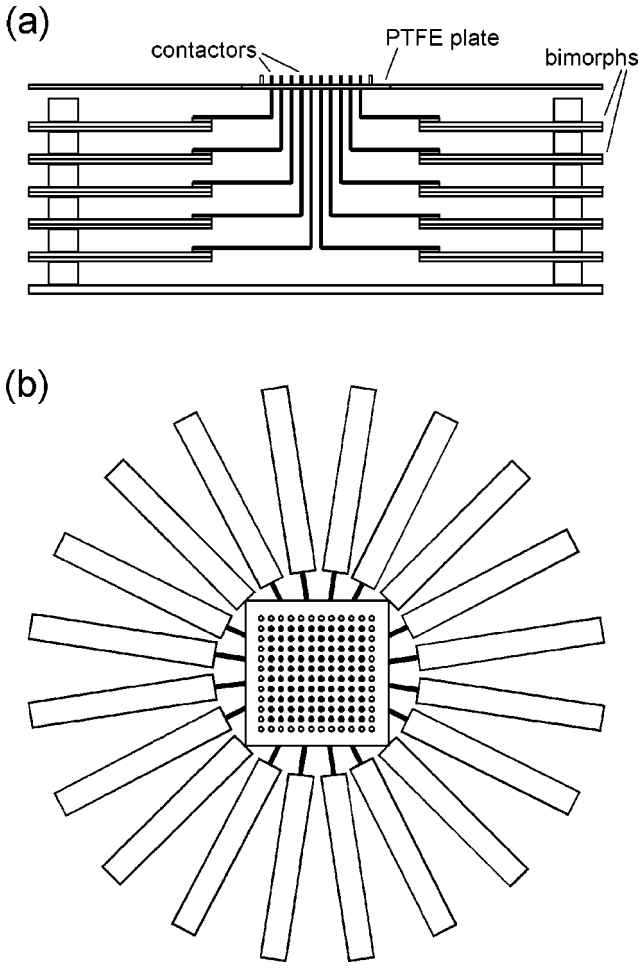


FIG. 1. Schematic diagrams of the array mechanism (not to scale): (a) side view (section) showing 2 of the 20 stacks of piezoelectric bimorphs, each stack holding 5 bimorphs and each bimorph carrying an L-shaped wire link whose free end acts as a contactor; (b) top view showing the arrangement of the 20 stacks to give 5 tiers of bimorphs, each tier consisting of a ring of 20 bimorphs—the exact geometry of the wire links, which varies from tier to tier so as to address all 100 contactor positions, is not shown; it appears that this arrangement requires some of the wire links to be grossly ''bent,'' but this is an artifact of showing the contactor matrix at larger than scale size and in practice such geometrical distortions are small.

mensions: width 3 mm, thickness 0.53 mm, length 40 mm (free length 34 mm). When in contact with the skin, each bimorph/link combination has a broad resonance at around 140 Hz (Q-factor~ 2), allowing effective operation over a frequency range of 20–400 Hz. The dimensions of the bimorph determine its bending stiffness which is critical to the successful operation of the array. Too much stiffness results in insufficient output because the bimorph is unresponsive to drive signals; too little stiffness also results in insufficient output—in this case, the bimorph deforms because it cannot withstand the blocking force from the mechanical (skin) load. (Loosely speaking, the problem is one of ''impedance matching'' between the bimorph and the skin.)

In normal operation, the drive signals in each of the 100 parallel channels are independently specified in software with a sample rate of 2560 Hz and 8-bit resolution. Within the working bandwidth of 20–400 Hz, a wide variety of waveforms is possible for each drive signal, allowing spatiotemporal patterns of displacement at the skin to be specified with great flexibility.

B. Calibration of the array

Calibration of the system presents problems for two reasons: (i) there is insufficient space within the array mechanism to fit measurement transducers for each of the 100 channels; (ii) the response of each channel depends on the mechanical load presented to the corresponding contactor. In order to check the uniformity of response across the array, subjective cross comparison was carried out between one ''reference'' channel and each of the other ''test'' channels in turn, with author CMC as the test subject and his right index finger as the ''standard'' load. A 320-Hz sine wave was alternated between the reference channel and the test channel, with the drive signal in the reference channel set to give a comfortable sensation level and the drive signal in the test channel adjusted by the subject to give a sensation level equal to that of the reference. Some channels proved to have anomalously strong or anomalously weak response (attributed to variation of friction in the mechanism). In subsequent experiments, using computer-generated stimulus waveforms, this variation in response between channels was compensated in software by applying appropriate normalization factors across the 100 channels. The only observed effect relating to nonuniform response across the array—localized stimuli appearing to ''bobble'' slightly in subjective intensity as they moved across the array—was eliminated with the normalization in place.

Absolute calibration of the system was carried out by measuring single-contactor detection thresholds, again with author CMC as the test subject and his right index finger as the standard load. Thresholds for sine waves at 40 and 320 Hz (the two frequencies used for subsequent experiments) were determined separately for nine representative contactors on the array, measured in terms of bimorph drive voltage in each of the nine channels. Average values over the nine channels were calculated, taking into account the normalization factors discussed in the previous paragraph. In addition, equivalent displacement-threshold measurements at 40 and 320 Hz were made on a single-channel device designed to be a close copy of a section of the array: a bimorph and wire link to provide a single moving contactor, surrounded by eight fixed contactors to give a 3×3 assembly overall. Displacement output was measured by a miniature accelerometer (Knowles type BU1771). Hence, thresholds were measured both in terms of drive voltage on the array and in terms of displacement on the array (equivalent), allowing the conversion between drive voltage and displacement output to be estimated at each of the two measurement frequencies: 0.55 μm V^{-1} at 40 Hz and 0.065 μm V^{-1} at 320 Hz. (The maximum available drive voltage of 85 V peak-to-peak produces a displacement output of around 50 μm peak-to-peak at 40 Hz and around 6 μm peak-to-peak at 320 Hz.) This imbalance between the two frequencies is intentional: it derives from setting the resonant frequency of the mechanical system at around 140 Hz and is intended to compensate for the difference in sensitivity of the skin between 40 and 320 Hz. Results from preliminary trials indicated that a 10:1 ratio

of amplitudes at 40 and 320 Hz would give approximately the same sensation level for corresponding 40- and 320-Hz stimuli. (See below for a discussion of sensation level in this context.) Thus, in the experiments detailed below, the imbalance between the two frequencies was increased from around 8 (due to the mechanical response) to around 10 by means of an additional factor of 1.25 in software.

Since a 10:1 ratio of displacement amplitudes at 40 and 320 Hz is required to give approximately the same sensation level (all other things being equal), it is convenient, when describing stimuli, to factor out this imbalance between the two frequencies and work in terms of normalized units for amplitude (with different normalization factors at 40 and 320 Hz—in the ratio 10:1). Relative amplitude in normalized units may then be taken as an approximate indication of relative sensation level. Thus, in the description of experiments given below, amplitudes are given in normalized units: 1 normalized unit corresponds to a displacement amplitude of approximately 0.4 μm peak-to-peak at 40 Hz and to a displacement amplitude of approximately 0.04 μm peak-to-peak at 320 Hz.

III. EXPERIMENT 1

A. Rationale

Philips *et al.* (1992) describe an ingenious experiment in which the responses of different types of mechanoreceptor to embossed dot patterns scanned across the fingertip are recorded as electrical activity in the corresponding afferent nerves, allowing the spatial resolutions of the different types of receptor to be compared. Experiment 1 is essentially an attempt to produce, using the stimulator array described above, a psychophysics equivalent of the Philips *et al.* experiment—a localized target, produced by vibration of a small number of adjacent contactors, is scanned across the array and discrimination of the direction of movement is compared at different vibration frequencies. However, as described, the discrimination task is too easy and ceiling effects cause convergence of data for different vibration frequencies. [A similar problem was observed by Hasser and Weisenberger (1993) in their investigation of pattern identification at different frequencies.] In order to avoid these ceiling effects it is necessary to make the task more difficult—this is conveniently achieved by introducing a uniform low-level ''background'' vibration across the entire array so as to partially mask the moving target. A particular attraction of this ''signal-in-noise'' paradigm is that it allows investigation of spatial acuity (expected to vary between different receptor populations) because good spatial acuity is required to take full advantage of the contrast between the localized target and the background which surrounds it. [This strategy for investigating spatial resolution has parallels with signal-in-noise masking techniques used to measure frequency resolution in auditory perception, dating from the work of Fletcher (1940). For a related tactile experiment, see Loomis (1993).]

B. General description

As outlined in the previous section, this is an experiment on the perception of a moving target (created by contactors driven at amplitude S) against a uniform background (created by the remainder of the array driven at amplitude N, with $N \leqslant S$). Subjects are required to identify the direction of motion: up, down, left, or right (viewed in the plane of the array). The target is in the form of a line of L array elements ($L = 1, 3, 5, 7$), moving across the array in a direction perpendicular to the line: a horizontal line moves up or down and a vertical line moves left or right. (For $L = 1$, the ''dot'' moves in any one of the four directions.)

Each stimulus has an overall duration of 2.0 s. The background amplitude ramps up from zero to N over the first 0.2 s, remains constant at N for 1.6 s, and then ramps down from N to zero over the last 0.2 s. In all cases the target with amplitude S is presented for 1.0 s at the midpoint of the stimulus, i.e., it appears 0.5 s after stimulus onset and disappears 1.5 s after stimulus onset. During this time it moves from one side of the array to the other in increments of 1 mm at an average speed of 10 mm s^{-1}, i.e., one increment every 0.1 s. There is no delay between offset of vibration at one target position and onset of vibration at the next position, i.e., each position is indicated by vibration for the full 0.1 s. It should be noted that the background surrounds the target rather than covering the whole array uniformly, i.e., the ''target'' contactors carry only the target, not both target and background.

In this experiment, all the contactor pins (both target and background) are driven with sine waves at the same frequency f (and with the same phase). With the target amplitude S held constant, discrimination is investigated as a function of the background amplitude N. As outlined above, this ''signal-in-noise'' task provides a measurement of some sort of spatial acuity at frequency f. Half of the test blocks were run with $f = 40$ Hz and the remainder with $f = 320$ Hz. These frequencies were chosen in the hope of targeting the measurement to nonpacinian and pacinian touch receptors, respectively, following the suggestion of Bernstein (1989) and the findings of Bolanowski *et al.* (1988)—it should be noted that the latter investigation relates to the thenar eminence and so it may not be entirely relevant to the present study on the fingertip, where the detection threshold is lower (Verrillo, 1971) and the density of pacinian receptors is higher (Stark *et al.*, 1998).

C. Subjects

These were 8 adults (age range 21 to 29), 6 male and 2 female. They were not paid for their participation in the experiment.

D. Test procedure

The subject rested the right index fingertip on the array. A cushioned support was provided for the forearm. He/she was requested to indicate the direction of movement of each test item (up, down, left, or right) using keyboard entry. Each subject underwent eight test sessions, one for each of the eight test conditions (four values of target length L with two values of frequency f). The order of test conditions was permutated across the subject group to balance any learning

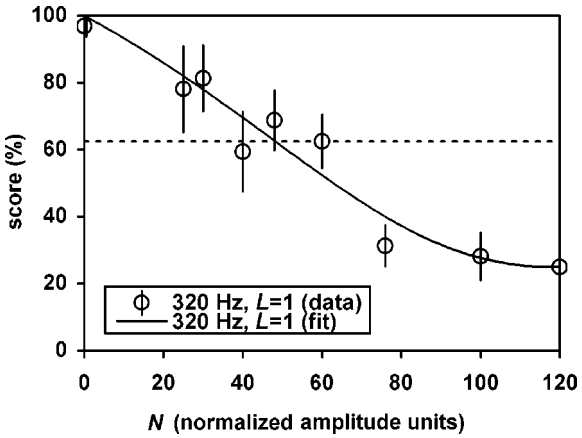


FIG. 2. Mean discrimination scores from 8 subjects in experiment 1, for a target of length $L=1$ at amplitude $S=120$, plotted as a function of background amplitude N. (See the text for explanation of normalized units for amplitude.) Both target and background are at 320 Hz. The error bars indicate the standard error in each score. The data point at $N=120$ is inferred at the chance score of 25%. Also shown is a fit to the data, using an appropriate psychometric function (see the text). The dotted line (score of 62.5%) intercepts the psychometric function at the critical value of N, in this case 48 ± 13.

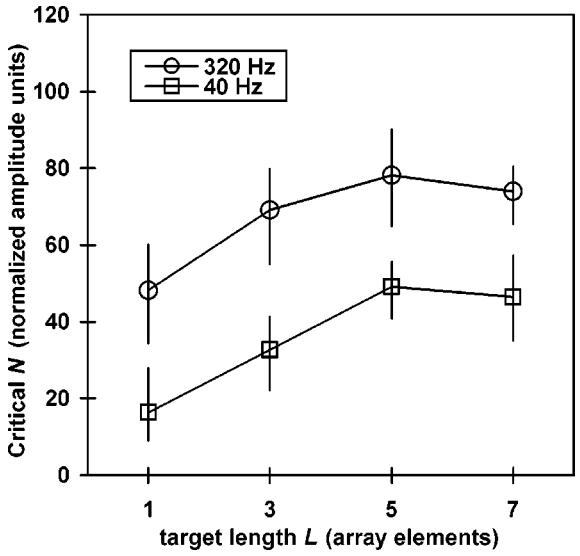


FIG. 3. Critical values of background amplitude N (at which the direction of movement of a target of amplitude $S=120$ is discriminable with 62.5% success rate) as a function of the length L of the target. (See the text for explanation of normalized units for amplitude.) For these data, both target and background are at the same frequency f: the two data sets are for $f=40$ Hz and $f=320$ Hz. Error bars are estimated from consideration of the individual curve fits (as in Fig. 2).

effects. The eight test sessions were grouped in pairs, with each subject attending for testing on four separate occasions.

In each test session the test sequence consisted of a demonstration block (8 items, each with the direction of movement identified to the subject via a computer screen) followed by a training block (8 items, each identified after the subject had made a response) and a test block (40 items, with the first 8 acting as a ''warm up'' and only the last 32 being scored; no feedback was provided during the test block). Subjects were able to request one or more repetitions of the complete demonstration block, if required, and to request repetitions of individual stimuli within the demonstration block or the training block. (It should be noted that the task is relatively ''intuitive''—touch stimuli which move across the fingertip are familiar from everyday situations—and hence subjects are able to understand the task with only a little practice.) For each test condition the 32 scored test items were equally distributed over 8 different values of background amplitude N. These values varied to some extent between the test conditions: they were chosen on the basis of preliminary trials so that floor and ceiling effects were avoided for the great majority of test items.

E. Results

Data for each test condition (mean discrimination scores over the eight subjects as a function of the background amplitude N, with the target amplitude S held constant) were fitted with an appropriate psychometric function $p(N)$ for probability of correct response, and in each case an estimate was obtained for the critical value of N at which the targets are just discriminable. The psychometric function takes the value $p(N)=1$ when $N=0$ and $p(N)=0.25$ when $N=S$; the critical value of N is taken from the ''halfway'' point of the function when $p(N)=0.625$. An example of experimental data and curve fit is shown in Fig. 2, for the case $L=1$, $f=320$ Hz. [Further examples of data from this experiment—for the cases $L=5$, $f=40$ Hz and $L=5$, $f=320$ Hz—can be found in Fig. 4, below, together with corresponding curve fits (dotted lines).]

The background amplitude N is expressed here in normalized units ($N \leqslant 120$), with a target amplitude S of 120 in the same units. At 40 Hz, on the basis of the calibration detailed above, amplitude 120 corresponds to a displacement amplitude of around 45 μm peak-to-peak; at 320 Hz, amplitude 120 corresponds to around 4.5 μm peak-to-peak. The conversion between displacement amplitude and a quantitative estimate of sensation level (i.e., level in decibels with respect to detection threshold) is not straightforward because of spatial summation: the detection threshold varies with the size of the area stimulated and, moreover, is expected to vary differently at the two experimental frequencies—little or no effect at 40 Hz, but significant spatial summation at 320 Hz (Verrillo, 1963; Rabinowitz *et al.*, 1987). However, the test stimuli are all at a ''comfortable'' sensation level. (As noted above, the 10:1 ratio of amplitudes at 40 and 320 Hz gives approximately the same sensation levels at the two frequencies).

Figure 3 shows the critical values of N, as obtained from the curve fits described above, for each of the eight test conditions. The error bars here are estimated from consideration of errors for the data used as the basis for each curve fit. (Standard errors in the mean discrimination scores are around 10 percentage points, except at the top of the range where they are reduced by ceiling effects—see Fig. 2, for example.) In Fig. 3, at each value of target length L, there is a clear difference between the data for $f=40$ Hz and those for $f=320$ Hz. The limit for discrimination occurs at higher values of target/background ratio S/N for 40-Hz stimuli than for 320-Hz stimuli. Both data sets ($f=40$ Hz and $f=320$ Hz) indicate that critical values of N increase as L

increases from 1 to 5, i.e., larger targets are more resistant to masking by the background, as might be expected.

F. Discussion

The clear difference between the data sets at 40 and at 320 Hz suggests that different populations of receptors may have been targeted with the two different stimulus frequencies. However, there is a problem with this interpretation: there are strong suggestions in the literature that nonpacinian receptors have greater spatial acuity than pacinian receptors [for both vibratory stimuli (Sherrick *et al.*, 1990) and "real" stimuli (Johnson and Lamb, 1981; Philips *et al.*, 1992)] and so it might be expected that discrimination in this signal-in-noise task would be easier at 40 Hz than at 320 Hz. However, the experimental data suggest that the converse is true: for equivalent performance, higher values of target/background ratio S/N are required for 40-Hz stimuli than for 320-Hz stimuli. This point is considered in more detail in the Overall Discussion towards the end of the paper.

In order to further explore the possibility that the two experimental frequencies target different populations of receptors, a second experiment was carried out in which the target and background are at different frequencies. If backgrounds at a different frequency produce much less masking of the target than backgrounds at the same frequency, involvement of different receptor populations is suggested. [Again, there are parallels here with the techniques of auditory psychophysics, deriving from the critical-band theory of Fletcher (1940). For a similar discussion relating to tactile perception, see Makous *et al.* (1995).]

IV. EXPERIMENT 2

A. General description

This experiment is very similar to experiment 1, both in terms of the overall design and in terms of the detailed test procedure, except that there are only two test conditions: because the effect of stimulus frequency in experiment 1 was largely independent of target length L (see Fig. 3), to save on subjects' time it was decided to use a single target length only ($L=5$) with (i) 40-Hz target and 320-Hz background; (ii) 320-Hz target and 40-Hz background. As in experiment 1, the background *surrounds* the target, i.e., the target contactors carry only the target, not both target and background. The eight subjects for experiment 2 were the same as those who participated in experiment 1. The order of the two test conditions was permutated across the subject group to balance any learning effects.

B. Results

Data for each test condition (mean discrimination scores over the eight subjects as a function of the background amplitude N, with the target amplitude S held constant at $S=120$) are shown in Fig. 4, together with the corresponding data from experiment 1 (i.e., for $L=5$). [Amplitudes are again expressed in normalized units (see above): at 40 Hz, amplitude 120 corresponds to a displacement amplitude of around 45 μm peak-to-peak; at 320 Hz, amplitude 120 corresponds to around 4.5 μm peak-to-peak.] It is clear that the "different-frequency" data from experiment 2 do not have the same form as the "same-frequency" data from experiment 1—backgrounds at a different frequency appear to provide much less masking of the target than backgrounds at the same frequency.

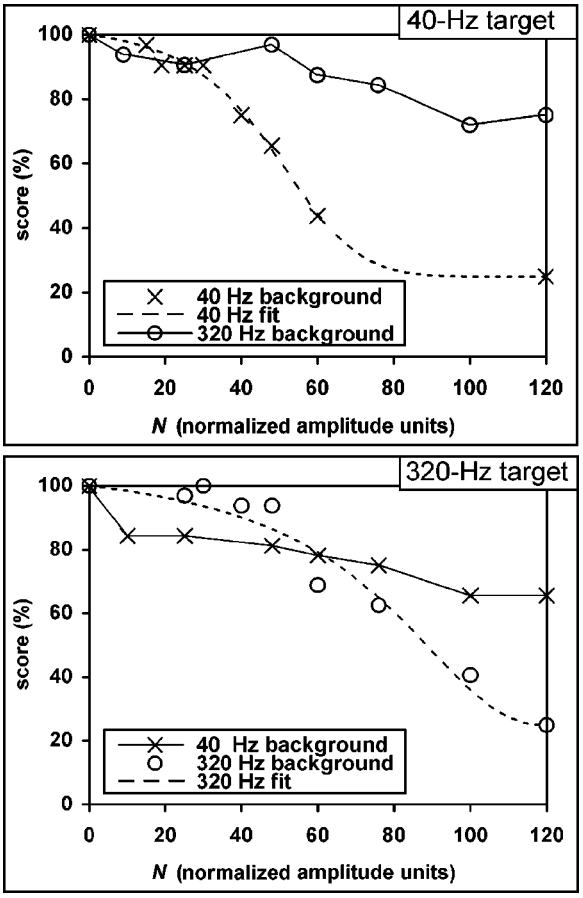


FIG. 4. Mean discrimination scores from 8 subjects, for a target of length $L=5$ at amplitude $S=120$, plotted as a function of background amplitude N. (See the text for explanation of normalized units for amplitude.) Upper panel: 40-Hz target in a 320-Hz background (experiment 2) and in a 40-Hz background (experiment 1, $L=5$ data); lower panel: 320-Hz target in a 40-Hz background (experiment 2) and in a 320-Hz background (experiment 1, $L=5$ data). In each case the experiment-1 data are fitted with an appropriate psychometric function, as described in the text.

There is a complicating factor which must be considered when comparing data from the two experiments on the same axes: the array does not necessarily give the same effective intensities for a 40-Hz background and 320-Hz background when these have the same value of background amplitude N. (As discussed above, equal amplitude in normalized units is expected to produce backgrounds whose sensation levels are approximately the same but are unlikely to be exactly the same, and in any case it is not obvious that the sensation level of the background equates exactly with its effective intensity in terms of masking a target.) Thus, the data in Fig. 4 (upper panel) for a 40-Hz target might be simply explained in terms of the 320-Hz background being effectively weaker than the 40-Hz background when these have the same nominal amplitude N. Similarly, the data in Fig. 4 (lower panel) for a 320-Hz target might be simply explained in terms of the

40-Hz background being effectively weaker than the 320-Hz background. However, these two explanations are clearly mutually exclusive, and simultaneous consideration of both panels in Fig. 4 suggests that the effective intensities for a 40-Hz background and a 320-Hz background are broadly similar, as is to be expected on the basis of the calibration procedures discussed above. We may thus conclude that the data in Fig. 4 do indeed indicate that different-frequency backgrounds provide much less masking of the target than same-frequency backgrounds.

The discussion in the previous paragraph indicates the potential pitfalls in a quantitative comparison of the data from experiment 1 and experiment 2. However, the following simple statistical analysis is probably justified: With a 40-Hz target (see Fig. 4, upper panel) the $N=120$ data from experiment 2 show a mean score of 75%; with a 320-Hz target (see Fig. 4, lower panel) the $N=120$ data from experiment 2 show a mean score of 66%. The 70% average of these two scores may be compared to the corresponding average from experiment 1, i.e., the average of the (inferred, chance) scores for $N=120$, $L=5$. This has a binomial distribution (64 trials, 0.25 probability of success), for which the maximum expected score is 42% ($p<0.001$). Hence, the difference between the experiment-1 data and the experiment-2 data is highly significant.

C. Discussion

The results of experiment 2, taken in conjunction with the results of experiment 1, show that discrimination of 40-Hz targets against 40-Hz backgrounds is much more difficult than against 320-Hz backgrounds of the same nominal amplitude N. Similarly, discrimination of 320-Hz targets against 320-Hz backgrounds is much more difficult than against 40-Hz backgrounds. This provides further evidence that the different experimental frequencies may be targeting different populations of receptors: The results are easily explained in terms of two populations of receptors, one more sensitive at 40 Hz and the other more sensitive at 320 Hz. In the ''same-frequency'' cases, the population which is predominantly excited by the target is also predominantly excited by the background; in the ''different-frequency'' cases, the population which is predominantly excited by the target is different from that predominantly excited by the background, allowing less masking of the target by the background.

V. OVERALL DISCUSSION

A. General observations

In the absence of a background (i.e., with $N=0$) the discrimination task is very easy. In 320 trials with $N=0$ in experiments 1 and 2 there were only two response errors. In the presence of a background, i.e., when there are appreciable numbers of response errors, no significant differences were observed between scores for the various directions of stimulus motion (although the experiments were not designed with this comparison in mind—differences might become apparent if more data were to be taken). Stark *et al.* (1998) report the distribution of pacinian receptors in the fingertip to be nonuniform, with clustering around branches of the digital nerve, and this might in principle be expected to produce anisotropic performance in a motion-detection task which involves these receptors.

As described above, both experiments suggest that vibratory stimulation at the different experimental frequencies may be targeting different populations of receptors in the skin. In this regard it is interesting to note that the reported nature of the subjective sensation is quite different at the two frequencies: 40-Hz stimuli tend to give the impression of a real object moving beneath the fingertip, whereas 320-Hz stimuli tend to simply give the impression of a superficial vibration moving over the skin surface. For moving 320-Hz stimuli with no background, subjects report that the individual increments of the motion (1-mm steps at 0.1-s intervals) can be distinguished to some extent, and the stimulus may appear as a sequence of individual quasistationary elements. This is not the case for 40-Hz stimuli with no background, for which the individual increments appear to merge into a steady motion and the stimulus is perceived as a single, moving entity. These findings are comparable to those of Gardner and Sklar (1994), using an Optacon array to deliver pulse stimuli that excite both pacinian and nonpacinian receptors on the fingertip: They report that 1.2-mm increments in the motion of a moving bar are individually distinguishable by subjects at 40-ms intervals (scan speed $30\ \mathrm{mm\ s^{-1}}$) but not at 10-ms intervals (scan speed $120\ \mathrm{mm\ s^{-1}}$).

B. Targeting of different receptor populations

As discussed above, the experimental frequencies—40 and 320 Hz—were chosen in the hope of targeting measurements to nonpacinian and pacinian touch receptors, respectively. In the event, both experiments provide some evidence that this targeting has been realized: In experiment 1, equivalent performance is obtained at much higher values of target/background ratio for 40-Hz stimuli than for 320-Hz stimuli. In experiment 2, there is much less masking of the target in the different-frequency cases (target at 40 Hz, background at 320 Hz, or vice versa) than in the same-frequency cases (both target and background at 40 or 320 Hz). However, as mentioned above, there is a problem with this interpretation in relation to experiment 1, since results from that experiment indicate better spatial acuity at 320 Hz than at 40 Hz, contrary to expectations based on the known characteristics of pacinian and nonpacinian receptors.

This apparent contradiction warrants a more detailed consideration of situations similar to those in experiment 1: For a localized target surrounded by a uniform background, the perceived contrast C_p is related to the objective contrast C_o (as specified in the stimulus) by the spatial acuity of the perceptual mechanism. For example, for an extended line target we may estimate $C_p=kC_o/s$, where s is a distance which characterizes the spatial acuity and k is a constant. For discrimination of the target we may assume that a certain threshold C_t for perceived contrast must be attained, i.e., $C_p>C_t$ or $C_o>C_ts/k$. Similar arguments apply to line targets of finite length, i.e., to all cases in experiment 1. The limiting objective contrast thus relates not only to the spatial

acuity of the perceptual mechanism, but also to the threshold for perceived contrast for that perceptual mechanism. If different perceptual mechanisms are involved, e.g., different populations of receptors, observed differences in the limiting objective contrast can be attributed to differences in spatial acuity only if the threshold for perceived contrast is similar for the two mechanisms.

There is another possible explanation of the experiment 1 results: Given the known characteristics of the various mechanoreceptors (Bolanowski *et al.*, 1988; Gescheider *et al.*, 2001) it is likely that a 320-Hz target in a 320-Hz background will excite nonpacinian (NP II) receptors to some extent, as well as pacinian receptors, and in principle the most useful spatial cues may be provided by whichever channel has the greater target/background ratio (whether or not that channel has the higher sensation level for the target). Hence, it is possible that the observed discrimination of 320-Hz targets derives primarily from nonpacinian receptors, and the difference between data at the two experimental frequencies derives from differences between the nonpacinian (NP I, NP III) response at 40 Hz and the nonpacinian (NP II) response at 320 Hz. This explanation is conjectural, but in this context it is interesting to note that Gescheider *et al.* (2001) observed strong evidence for excitation of both pacinian and nonpacinian (NP II) receptors when stimulation at around 250 Hz was applied via a contactor of area 0.008 cm^2. [As mentioned above, pacinian receptors exhibit strong spatial summation: the apparent sensitivity of the pacinian system falls as the area of stimulation is reduced. For vibratory stimuli in the frequency range where pacinian receptors are most sensitive (around 250 Hz) the pacinian sensitivity is much higher than the nonpacinian (NP II) sensitivity when stimulation is over a large area, but this may not be the case when stimulation is over a small area.]

However, a further complication is that simple analyses of the sort given in the previous paragraphs may not be appropriate for the situation in experiment 1. This is because such analyses assume what may be termed ''quasistatic'' perception, i.e., that moving stimuli are perceived primarily as a sequence of ''snapshot'' spatial distributions of stimulation and that the percept of motion is secondary, based on a comparison of spatial distributions at different instants in time. In practice, it may well be that tactile perception involves motion detectors at the lowest levels of the cognitive hierarchy, producing primary sensations in their own right, as in visual perception (Fredericksen *et al.*, 1993; Lu and Sperling, 1995). Some evidence for this is provided by the results of Gardner and Sklar (1994), using an Optacon array to deliver pulse stimuli to the fingertip: sequential stimulation at only two locations to mark the start and finish of a moving stimulus was found to give insufficient information to identify the direction of motion; improved identification was obtained with stimulation (on the same time scale) which included intermediate locations.

Despite these uncertainties about the analysis of the results from experiment 1, the tentative conclusion that subjects' performance is mediated by better spatial acuity at 320 Hz than at 40 Hz is supported by subjects' reports on the subjective nature of the stimuli (see the previous section), to the effect that the individual increments of the motion (1-mm steps at 0.1-s intervals) may be distinguished at 320 Hz but not at 40 Hz. It is possible to conclude that all experimental observations in the present study are consistent with better spatial acuity at 320 Hz than at 40 Hz, even if the reasons for this are difficult to understand. (In principle, an explanation might be found in terms of mechanical spread of the vibratory stimulation from contactor to contactor, either across the skin or through the mechanism of the array. However, it is not obvious that any such an effect would be so markedly different at the two experimental frequencies as to convert an intrinsic advantage of spatial acuity at 40 Hz into the observed disadvantage.) It is interesting to note that Rogers (1970), in one of the very few studies on localization of sinusoidal vibratory stimuli at different frequencies on the fingertip, found spatial acuity to be (slightly) better at 250 Hz than at 10 Hz. Rogers' method, involving simultaneous presentation of a strong vibration and a weak vibration at two adjacent sites, has some similarities with the methods of the present investigation. In the same paper, Rogers describes a different experiment involving recognition of letters moving over an air-jet array with pulse output: better results were again obtained at higher stimulation frequency (160 Hz) than at lower frequency (20 Hz), although in this case the use of pulse stimuli rather than sine waves hinders a comparison with the present investigation (because pulse stimuli have components at frequencies other than the fundamental frequency).

VI. CONCLUSION

The results of this study demonstrate that the design goals for the array (i.e., to allow the use of different stimulation frequencies to investigate the various aspects spatiotemporal perception which are mediated by different populations of mechanoreceptors) have been satisfied, and indicate the potential of broadband array stimulators as a tool for measurements on the sense of touch in glabrous skin. Questions remain to be answered in relation to the response of different receptor populations to stimuli from such arrays. The production of virtual textures will be of particular interest in future studies—this aspect of tactile perception is of considerable importance in the context of virtual environments, and there is currently little information on the relation between objective and subjective aspects of artificial stimuli intended to evoke texture sensations. The array used in the present study allows only passive perception, i.e., there is no intervention by the subject in the generation of the tactile stimulus. It is possible that perception in an active scenario may be very different from that observed in the passive case. Development of a movable array which will allow active exploration of a virtual tactile environment is an interesting possibility for future work.

ACKNOWLEDGMENTS

Author C.M.C. was supported by the Engineering and Physical Sciences Research Council. Janet Weisenberger and

Jim Craig provided helpful discussion at early stages of the project. Three anonymous reviewers are thanked for their useful suggestions.

Bernstein, L. E., Eberhardt, S. P., and Demorest, M. E. (**1989**). ''Single-channel vibrotactile supplements to visual perception of intonation and stress,'' J. Acoust. Soc. Am. **85**, 397–405.

Bliss, J. C., Katcher, M. H., Rogers, C. H., and Shepard, R. P. (**1970**). ''Optical-to-tactile image conversion for the blind,'' IEEE Trans. Man-Machine Sys. **11**, 58–65.

Bolanowski, S. J., Gescheider, G. A., Verrillo, R. T., and Checkowsky, C. M. (**1988**). ''Four channels mediate the mechanical aspect of the sense of touch,'' J. Acoust. Soc. Am. **84**, 1680–1694.

Cholewiak, R. W., and Collins, A. A. (**1995**). ''Vibrotactile pattern discrimination and communality at several body sites,'' Percept. Psychophys. **57**, 724–737.

Cholewiak, R. W., and Collins, A. A. (**2000**). ''The generation of vibrotactile patterns on a linear array: Influences of body site, time, and presentation mode,'' Percept. Psychophys. **62**, 1220–1235.

Craig, J. C. (**1980**). ''Modes of vibrotactile pattern generation,'' J. Exp. Psychol. Hum. Percept. Perform. **6**, 151–166.

Craig, J. C. (**2000**). ''Processing of sequential tactile patterns: Effects of a neutral stimulus,'' Percept. Psychophys. **62**, 596–606.

Craig, J. C., and Xu, B. (**1990**). ''Temporal order and tactile patterns,'' Percept. Psychophys. **47**, 22–34.

Essick, G. K., McGuire, M., Joseph, A. K., and Frantzen, O. (**1992**). ''Characterization of the percepts evoked by discontinuous motion over the perioral skin,'' Somatosens Mot Res. **9**, 175–184.

Fletcher, H. (**1940**). ''Auditory patterns,'' Rev. Mod. Phys. **12**, 47–65.

Fredericksen, R. E., Verstraten, F. A. J., and Van de Grind, W. A. (**1993**). ''Spatiotemporal characteristics of human motion perception,'' Vision Res. **33**, 1193–1205.

Gardner, E. P., Palmer, C. I., Hamalainen, H. A., and Warren, S. (**1992**). ''Simulation of motion on the skin 5: Effect of stimulus temporal frequency on the representation of moving bar patterns in primary somatosensory cortex of monkeys,'' J. Neurophysiol. **67**, 37–63.

Gardner, E. P., and Sklar, B. F. (**1994**). ''Discrimination of the direction of motion on the human hand: A psychophysical study of stimulation parameters,'' J. Neurophysiol. **71**, 2414–2429.

Gescheider, G. A., Bolanowski, S. J., and Hardick, K. R. (**2001**). ''The frequency selectivity of information-processing channels in the tactile sensory system,'' Somatosens Mot Res. **18**, 191–201.

Hasser, C. J., and Weisenberger, J. M. (**1993**). ''Preliminary evaluation of a shape-memory-alloy tactile feedback display,'' Proc. ASME DSC **49**, 73–80.

Horner, D. T. (**1997**). ''The effect of shape and location on temporal masking of spatial vibrotactile patterns,'' Percept. Psychophys. **59**, 1255–1265.

Horner, D. T. (**2000**). ''Perceptual processing at adjacent locations an a single finger: Masking and response competition,'' Percept. Psychophys. **62**, 1545–1555.

Johnson, K. O., and Lamb, G. D. (**1981**). ''Neural mechanisms of spatial tactile discrimination: Neural patterns evoked by braille-like dot patterns in the monkey,'' J. Physiol. (London) **310**, 117–144.

Johnson, K. O., Yoshioke, T., and Vega-Bermudez, F. (**2000**). ''Tactile functions of mechanoreceptive afferents innervating the hand,'' J. Clin. Neurophysiol. **17**, 539–558.

Lederman, S. J., and Klatzky, R. L. (**1999**). ''Sensing and displaying spatially distributed fingertip forces in haptic interfaces for teleoperator and virtual environment systems,'' Presence-Teleoper. Virtual Env. **8**, 86–103.

Loomis, J. M. (**1993**). ''Counterexample to the hypothesis of functional similarity between tactile and visual pattern perception,'' Percept. Psychophys. **54**, 179–184.

Lu, Z.-L., and Sperling, G. (**1995**). ''The functional architecture of human visual motion perception,'' Vision Res. **35**, 2697–2722.

Makous, J. C., Friedman, R. M., and Vierck, C. J. (**1995**). ''A critical band filter in touch,'' J. Neurosci. **15**, 2808–2818.

Pawluk, D. T. V., van Buskirk, C. P., Killebrew, J. H., Hsiao, S. S., and Johnson, K. O. (**1998**). ''Control and pattern specification for a high density tactile display,'' Proc. ASME DSC **64**, 97–102.

Philips, J. R., Johansson, R. S., and Johnson, K. O. (**1992**). ''Responses of human mechanoreceptive afferents to embossed dot arrays scanned across fingerpad skin,'' J. Neurosci. **12**, 827–839.

Rabinowitz, W. M., Houtsma, A. J. M., Durlach, N. I., and Delhorne, L. A. (**1987**). ''Multidimensional tactile displays: Identification of vibratory intensity, frequency, and contactor area,'' J. Acoust. Soc. Am. **82**, 1243–1253.

Raj, A. K., McGrath, B. J., Rochlis, J., Newman, D. J., and Rupert, A. H. (**1998**). ''The application of tactile cues to enhance situation displays,'' Proceedings of the 3rd Annual Symposium on Situational Awareness in the Tactical Environment, Patuxent River, MD, pp. 77–84.

Rogers, C. H. (**1970**). ''Choice of stimulator frequency for tactile arrays,'' IEEE Trans. Man-Machine Sys. **11**, 5–11.

Schneider, W. (**1988**). ''The tactile array stimulator,'' Johns Hopkins APL Tech. Dig. **9**, 39–43.

Sherrick, C. E., Cholewiak, R. W., and Collins, A. A. (**1990**). ''Localization of low- and high-frequency vibrotactile stimuli,'' J. Acoust. Soc. Am. **88**, 169–179.

Srinivasan, M. A., and Basdogan, C. (**1997**). ''Haptics in virtual environments: Taxonomy, research status, and challenges,'' Comput. Graph. **21**, 393–404.

Stark, B., Carlstedt, T., Hallin, R. G., and Risling, M. (**1998**). ''Distribution of human pacinian corpuscles in the hand: a cadaver study,'' J. Hand Surg. [Br] **23B**, 370–372.

Szaniszlo, J. A., Essick, G. K., Kelly, D. G., Joseph, A. K., and Bredehoeft, K. R. (**1998**). ''Evocation and characterization of percepts of apparent motion on the face,'' Percept. Psychophys. **60**, 785–804.

Tan, H., Lim, A., and Traylor, R. (**2000**). ''A psychophysical study of sensory saltation with an open response paradigm,'' Proc. ASME DSC **69-2**, 1109–1115.

Taylor, P. M., Moser, A., and Creed, A. (**1998**). ''A sixty-four element tactile display using shape memory alloy wires,'' Displays **18**, 163–168.

Van Doren, C. L., Pelli, D. G., and Verrillo, R. T. (**1987**). ''A device for measuring tactile spatiotemporal sensitivity,'' J. Acoust. Soc. Am. **81**, 1906–1915.

van Erp, B. F., and van Veen, A. H. C. (**2001**). ''Vibro-tactile information presentation in automobiles,'' Proc. Eurohaptics 2001, Birmingham, U.K., pp. 99–104.

Vega-Bermudez, F., and Johnson, K. O. (**1999**). ''SA1 and RA receptive fields, response variability, and population responses mapped with a probe array,'' J. Neurophysiol. **81**, 2701–2710.

Verrillo, R. T. (**1963**). ''Effect of contactor area on the vibrotactile threshold,'' J. Acoust. Soc. Am. **35**, 1962–1966.

Verrillo, R. T. (**1971**). ''Vibrotactile thresholds measured at the finger,'' Percept. Psychophys. **9**, 329–330.

Weisenberger, J. M., and Craig, J. C. (**1982**). ''A tactile metacontrast effect,'' Percept. Psychophys. **31**, 530–536.

Chaotic vibration induced by turbulent noise in a two-mass model of vocal folds

Jack J. Jiang[a] and Yu Zhang
Department of Surgery, Division of Otolaryngology Head and Neck Surgery, University of Wisconsin Medical School, Madison, Wisconsin 53792-7375

(Received 3 January 2002; revised 19 July 2002; accepted 24 July 2002)

The contribution of turbulent noise was modeled in symmetric vocal folds. A two-mass model was used to simulate irregular vocal fold vibrations. The threshold values of system parameters to produce irregular vibrations were decreased as a result of turbulent airflow. Periodic vibrations were then driven into the regions of irregular vibrations. Using nonlinear dynamics including Poincaré map and Lyapunov exponents, irregular vibrations were demonstrated as chaos. For the deterministic vocal-fold model with noise free and steady airflow, a fine period-doubling bifurcation cascade was shown in a bifurcation diagram. However, turbulent noise added to the vocal-fold model would induce chaotic vibrations, broaden the regions of irregular vocal fold vibrations, and inhibit the fine period-doubling bifurcations in the bifurcation diagrams. The perturbations from neurological and biomechanical effects were simulated as a random variation of the vocal fold stiffness. Turbulent noise as an external random source, as well as random stiffness perturbation as an internal random source, played important roles in the presence of irregular vocal fold vibrations.
 [DOI: 10.1121/1.1509430]

PACS numbers: 43.70.Aj, 47.27.Sd [AL]

I. INTRODUCTION

Chaotic or turbulent behaviors have been observed in physical, chemical, and biological systems. The inevitable existence of noise and fluctuations in real systems provides the basic reason for studying the effects of stochastic noises on chaotic systems. Recent studies have demonstrated that the interaction between dynamic and stochastic forces can produce a variety of interesting phenomena, such as jumps between different attractors,[1] escaping from the attractor,[2] deformation of the attractor,[3] delayed bifurcation,[4] instability of an integrable system,[5] stabilization by noise,[6] noise-induced chaos,[7–9] noise-induced order,[10,11] and stochastic resonance.[12–14] Thus, in recent years, noise-induced dynamics has become an intriguing subject with ongoing intensive experimental and theoretical investigations.

A noise-free deterministic mechanism has been used to describe nonlinear oscillations of vocal folds. The combined effects of nonlinear biomechanical and aerodynamic actions can produce complicated vibrations, including bifurcation and chaos.[15–21] Differing from only the first few period-doubling bifurcations observed in excised larynx experiments and speech analysis,[16,22–24] the very fine period-doubling bifurcation cascades[18–21] can be theoretically predicted using vocal-fold models with the assumptions of noise-free and steady airflow. However, the larynx is situated in physical proximity to major blood vessels, muscles, and nerves. It is especially susceptible to small fluctuations in neurological, biomechanical, aerodynamic, and acoustic systems.[25] In particular, recent research has shown that unsteady jet airflow radiating from the glottal exit shows the pattern of strong turbulence.[26–31] Numerous vortices generating and growing in turbulence downstream have been experimentally observed and numerically simulated. Fully developed turbulent flow results in the rapid and random fluctuations of velocity and pressure,[32] as well as aerodynamic noise.[33] Therefore, modeling the influences of these random perturbations on vocal-fold vibrations is important for a better understanding of irregular phenomena in excised larynx experiments and in voice analysis.

Differing from previous research based on the traditional steady airflow assumption, in this paper, we consider the effects of airflow at two regimes: the steady Bernoulli regime and the turbulent jet regime. The latter takes effect in the turbulent jet region of vocal folds. A two-mass model is applied to simulate the influences of turbulent noise and other random perturbations. Furthermore, employing Poincaré map and Lyapunov exponents study nonlinear dynamics of vocal-fold vibrations. Bifurcation diagrams illustrate the period-doubling bifurcation cascades. The results demonstrate that turbulent noise and random stiffness perturbation may qualitatively change the dynamics of the vocal-fold model, such as inducing chaos, broadening the intervals of irregular vocal fold vibrations, and inhibiting fine period-doubling bifurcations.

II. THE VOCAL-FOLD MODEL WITH TURBULENT NOISE

The two-mass model has been widely used to study nonlinear oscillations of vocal folds in voice production. The complicated vibratory patterns, such as bifurcation and chaos, can be observed in this vocal-fold model.[18–21] To investigate the influence of turbulent glottal flow on nonlinear dynamics of laryngeal systems, we will generalize the two-mass model of Steinecke and Herzel.[19,21] The dynamic equations can be described by two coupled nonlinear oscillators:

[a]Electronic mail: Jiang@surgery.wisc.edu

$$m_1\ddot{x}_1+r_1\dot{x}_1+k_1x_1+\Theta(-a_1)c_1\frac{a_1}{2l}+k_c(x_1-x_2)=F_1,$$
$$m_2\ddot{x}_2+r_2\dot{x}_2+k_2x_2+\Theta(-a_2)c_2\frac{a_2}{2l}+k_c(x_2-x_1)=F_2, \tag{1}$$

where x_i $(i=1,2)$ is the displacement of mass m_i, r_i is the viscous damping constant, k_i is the stiffness constant, d_i is the thickness of mass m_i, c_i is the additional spring constants during the collision between two sides of vocal folds, l is the length of the glottis, and k_c is the coupling stiffness between two masses m_1 and m_2. a_1 and a_2 denote the lower and upper glottal areas, respectively. For the symmetric model, the left side of vocal folds is the same as the right side, so that the glottal areas can be given as $a_1=a_{01}+2lx_1$ and $a_2=a_{02}+2lx_2$. a_{0i} is the glottal rest area. The function $\Theta(x)$ describing the collision is approximated by

$$\Theta(x)=\begin{cases}\tanh(50x/x_0), & x>0,\\ 0, & x\leqslant 0.\end{cases} \tag{2}$$

The driving forces F_i of two masses are expressed as $F_i=ld_iP_i$. The following assumptions are made in order to deduce the pressure P_i inside the glottis.

(1) The inertia of glottal air is small enough to be neglected.
(2) The interactions of the glottal flow with subglottal and supraglottal resonances are neglected.[34] Subglottal pressure P_s remains constant and the static component of supraglottal pressure P_0 is zero.
(3) The airflow detaches at the minimum glottal diameter $a_{\min}$, where

$$a_{\min}=\begin{cases}a_1, & x_1<x_2\\ a_2, & x_2\leqslant x_1\end{cases}.$$

The effects of airflow are separated into two regimes by the narrowest glottal gap: Bernoulli regime and jet regime.[35,36] Any point upstream (i.e., the region from the subglottal region to the minimum glottal diameter) satisfies the steady Bernoulli equation:

$$P(a)=P_s-\frac{\rho}{2}\left(\frac{U}{a}\right)^2, \tag{3}$$

where $P(a)$ denotes the pressure at the point with area a. However, downstream (i.e., the region from the minimum glottal diameter to the glottal exit), fully developed turbulence generates numerous vortices at the glottis exit. The masses in the turbulent region are perturbed by extremely irregular fluctuations. We assume that turbulence in the unsteady jet region acts as an external random force to drive vocal folds, and the additional pressure perturbation in this region is a random force with Gaussian distribution. This assumption presents a description differing from previous research where the contribution of turbulent noise was neglected, so that the pressure in the turbulent jet region was simply set to be zero.[18–21,28,35,36]

According to these assumptions, the pressure equations can be written as

$$P_1=P_{10}+\mu P_{11},\quad P_2=P_{20}+\mu P_{21}, \tag{4}$$

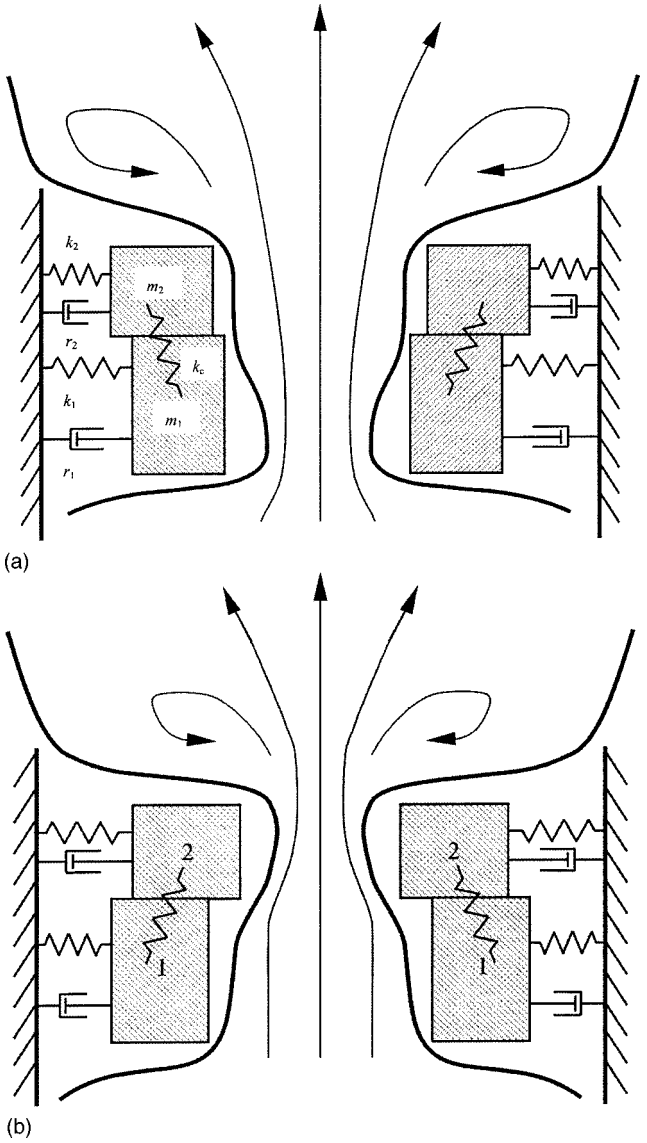


FIG. 1. The flow patterns in the glottis: (a) divergent glottis, where the jet region is in both the upper and lower sections; (b) convergent glottis, where the Bernoulli region is the lower section, and the jet region is the upper.

where P_1 and P_2 represent the pressures acting on m_1 and m_2, respectively. P_{10} and P_{20} describe the steady pressure components in the Bernoulli region, respectively. P_{11} and P_{21} denote the random forces in the turbulent jet region. μ is defined as the region factor having two values: 0 or 1. Clearly, the region factor $\mu=1$ corresponds to the vocal-fold model perturbed by turbulent noise. However, for $\mu=0$, the model will be reduced to the deterministic model with steady airflow.[18–21,28,35,36] Furthermore, P_{10}, P_{11}, P_{20}, and P_{21} can be determined by these two cases.

(i) With the divergent glottis $a_1<a_2$ [see Fig. 1(a)], two masses are both located in the turbulent jet region, which results in

$$P_{10}=P_0=0,\quad P_{11}=\xi_1(t),$$
$$P_{20}=P_0=0,\quad P_{21}=\xi_2(t), \tag{5}$$

where $\xi_1(t)$, $\xi_2(t)$ are Gaussian white noise with mean value 0 and mean squared value σ^2, and satis-

fying $\langle \xi_i(t)\rangle_t = 0$, $\langle \xi_i(t)\xi_i(t')\rangle_t = \sigma^2\delta(t-t')$ $(i=1,2)$. $\langle\cdot\rangle_t$ denotes the average over time t.

(ii) With the convergent glottis $a_1 \geqslant a_2$ [see Fig. 1(b)], mass 1 is within the steady Bernoulli region, but mass 2 is located in the turbulent jet region; then we have

$$P_{10} = P_s\left[1-\Theta(a_2)\left(\frac{a_2}{a_1}\right)^2\right], \quad P_{11}=0,$$
$$P_{20}=P_0=0, \quad P_{21}=\xi_1(t). \tag{6}$$

Combining these equations, we rewrite Eq. (1) as the following vocal-fold model in the presence of turbulent noise:

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x},\mathbf{a}) + \mu\mathbf{U}(\boldsymbol{\xi}), \quad \mathbf{x}\in R^4, \tag{7}$$

where the vector $\mathbf{x}$ is written as $\mathbf{x}=\{x_1,\nu_1,x_2,\nu_2\}^T$, and $\mathbf{a}$ is the parameter vector describing mass, stiffness, damping, and geometrical sizes of the vocal folds. The velocity vectors $\mathbf{v}$ and $\mathbf{U}$ represent the steady and turbulent components, which are described as

$$\mathbf{v}(\mathbf{x},\mathbf{a}) = \begin{pmatrix} \nu_1 \\ \frac{1}{m_1}\left(P_{10}ld_1\Theta(a_1) - r_1\nu_1 - k_1x_1 - \Theta(-a_1)c_1\frac{a_1}{2l} - k_c(x_1-x_2)\right) \\ \nu_2 \\ \frac{1}{m_2}\left(P_{20}ld_2\Theta(a_2) - r_2\nu_2 - k_2x_2 - \Theta(-a_2)c_2\frac{a_2}{2l} - k_c(x_2-x_1)\right) \end{pmatrix},$$

and

$$\mathbf{U}(\xi) = \begin{pmatrix} 0 \\ P_{11}ld_1\Theta(a_1)/m_1 \\ 0 \\ P_{21}ld_2\Theta(a_2)/m_2 \end{pmatrix},$$

respectively. The pressure functions satisfy Eqs. (5) and (6). The noise vector denotes $\boldsymbol{\xi}=\{\xi_1,\xi_2\}^T$. The relative amplitude $\Delta=\sigma/P_s$ is defined. The amplitude of turbulent noise can be estimated from experimental measures, numerical simulations, and speech signal analysis. Alipour and Scherer[29] experimentally separated the turbulent component from the periodic component using the phase-averaging technique. Based on the finite element analysis, Iijima *et al.*[27] calculated the pressure distributions of the two-dimensional unsteady viscous flow from the Bernoulli region to the turbulent jet region in a static glottis. In a glottal waveform model, Childers and Lee[37] applied the typical 0.25% ratio of the turbulent noise amplitude to the glottal waveform amplitude. According to these results, we set the relative amplitudes of stochastic noises at $\Delta \leqslant 1\%$.

III. NUMERICAL CALCULATIONS AND DISCUSSIONS

A. Chaotic vibrations of the vocal-fold model

Chaotic vibrations may occur when the mechanical parameters and tissue geometry of vocal folds, such as stiffness, mass, and glottal rest area, are out of normal ranges.[18–21] Furthermore, in this study, it will be found that turbulent noise is also an important factor in causing chaotic vibrations of vocal folds. A periodic state will be forced into chaos when a certain random force, resulting from turbulent noise, acts on vocal folds. To demonstrate this, we will employ nonlinear dynamic methods including the Poincaré map and Lyapunov exponent that have been found as the effective descriptors for chaotic vocal-fold vibrations.[21] The Poincaré map projects an n-dimensional dynamic flow into an $(n-1)$-dimensional hypersurface. In the Poincaré section, the dynamic types of attractors can be illustrated as follows: a simple periodic orbit appears as a single point; a periodic orbit with finite commensurable frequency components gives rise to some discrete points; and a chaotic attractor is shown as scattered points with a certain fractal structure. To avoid the interpolation error and locate the intersection point accurately, we will use the method suggested by Henon[38] to calculate the Poincaré map after a 1000 ms transient time. The total time of numerical calculation is 2000 ms. For the weak random perturbations, a fourth-order Runge–Kutta method[13] is applied to numerically integrate Eq. (7) with the time step $h=0.001$, and the initial conditions $x_1(0)=0.1$, $\nu_1(0)=0$, $x_2(0)=0.1$, $\nu_2(0)=0$.

In the deterministic vocal-fold model with steady airflow (the region factor $\mu=0$), the following parameters are used to produce an eightfolding-period vibration shown in Fig. 2: $m_1=0.125$, $m_2=0.025$, $r_1=r_2=0.01$, $k_1=0.08$, $k_2=0.008$, $k_c=0.0451$, $c_1=3k_1$, $c_2=3k_2$, $d_1=0.25$, $d_2=0.05$, $a_{01}=a_{0.2}=0.1$, $l=1.4$, $Ps=0.015$, and $\rho=0.001\,13$. All parameters are given in units of cm, g, ms, and their corresponding combinations. The eightfolding-period orbit can be seen as eight discrete points in the Poincaré map [see Fig. 2(a)]. Figure 2(b) illustrates the time evolutions of the Lyapunov exponents where the Lyapunov exponents asymptotically converge to $\lambda_1=0$, $\lambda_2=-0.017$, $\lambda_3=-0.147$, and $\lambda_4=-0.316$. The Kaplan–Yorke dimension can be calculated using the definition $D_L=k+\Sigma_{i=1}^k\lambda_i/|\lambda_{k+1}|$, where k is the largest integer satisfying $\Sigma_{i=1}^k\lambda_i \geqslant 0$.[39] The discrete points in the Poincaré map, the maximal Lyapunov exponent $\lambda_1=0$, and Kaplan–Yorke dimension $D_L=1$ determine the regular and periodic motion of the vocal-fold model with steady airflow under this group of system parameters.

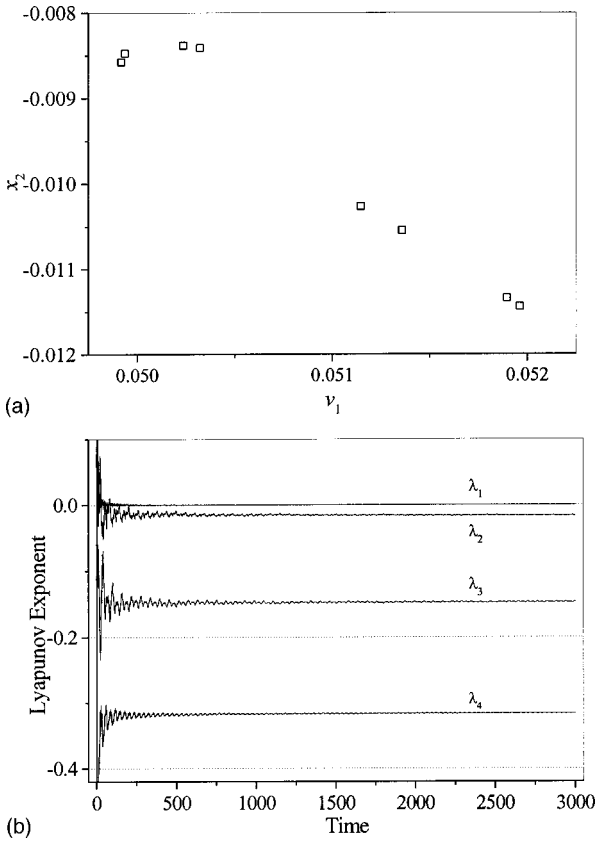


FIG. 2. The periodic vibration of the vocal-fold model with steady airflow ($\mu=0$), where the typical values of parameters are $m_1=0.125$, $m_2=0.025$, $r_1=r_2=0.01$, $k_1=0.08$, $k_2=0.008$, $k_c=0.0451$, $c_1=3k_1$, $c_2=3k_2$, $d_1=0.25$, $d_2=0.05$, $a_{01}=a_{02}=0.1$, $l=1.4$, $\rho=0.001\,13$, and $P_s=0.015$. (a) Poincare map; (b) Lyapunov exponents, where the curves from the top correspond to λ_1, λ_2, λ_3, and λ_4, respectively.

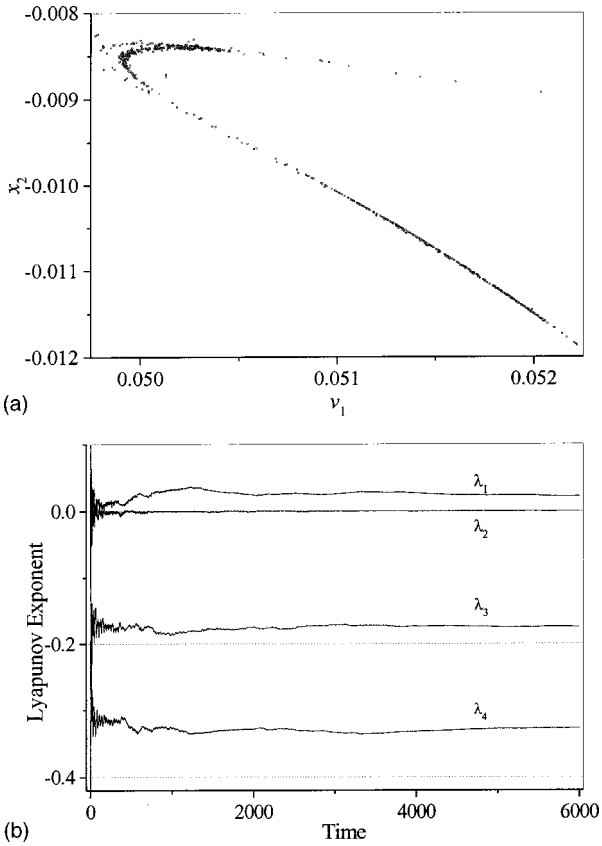


FIG. 3. The chaotic vibration induced by turbulent noise ($\mu=1$). (a) Poincare map; (b) Lyapunov exponents, where the curves from the top correspond to λ_1, λ_2, λ_3, and λ_4, respectively.

However, when considering the effects of turbulent noise ($\mu=1$), vocal-fold vibrations become much more complicated. The output of Eq. (7) will be forced from an eightfolding-periodic motion into an irregular vibratory pattern, even when the weak amplitude $\Delta=0.5\%$ of turbulent noise is applied. The continuous trajectory of this irregular vibration is projected into a group of scattered points with a complex structure in the Poincaré map [see Fig. 3(a)]. The time evolutions of Lyapunov exponents are displayed in Fig. 3(b), where $\lambda_1=0.022$, $\lambda_2=0$, $\lambda_3=-0.175$, $\lambda_4=-0.327$, and the corresponding Kaplan–Yorke dimension can be calculated as $D_L=2.13$. The chaotic attractor shows sensitivity to initial conditions. The error will be amplified exponentially according to the rule described by the positive Lyapunov exponent. The calculated positive Lyapunov exponent λ_1 demonstrates the chaotic nature of this irregular vibratory pattern. Therefore, turbulent noise increases the complexity of the vocal-fold vibrations, and produces a qualitative change to the dynamic behavior of vocal folds, such as inducing chaos from periodic vibrations.

B. Bifurcation diagram

Different dynamic solutions, such as fix point, period, and chaotic attractor, can be shown in a bifurcation diagram by continuously changing a system parameter. Generally, the coupling stiffness k_c was used as a variable parameter in modeling nonlinear vocal-fold vibrations.[18,35,40] The above results examine the effect of turbulent noise on a specific eightfolding-periodic state. To investigate the effects of turbulent noise on different dynamic solutions, the bifurcation diagram will be used. When calculating a bifurcation diagram, we hold the coupling stiffness k_c constant for 2000 ms, and calculate the Poincaré map after the 1000 ms transient time. Subsequently, the parameter k_c increases incrementally to produce new groups of model output. In this way, the entire bifurcation diagram can be obtained.

1. Turbulent noise

The deterministic vocal-fold model with steady airflow is a special case of Eq. (7) when setting the region factor $\mu=0$. Figure 4(a) illustrates the bifurcation diagram of the deterministic model. To reveal the detailed period-doubling bifurcation characteristics, an enlargement of the region $k_c=0.0443$ to $0.045\,15$ of Fig. 6 is shown in Fig. 4. Here, a fine period-doubling bifurcation cascade can be found in such an enlarged zone of parameter space, which is consistent with the result presented by Herzel and Knudsen.[18] All threshold points of period-doubling bifurcations in Fig. 4(b) can be estimated using the relation $\lambda_2=0$. When $k_c>0.044\,48$, the limit cycle with period $1/f_0$ will become unstable; and a stable period-doubling oscillation with period $2/f_0$ sets in instead. Next, a period-doubling bifurcation series of period $2^n/f_0$ ($n=2,3,...$) appears by increasing k_c. These regular

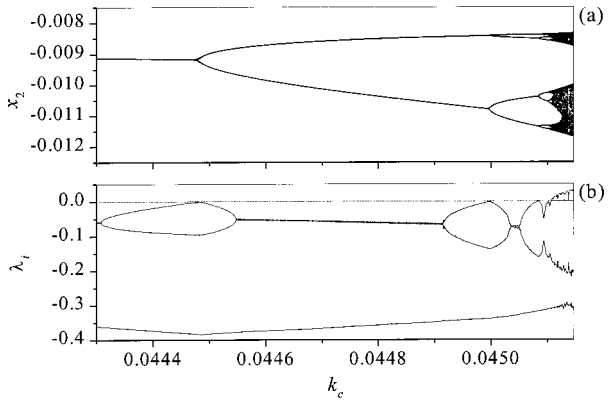


FIG. 4. The deterministic vocal-fold model with steady airflow ($\mu=0$). (a) Bifurcation diagram; (b) Lyapunov exponents λ_i versus coupling stiffness k_c, where the curves from the top correspond to λ_1, λ_2, λ_3, and λ_4, respectively.

vibratory patterns, including limit cycle and period-doubling series, are defined within the parameter intervals corresponding to $\lambda_1=0$. However, when $k_c>0.04511$, $\lambda_1>0$, and $\lambda_2=0$ are yielded, such that the vibrations of the vocal-fold model become irregular, then chaos with one positive Lyapunov exponent occurs.

When the contribution of turbulent noise is included ($\mu=1$), bifurcation diagrams show that turbulent noise broadens the intervals of chaotic vibrations. With the increase of k_c, Fig. 5 illustrates the effects of turbulent noise on the period-doubling bifurcation series where the noise amplitude is $\Delta=0.5\%$. Here, the corresponding periodic solutions with high periodicities $2^n/f_0$ ($n\geqslant 2$) in Fig. 4 are replaced by irregular patterns in the presence of turbulent noise. A typical example of such a change has been presented in Sec. III A, where noise-induced dynamics has been studied at the specific value of $k_c=0.0451$. Some period-doubling bifurcation details become obscure and even unobservable in the bifurcation diagram, Fig. 5(a). The period-eightfolding and period-quadrupling windows disappear, and, instead, the regions of noise-induced chaos appear. Noise inhibits the fine period-doubling bifurcations. Based on the calculated Lyapunov exponents in Fig. 5(b), the threshold value of k_c from period-doubling bifurcation to

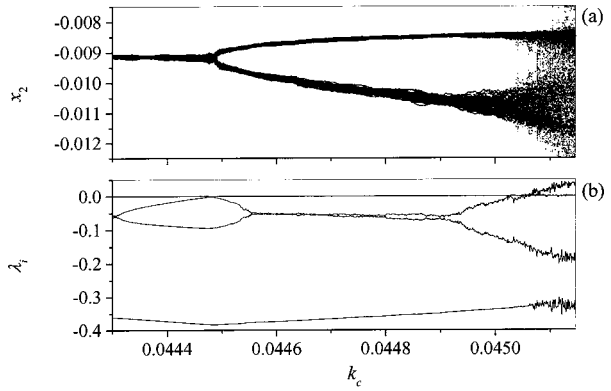


FIG. 5. The vocal-fold model with turbulent noise ($\mu=1$), where the amplitude of turbulent noise is $\Delta=0.5\%$. (a) Bifurcation diagram; (b) λ_i vs k_c, where the curves from the top correspond to λ_1, λ_2, λ_3, and λ_4, respectively.

chaos is about 0.045 06, which is lower than the value of 0.045 11 for steady airflow ($\mu=0$). Here, the small increase of the threshold value of k_c results from the 0.5% weak amplitude of turbulent noise. When the amplitude of turbulent noise is further increased, the threshold value of k_c will be correspondingly decreased. When the coupling stiffness is above the threshold value, the chaotic vibrations with $\lambda_1>0$ will be reached. Therefore, turbulent noise lowers the threshold value of the system parameters, inhibits period-doubling bifurcations, and broadens the intervals of irregular vibration in the bifurcation diagrams.

The influence of turbulent noise on the vibratory patterns of the vocal-fold model is important. When the system is unperturbed, the trajectories will move within their own basins of attraction in the regular fashion. When turbulent noise acts as a kind of random force, the trajectories may randomly wander around the periodic orbits. For large noise perturbations, the trajectories will even escape from the basins of the regular attraction, and then reside within the nearby chaotic attractors, which results in the qualitative changes of dynamic behaviors. Generally, the manifolds of chaotic attractors are insensitive or robust to external perturbations, and then the trajectories will stay within the manifolds. Therefore, turbulent noise may decrease the threshold values of the system parameters to induce chaotic vibrations.

2. The stiffness perturbation

Turbulent noise takes the effect of an external aerodynamic force. On the other hand, the random perturbations of system parameters that come from neurological and biomechanical effects constitute the internal random forces of vocal folds.[25,41,42] Some neurological fluctuations such as muscle twitches cause unsteadiness and randomness of the fundamental frequency. The proportion of the frequency perturbation (jitter) to the tension perturbation determines the random characteristics of the vocal-fold stiffness perturbation. The pulsatile blood flows and unsteady lung pressure will also introduce the additional random perturbations to biomechanical parameters of vocal folds. Considering these random perturbation factors, the dynamic equation of the vocal-fold model can be further written as

$$\frac{d\mathbf{x}}{dt}=\mathbf{v}(\mathbf{x},\mathbf{a}')+\mu\mathbf{U}(\xi),\quad \mathbf{x}\in R^4,\quad \text{and}\quad \mathbf{a}'=\mathbf{a}+\tilde{\mathbf{a}},\quad (8)$$

where $\tilde{\mathbf{a}}$ denotes the random perturbation to the default parameter $\mathbf{a}$. For the random stiffness perturbation, $\mathbf{a}=\{k_1,k_2,k_c\}^T$ gives the normal stiffness vector, and $\tilde{\mathbf{a}}$ is the random component satisfying $\tilde{\mathbf{a}}=\Delta'\mathbf{a}$, and Δ' is a relative ratio.

To investigate the influence of the vocal-fold stiffness perturbation exclusively, the bifurcation diagram for steady airflow ($\mu=0$) is shown. Figures 6(a) and (b) correspond to the stiffness perturbation $\Delta'=0$ and $\Delta'=0.1\%$, respectively. Differing from the fine period-doubling bifurcation of the deterministic model in Fig. 6(a), random stiffness perturbation, just as turbulent noise, destroys the fine structure of the attractors, inhibits the period-doubling bifurcation cascade, and broadens the intervals of irregular vibration in the bifurcation diagram [see Fig. 6(b)]. From the nonlinear dy-

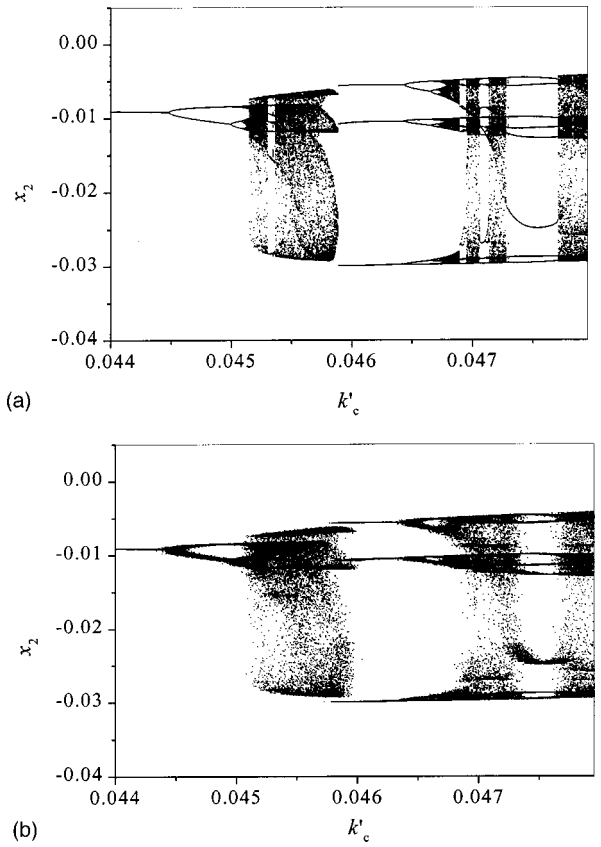


FIG. 6. The influence of the random stiffness perturbation. (a) Bifurcation diagram of the deterministic model where $\mu=0$, and the amplitude of the stiffness perturbation $\Delta'=0$; (b) bifurcation diagram of the vocal-fold model in presence of random stiffness perturbation where $\mu=0$, and $\Delta'=0.1\%$.

namic point, when the system parameters are randomly perturbed, chaotic systems correspondingly exhibit dynamic changes. The dynamic state will frequently be switched between different attractors since the manifolds of dynamic systems will be completely determined by the system parameters. For a sufficiently large random perturbation to system parameters, particularly to those parameters whose values are close to the threshold value for chaos, trajectories will easily be led to nearby chaotic attractors and will stably stay within the manifolds of chaotic attractors.

Random perturbations including turbulent noise and random stiffness perturbation will qualitatively change nonlinear dynamic behaviors of vocal folds, such as inducing chaos and inhibiting fine period-doubling bifurcations. For a noise-free deterministic system ($\mu=0$ and $\Delta'=0$), a period-doubling series is illustrated in the spectrogram of Fig. 7(a). However, in the presence of stochastic noises, high-period bifurcations will be inhibited. Figure 7(b) shows the spectrogram of the vocal-fold model perturbed by the turbulent noise and random stiffness perturbation k'_c, where the coupling stiffness k_c increases linearly with time. The length of the voice sample is 26.5 s, and the amplitudes Δ, Δ' of turbulent noise and stiffness perturbation are 0.5%, and 0.1%, respectively. The $f_0/2$ subharmonic frequency can be seen in this spectrogram. The higher period-doubling bifurcations, including $f_0/4$, $f_0/9$, and $f_0/10$, disappear, and instead, irregular regions are broadened. When the amplitudes

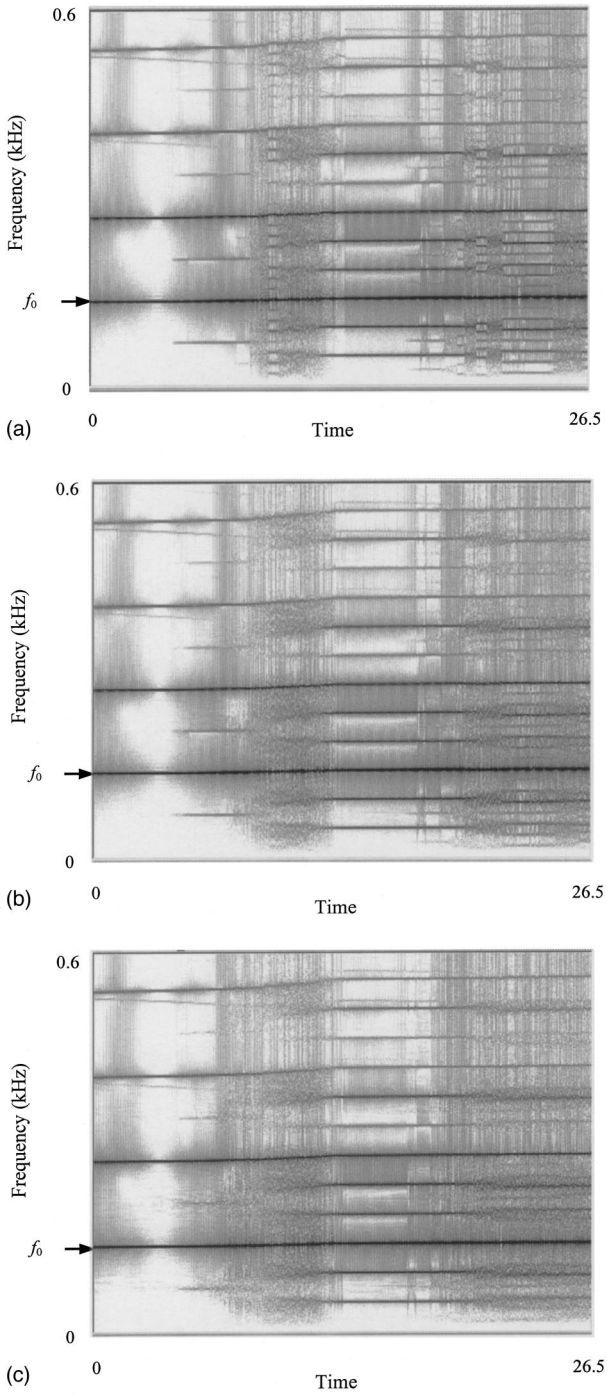


FIG. 7. The spectrograms of vocal-fold vibrations. (a) The noise-free deterministic model where $\mu=0$, $\Delta'=0$; (b) the model with turbulent noise and random stiffness perturbation where $\mu=1$, $\Delta=0.5\%$, and $\Delta'=0.1\%$, respectively; (c) the model with turbulent noise and random stiffness perturbation where $\mu=1$, $\Delta=1\%$, and $\Delta'=1\%$, respectively.

of noises are further increased to $\Delta=\Delta'=1\%$, the bifurcation solutions with frequency components $f_0/6$ and $f_0/8$ are both inhibited [see Fig. 7(c)]. The irregular regions will be further broadened. The noise-inhibited period-doubling bifurcation reveals that only a few period-doubling bifurcation series, such as f_0, $f_0/2$, and $f_0/3$, can be seen in the spectrogram. This phenomenon has also been experimentally observed in the excised larynx by varying the asymmetry of

vocal fold elongation,[22] or by increasing the subglottal pressure into the region of above range phonation.[21,24]

IV. CONCLUSION

Based on the vocal-fold model in the presence of turbulent noise, it was found that turbulent noise might qualitatively change the dynamic behaviors, such as inducing chaos from periodic vibrations. Nonlinear dynamic methods, including the Poincaré map and Lyapunov exponent, demonstrated the chaotic nature of irregular vibrations. To investigate the influences of turbulent noise and other random perturbations on different attractors, bifurcation diagrams were calculated. Turbulent noise as well as the random stiffness perturbation inhibited the fine period-doubling bifurcations and broadened the intervals of irregular vibrations in bifurcation diagrams. Chaos induced by external noise had been observed in the semiconductor laser model.[8,9] Our research presented the evidence of this interesting phenomenon in the vocal-fold model. The weak intensities of random noises to vocal fold parameters were used, i.e., $\Delta \leqslant 1\%$. However, they exerted important effects on producing irregular vibration of vocal folds. Noise-induced chaos and noise-inhibited period-doubling bifurcation might not be typically revealed in deterministic models of vocal folds. The study of noise-induced dynamics might present a valuable approach to simulate irregular vocal fold vibrations in the presence of random aerodynamic, neurological, and biomechanical factors.

ACKNOWLEDGMENT

This study was supported by NIH Grant No. (1-RO1DC05522-01) from the National Institute of Deafness and other Communication Disorders.

[1] F. T. Arecchi, R. Badii, and A. Politi, "Generalized multistability and noise-induced jumps in a nonlinear dynamical system," Phys. Rev. A **32**, 402–408 (1985).
[2] P. D. Beale, "Noise-induced escape from attractors in one-dimensional map," Phys. Rev. A **40**, 3998–4003 (1989).
[3] M. Diestelhorst, R. Hegger, L. Jaeger, H. Kantz, and R. P. Kapsch, "Experimental verification of noise induced attractor deformation," Phys. Rev. Lett. **82**, 2274–2277 (1999).
[4] N. G. Stocks, R. Mannella, and P. V. E. McClintock, "Influence of random fluctuations on delayed bifurcations: The case of additive white noise," Phys. Rev. A **40**, 5361–5369 (1989).
[5] Z. Y. Chen, "Noise-induced instability," Phys. Rev. A **42**, 5837–5843 (1990).
[6] L. Arnold, H. Crauel, and V. Wihstutz, "Stabilization of linear system by noise," SIAM J. Control Optim. **21**, 451–461 (1983).
[7] J. P. Crutchfield and B. A. Huberman, "Fluctuations and the onset of chaos," Phys. Lett. A **77**, 407–410 (1980).
[8] J. B. Gao, S. K. Hwang, and J. M. Liu, "When can Noise induce chaos?," Phys. Rev. Lett. **82**, 1132–1135 (1999).
[9] S. K. Hwang, J. B. Gao, and J. M. Liu, "Noise-induced chaos in an optically injected semiconductor laser model," Phys. Rev. E **61**, 5162–5170 (2000).
[10] K. Matsumoto and I. Tsuda, "Noise-induced order," J. Stat. Phys. **31**, 87–106 (1983).
[11] H. Herzel and B. Pompe, "Effects of noise on a nonuniform chaotic map," Phys. Lett. A **122**, 121–125 (1987).
[12] B. McNamara and K. Wiesenfeld, "Theory of stochastic resonance," Phys. Rev. A **39**, 4854–4869 (1989).
[13] J. A. Blackburn, N. Gronbech-Jensen, and H. J. T. Smith, "Stochastic noise and chaotic transients," Phys. Rev. Lett. **74**, 908–911 (1995).
[14] I. Hidaka, D. Nozaki, and Y. Yamamoto, "Functional stochastic resonance in the human brain: noise induced sensitization of Baroreflex system," Phys. Rev. Lett. **85**, 3740–3743 (2000).
[15] 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.
[16] W. Mende, H. Herzel, and K. Wernke, "Bifurcation and chaos in newborn infant cries," Phys. Lett. A **145**, 418–424 (1990).
[17] D. A. Berry, H. Herzel, I. R. Titze, and K. Krischer, "Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions," J. Acoust. Soc. Am. **95**, 3595–3604 (1994).
[18] H. Herzel and C. Knudsen, "Bifurcation in a vocal fold model," Nonlinear Dyn. **7**, 53–64 (1995).
[19] I. Steinecke and H. Herzel, "Bifurcation in an asymmetric vocal-fold model," J. Acoust. Soc. Am. **97**, 1874–1884 (1995).
[20] H. Herzel, D. Berry, I. R. Titze, and I. Steinecke, "Nonlinear dynamics of the voice: Signal analysis and biomechanical modeling," Chaos **5**, 30–34 (1995).
[21] J. J. Jiang, Y. Zhang, and J. Stern, "Modeling of chaotic vibrations in symmetric vocal folds," J. Acoust. Soc. Am. **110**, 2120–2128 (2001).
[22] D. A. Berry, H. Herzel, I. R. Titze, and B. H. Story, "Bifurcation in excised larynx experiments" J. Voice **10**, 129–138 (1996).
[23] J. G. Svec, H. K. Schutte, and D. G. Miller, "On pitch jumps between chest and falsetto registers in voice: Data from living and excised human larynges," J. Acoust. Soc. Am. **106**, 1523–1531 (1999).
[24] J. J. Jiang, J. R. Raviv, and D. G. Hanson, "Quantitative measurement of unstable phonation," presented at the Voice Foundation Meeting, Philadelphia, PA, June 2000.
[25] I. R. Titze, *Principles of Voice Production* (Prentice–Hall, Englewood Cliffs, NJ, 1994), p. 279.
[26] J. Liljencrants, "Numerical simulations of glottal flow," in *Proceedings of EUROSPEECH*, Genova, 1991, pp. 255–258.
[27] H. Iijima, N. Miki, and N. Nagai, "Glottal impedance based on a finite element analysis of two-dimensional unsteady viscous flow in a static glottis," IEEE Trans. Signal Process. **40**, 2125–2135 (1992).
[28] X. Pelorson, A. Hirschberg, R. R. van Hassel, A. P. J. Wijnands, and Y. Auregan, "Theoretical and experimental study of steady-flow separation within the glottis during phonation. Application to a modified two-mass model," J. Acoust. Soc. Am. **96**, 3416–3431 (1994).
[29] F. Alipour and R. C. Scherer, "Pulsatile airflow during phonation: An excised larynx model," J. Acoust. Soc. Am. **97**, 1241–1248 (1995).
[30] L. Mongeau, N. Franchek, C. H. Coker, and R. A. Kubli, "Characteristics of a pulsating jet through a small modulated orifice, with application to voice production," J. Acoust. Soc. Am. **102**, 1121–1133 (1997).
[31] A. Barney, C. H. Shadle, and P. O. A. L. Davies, "Fluid flow in a dynamic mechanical model of the vocal folds and tract. I. Measurements and theory," J. Acoust. Soc. Am. **105**, 444–455 (1999).
[32] M. T. Landahl and E. Mollo-Christensen, *Turbulence and Random Processes in Fluid Mechanics* (Cambridge University Press, Cambridge, 1992).
[33] M. J. Lighthill, "On sound generated aerodynamically. I. General theory," Proc. R. Soc. London, Ser. A **211**, 564–587 (1952).
[34] T. Baer, "Investigation of the phonatory mechanism," ASHA Report **11**, 1981, pp. 38–47.
[35] B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. **97**, 1249–1260 (1995).
[36] J. C. Lucero, "A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset," J. Acoust. Soc. Am. **105**, 423–431 (1999).
[37] D. G. Childers and C. K. Lee, "Vocal quality factors: Analysis, synthesis, and perception," J. Acoust. Soc. Am. **90**, 2394–2410 (1991).
[38] M. Henon, "On the numerical computation of Poincare maps," Physica D **5**, 412 (1982).
[39] J. L. Kaplan and J. A. Yorke, "Chaotic behavior of multidimensional difference equations," in *Functional Differential Equations and Approximation of Fixed Point*, Lecture Notes in Mathematics 730, edited by H.-O. Peitgen and H.-O. Walther (Springer-Verlag, Springer-Berlin, 1979), p. 204.
[40] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Syst. Tech. J. **51**, 1233–1268 (1972).
[41] I. R. Titze, "A model for neurologic sources of aperiodicity in vocal fold vibration," J. Speech Hear. Res. **34**, 460–472 (1991).
[42] R. F. Orlikoff and R. J. Baken, "The effect of the heartbeat on fundamental frequency perturbation," J. Speech Hear. Res. **35**, 576–583 (1989).

Computational aeroacoustics of phonation, Part I: Computational methods and sound generation mechanisms

Wei Zhao,[a] Cheng Zhang,[b] Steven H. Frankel,[c] and Luc Mongeau
School of Mechanical Engineering, Purdue University, West Lafayette, Indiana 47907

(Received 25 October 2001; revised 3 July 2002; accepted 19 July 2002)

The aerodynamic generation of sound during phonation was studied using direct numerical simulations of the airflow and the sound field in a rigid pipe with a modulated orifice. Forced oscillations with an imposed wall motion were considered, neglecting fluid–structure interactions. The compressible, two-dimensional, axisymmetric form of the Navier–Stokes equations were numerically integrated using highly accurate finite difference methods. A moving grid was used to model the effects of the moving walls. The geometry and flow conditions were selected to approximate the flow within an idealized human glottis and vocal tract during phonation. Direct simulations of the flow and farfield sound were performed for several wall motion programs, and flow conditions. An acoustic analogy based on the Ffowcs Williams–Hawkings equation was then used to decompose the acoustic source into its monopole, dipole, and quadrupole contributions for analysis. The predictions of the farfield acoustic pressure using the acoustic analogy were in excellent agreement with results from the direct numerical simulations. It was found that the dominant sound production mechanism was a dipole induced by the net force exerted by the surfaces of the glottis walls on the fluid along the direction of sound wave propagation. A monopole mechanism, specifically sound from the volume of fluid displaced by the wall motion, was found to be comparatively weak at the frequency considered (125 Hz). The orifice geometry was found to have only a weak influence on the amplitude of the radiated sound. [DOI: 10.1121/1.1506693]

PACS numbers: 43.70.Bk [AL]

I. INTRODUCTION

The discipline of Computational Aeroacoustics (CAA) involves the numerical simulation of the generation and radiation of sound by unsteady flows. The nature, characteristics, and objectives of computational aeroacoustics problems are quite different from those commonly encountered in computational fluid dynamics (CFD) problems that do not require the calculation of the radiated sound. The quantities of interest in CAA problems include the amplitude and the directivity of the sound pressure radiated to the farfield. To ensure that the computed solution is uniformly accurate over long propagation distances from the sound source to the boundary of the computational domain, the numerical scheme must be almost free of numerical dispersion, dissipation, and anisotropy.[1] The application of CAA to study human phonation has the potential to address several important questions related to the precise relation between fluid dynamic and acoustic phenomena in laryngeal flows.

It is well known that during human speech a pulsating jet is generated through the glottis by self-sustained oscillations of the human vocal cords.[2] The glottis is a slit-like orifice between the vocal cords, which is located in the airway upstream of the vocal tract. A typical sequence of the approximate motion of the vocal cords (idealized for the purposes of this discussion) is sketched in Fig. 1. Initially, the glottis is nearly closed (A). Air at high pressure is supplied from the lungs and forces the vocal folds apart. A starting jet forms [(B) and (C)], and the pressure around the vocal cords decreases. Elastic forces and lower pressures then force the glottis to close again (D), and the cycle repeats. Again, the geometry shown in Fig. 1 is highly idealized, and is intended to show the basic features of the orifice shape at different times over one cycle. During any given cycle the glottis takes on complex converging, uniform, and diverging shapes. The pulsating jet, shown here as laminar, although it is, in reality, turbulent downstream of the orifice, features vortices shed from the orifice trailing edge. The acoustic waves generated by the pulsating jet flow constitute the main acoustic source for speech. They acoustically excite the cavities formed by the vocal tract, such as the mouth and the nasal cavity, before being radiated from the lips. The interactions of the flow field with the vocal cords, whose shape varies over any oscillation period, and the radiated sound have an important impact on voice production efficiency and quality.[2] The exact nature of the flow–sound interactions, however, is not very well understood. The possible contributions of the wall motion itself also need to be elucidated.

Several previous studies have motivated the present work. McGowan[3] pursued the use of an acoustic analogy approach to study the sources of sound during phonation. He postulated that, in addition to an unsteady volume velocity monopole source associated with the pulsating jet, an oscillating force may arise from the interaction between vortical structures shed from the glottis and the velocity field itself.

[a]Currently employed at Reaction Engineering International, Salt Lake City, Utah.
[b]Graduate student.
[c]Author to whom correspondence should be addressed. Electronic mail: frankel@ecn.purdue.edu

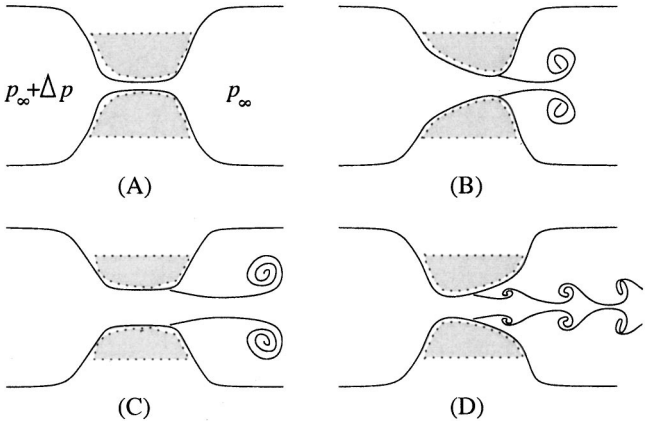


FIG. 1. A sketch illustrating an idealized time sequence of motion of the human glottis.

This was speculated to cause dipole acoustic sources. Hofmans[4] experimentally, analytically, and numerically studied vortex-generated sound in confined ducts. Impulsively started flows through various mechanical models combining nozzles and orifices were studied. A very complex behavior of interacting vortices due to the jet instability was observed. Frequencies of pressure perturbations associated with the vortex pairing/merging processes were in the audible range, and were postulated to contribute to human speech significantly. Hofmans[4] did not attempt to directly solve for the acoustic field numerically.

To investigate laryngeal flows, rigid static laryngeal models have been widely used both in experimental and numerical studies. Iijima *et al.*[5] numerically investigated the effects of glottal shape on flow characteristics and suggested that in order to study the effects of vorticity, dynamic simulations should be employed. A few pulsating flow simulations have also been reported for different laryngeal models.[6,7]

Attempts have also been made to experimentally investigate the flow through a dynamic mechanical model of the vocal cords. Mongeau *et al.*[8] measured the velocity profile development near the orifice in a dynamic mechanical model of the larynx. The purpose was to investigate the validity of the quasisteady assumption often utilized in so-called multimass models of phonation. Shadle *et al.*[9] and Barney *et al.*[10] made aerodynamic and acoustic measurements in a dynamic mechanical model of the larynx and the vocal tract. A pair of sinusoidally oscillating shutters modulated an otherwise steady air duct flow. They postulated that the interaction of vortices with the exit of the duct might be an important sound source in speech production.

In a previous study, the flow and sound radiation in an idealized *rigid model* of the human vocal tract was investigated numerically.[11] A *time-varying* volume flow rate was imposed at the inlet to generate a pulsating flow, with the cross-sectional area of the glottal orifice fixed. Three different geometric models, featuring either a uniform, diverging, or converging glottis shape, were considered to study the influence of flow separation on the flow field and the radiated sound. It was shown that the included angle of the glottis had a significant effect on the quadrupole sound source contribution. One important limitation of those simulations was the stationary rigid wall assumption. As mentioned, during actual speech production, the vocal cords are self-oscillating, resulting in a time-varying geometry and cross-sectional area of the glottis. The potential effects of generic wall motion on the flow and on the acoustic field were the focus of the present investigation.

In the present study, a numerical model of an idealized vocal tract with a time-varying geometry was developed. The model geometry and the flow field were assumed to be axisymmetric. The vocal fold motion was imposed, and so only forced oscillations were considered in this study. This is a logical first step toward the development of a self-oscillating (flow-induced) model accounting for fluid–structure interactions. A grid system moving with the wall boundary was employed to discretize the computational domain. The compressible Navier–Stokes equations were integrated in a moving axisymmetric coordinate system using a high-order finite-difference method in order to capture both the vortical flow structure as well as the sound radiated in the farfield. An acoustic analogy based on the Ffowcs Williams–Hawkings equation[12] was then applied to decompose the farfield sound into contributions from monopole, dipole, and quadrupole sound generation mechanisms.

In this paper we describe the approach and address questions of numerical accuracy, the effects of geometry on the flow and sound sources, as well as the accuracy of the acoustic analogy. The results of a parametric study examining the influence of subglottal pressure, of the glottal oscillation frequency, and of the presence of ventricular folds are presented in a separate paper.[13]

II. COMPUTATIONAL METHODS

A. Background

In CFD, the Navier–Stokes equations (NSE) are numerically integrated, subject to initial and boundary conditions, on a prespecified computational domain. The exact form of the equations, boundary conditions, and the domain depends upon the problem being studied and the information desired.

For example, for low-speed viscous flow, such as in phonation, the incompressible form of the NSE may be considered appropriate. The flow in the vocal tract can then be modeled as incompressible viscous flow, since the Mach number (the ratio of flow velocity to sound speed) is low (less than 0.1), and so all velocities are small compared to the speed of sound. The NSE for incompressible flow effectively assume an infinite sound speed. This precludes the simulation of acoustic waves and hence the direct simulation of flow-induced sound radiation. Hence, an acoustic analogy (to be described) is required to predict sound radiation in a CAA analysis based on an incompressible flow solution.

An alternative method is to solve the compressible form of the NSE, where both the flow and acoustic waves are simulated directly. No acoustic analogy is needed in this approach, assuming the computational domain can be extended into the farfield. The disadvantage of this approach for low Mach number flows is that most algorithms designed for compressible flows are inefficient or inaccurate for these

flow conditions.[14] This can be addressed with preconditioning techniques or semi-implicit formulations, but their efficacy for CAA has yet to be assessed. Here we handle this problem by another approach, i.e., artificially increase the Mach number of the flows to be simulated. Possible effects of the increased Mach number on the flow field and the acoustic radiation have been discussed in Zhao *et al.*[11]

Turbulence is widely believed to be a feature of the downstream glottal jet.[15] It is also important in frication. There are three main approaches to predicting turbulent flows. Direct numerical simulation or DNS involves numerical integration of the NSE on a grid fine enough to resolve all the scales of motion from the largest scales imposed by the vocal tract geometry to the smallest scales associated with viscous dissipation. Sufficiently accurate numerical methods with low dissipation and dispersion errors are required to simulate such a broad range of scales. The computational cost of DNS in terms of the number of grid points and time steps for geometries and flow conditions relevant to phonation makes the DNS approach prohibitive with current and near-term future computational resources. Thus, the approach is limited to simple geometries and low flow rates. A fully three-dimensional simulation of the unsteady flow and acoustic fields inside a realistic vocal tract model is prohibitive.

The Reynolds-averaged Navier–Stokes, or RANS, method involves ensemble averaging the NSE over all length scales of the flow. Only the mean (time-averaged) flow is predicted. The effects of turbulence on the mean flow are incorporated using a turbulence model. Mean flow predictions are far easier to perform than instantaneous, turbulent flow predictions. They require fewer grid points than DNS, and hence they are much more computationally affordable. Because these models must predict effects associated with large- and small-scale motions, they are sensitive to the geometry and flow conditions and must often be adjusted for each problem studied. They may be inaccurate for separated turbulent flows, such as the flow downstream of the glottis during phonation.

A compromise between DNS and RANS, large eddy simulations, or LES, involve the direct computation of large-scale flow motions, as in DNS. But unlike for DNS, the effects of the small scales on the large scales are modeled using a subgrid-scale (SGS) model. Since the small scales are believed to be universal (i.e., insensitive to the gross flow characteristics) in their behavior, SGS models are expected to be more robust for a variety of flows than the turbulence models used in RANS calculations. The advantage of LES over RANS is that estimates of both the instantaneous, albeit large-scale, flow and the mean flow are obtained. Flows that feature large-scale unsteadiness, such as the highly separated supraglottal flow, are expected to be better predicted by LES than by RANS. LES do require accurate numerical schemes and a sufficiently fine grid. Hence, they are more costly than RANS, but much more affordable than DNS.

In the present study, the approach of DNS was selected. The main simplification made was to consider the flow to be two-dimensional axisymmetric. This precludes the vortex stretching mechanism which is necessary to produce turbulence. Hence, the results from the simulations are laminar. Instantaneous vortical structures associated with the glottal jet, which are responsible for a portion of the generated acoustic signal were captured. The neglect of turbulence was expected to exaggerate the spatial coherence and strength of these vortical structures. It is well known that turbulence enhances mixing, and causes a breakdown of coherent vortical flow structures into disorganized, spatially random flow fluctuations. The consequences of ignoring these effects were pondered, and are the object of discussion in subsequent sections.

Planar flow models have also been used in glottal flow simulations.[5] Those models certainly provided a good approximation of geometry and flow near the glottis region. However, axisymmetric models allow the specification of the actual cross-sectional area ratio between the glottis and the vocal tract. The flow is better approximated as axisymmetric rather than planar far downstream of the glottis.

B. Equations of motion

The compressible Navier–Stokes equations in a cylindrical coordinate system were considered. Under axisymmetric and no-swirl (circumferential velocity $u_\phi = 0$) assumptions, these can be written as

$$\frac{\partial U}{\partial t} + \frac{\partial E}{\partial x} + \frac{\partial F}{\partial r} + \frac{G}{r} = 0, \tag{1}$$

where

$$U = [\rho, \rho u_x, \rho u_r, e]^T, \tag{2}$$

$$E = \begin{bmatrix} \rho u_x \\ \rho u_x^2 + p - \sigma_{xx} \\ \rho u_x u_r - \sigma_{xr} \\ (e+p)u_x - \sigma_{xx}u_x - \sigma_{xr}u_r + Q_x \end{bmatrix}, \tag{3}$$

$$F = \begin{bmatrix} \rho u_r \\ \rho u_x u_r - \sigma_{xr} \\ \rho u_r^2 + p - \sigma_{rr} \\ (e+p)u_r - \sigma_{xr}u_x - \sigma_{rr}u_r + Q_r \end{bmatrix}, \tag{4}$$

$$G = \begin{bmatrix} \rho u_r \\ \rho u_x u_r - \sigma_{xr} \\ \rho u_r^2 - \sigma_{rr} \\ (e+p)u_r - \sigma_{xr}u_x - \sigma_{rr}u_r + Q_r \end{bmatrix}, \tag{5}$$

$$e = \frac{1}{2}\rho(u_r^2 + u_x^2) + \frac{1}{\gamma - 1}p. \tag{6}$$

In these equations ρ is the density, p is the pressure, u_x and u_r are the velocity components in the streamwise and radial directions, respectively. Here σ_{ij} is the laminar viscous stress tensor and Q_j is the heat conduction vector. Formulas for these can be found in any standard fluid dynamics or CFD textbook.[14] In the above, a calorically perfect gas is assumed.

A curvilinear coordinate system in the $x-r$ plane was employed. This allows the generation of a grid with a time-varying geometry that resembles that of the human glottis. The transformed coordinates (ξ, ζ, τ) were obtained from the cylindrical coordinate system (x,r,t) using

$$\xi = x, \tag{7}$$

$$\zeta = \frac{r}{r_w(x,t)}, \tag{8}$$

$$\tau = t, \tag{9}$$

where $r_w(x,t)$ is a function describing the radius of the duct, defined in Sec. III. Note that the coordinate ζ is also a function of time, t, because of the vocal fold motion, whereas ξ is not. The compressible Navier–Stokes equations were then transformed to a new coordinate system, which take the form

$$\frac{1}{J}\frac{\partial U}{\partial \tau} + \frac{1}{J}\frac{\partial U}{\partial \zeta}\zeta_t + \frac{\partial E'}{\partial \xi} + \frac{\partial F'}{\partial \zeta} + \frac{1}{J}\frac{G}{r} = 0, \tag{10}$$

where

$$E' = \frac{E\xi_x + F\xi_r}{J}, \tag{11}$$

$$F' = \frac{E\zeta_x + F\zeta_r}{J}. \tag{12}$$

J is the Jacobian of the coordinate transformation, and ζ_t, ξ_x, ξ_y, etc. are the components of the transformation matrix, $\partial(\xi,\zeta,\tau)/\partial(x,r,t)$. The second term in Eq. (10) is due to the moving coordinate system.

C. Numerical schemes

A sixth-order compact finite-difference scheme[16] was used to approximate the first-order derivatives at the interior grid points (away from the boundaries). It can be written as

$$\frac{1}{3}f'_{j-1} + f'_j + \frac{1}{3}f'_{j+1} = \frac{7}{9\,\Delta\xi}(f_{j+1} - f_{j-1}) + \frac{1}{36\,\Delta\xi}(f_{j+2} - f_{j-2}), \tag{13}$$

where f'_j is the approximation of $\partial f/\partial\xi$ at point j and $\Delta\xi$ is the mesh size. For grid points at the boundary and adjacent to the boundary, the following schemes were used, respectively:

$$f'_1 + 2f'_2 = \frac{1}{2\,\Delta\xi}(-5f_1 + 4f_2 + f_3), \tag{14}$$

$$\frac{1}{4}f'_1 + f'_2 + \frac{1}{4}f'_3 = \frac{3}{4\,\Delta\xi}(f_3 - f_1). \tag{15}$$

The second-order derivatives in the diffusion terms were discretized by using the above compact finite-difference schemes twice. The standard fourth-order Runge–Kutta method was used for explicit time marching. At each substage of the Runge–Kutta scheme the computational grids and the transformation Jacobians were updated to ensure fourth-order accuracy in time. A sixth-order compact filter[16] was periodically applied to the dependent variables, U, to eliminate spurious numerical waves. This filter takes the form

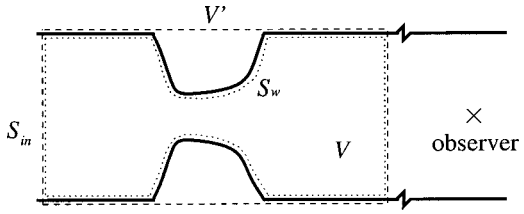


FIG. 2. Sketch of the integral domains and surfaces involved in the integral of the Ffowcs Williams–Hawkings equation for the acoustic analogy predictions.

$$\beta\overline{f}_{j-2} + \alpha\overline{f}_{j-1} + \overline{f}_j + \alpha\overline{f}_{j+1} + \beta\overline{f}_{j+2} = af_j + b(f_{j+1} + f_{j-1}) + c(f_{j+2} + f_{j-2}) + d(f_{j+3} + f_{j-3}), \tag{16}$$

where the definition of the coefficients α, β, a, b, c, and d can be found in Zhao.[7]

D. Boundary conditions

No-slip and adiabatic boundary conditions were specified on the surface of the impenetrable walls, i.e., the flow velocity on the wall surface is specified as $\partial r_w(x,t)/\partial t$, the wall surface velocity. Nonreflecting characteristic boundary conditions[17,18] were imposed at the inflow and outflow boundaries to minimize acoustic reflections while still allowing the passage of vortical structures into and out of the domain. In reality, voice production involves the excitation of many acoustic modes within the vocal tract, i.e., formants. The effects of such acoustic loading was ignored here in order to highlight the production of sound as opposed to sound propagation phenomena. Our focus in the present study is on sound production, i.e., flow/sound interactions within the vocal tract.

E. Acoustic source characterization

In the classical acoustic analogy,[19] the nonlinear problem of flow–sound interactions is assumed to be analogous to sound generation and propagation in a uniform acoustic medium at rest. The effects of turbulence, flow–structure interactions, or other nonlinear phenomena on the acoustic field are relegated to source terms in a nonhomogeneous linear wave equation. In the present study, a form of the acoustic analogy that accounts for the effects of moving surfaces, namely the Ffowcs Williams–Hawkings equation,[12] was applied. The solution of the Ffowcs Williams–Hawkings equation for the domain depicted in Fig. 2 is well known. It is given, for example, by Zhao *et al.*[11] and takes the form

$$p' = \frac{1}{2c_\infty A_0}\frac{\partial}{\partial t}\int\int\int_V [T_{11}]_{t*}\,dV(\mathbf{y}) \quad \text{(I)} - \frac{1}{2A_0}\int\int_{S_w} [p'\delta_{1j} - \sigma_{1j}]_{t*} n_j\,dS(\mathbf{y}) \quad \text{(II)} - \frac{1}{2A_0}\int\int_{S_w} [\rho_\infty c_\infty v_j]_{t*} n_j\,dS(\mathbf{y}) \quad \text{(III)} + \frac{1}{2A_0}\int\int_{S_{\text{in}}} [p' + \rho_\infty c_\infty u_1]_{t*}\,dS(\mathbf{y}), \quad \text{(IV)} \tag{17}$$

TABLE I. Details of the four cases.

Case	Δp	Rotational motion
1	0	No
2	9.8 cm H_2O	Yes
3	9.8 cm H_2O	No
4	9.8 cm H_2O	No

where S_w is a function describing the location of the wall surface, V is the volume containing the acoustic sources, A_0 is the cross-sectional area of the uniform rigid part of the duct, S_{in} is the inlet surface, n_j is the vector normal to the surface, c_∞ is the speed of sound, T_{11} is the axial kinetic energy component of Lighthill's stress tensor, and $t^*=t-(x_1-y_1)/c_\infty$ is the retarded time. The quantities in brackets are to be evaluated at the retarded time. The physical meaning of the four terms is as follows: Term (I) is a quadrupole source due to net unsteady flow inside the vocal tract; term (II) is a dipole source due to the net unsteady force in the x direction exerted by the surface of the vocal folds onto the fluid; term (III) is a monopole source due to the motion of the vocal fold in the throat region—if the wall is fixed this term is zero; and term (IV) is the monopole source enforced at the inlet. In the present study, term (IV) makes no acoustic contribution since a nonreflecting boundary condition was specified at the inlet. A discussion of the precise meaning of each term in the equation can also be found in Goldstein[20] and Hofmans.[4]

In simple terms, the dipole contribution is governed by the instantaneous pressure drop across the orifice. The pressure drop is, in turn, mostly a function of the ratio of the orifice area and the tube area. The orifice geometry (cross section and profile) may also play a role. The monopole contribution is due to displacement of fluid by the moving walls. Finally, the quadrupole component is the result of fluctuations in axial kinetic energy of the flow downstream of the orifice. It is the only contribution to the sound field that does not involve an interaction between the flow motion and the surface of the vocal fold walls.

III. MODEL DETAILS

Four different cases were studied to isolate the effects of the different possible sound generation phenomena involved. A summary of the input conditions for all cases is shown in Table I.

A. Case 1: No mean flow

In Case 1, no mean pressure gradient was imposed between the regions upstream and downstream of the glottis. The included angle of the glottal opening was kept zero, and the diameter of the glottis varied with time. This case was used to highlight the basic contributions of the different source terms in Eq. (17), for a motion of the vocal folds that is relatively simple. The time-dependent shape of the glottal orifice for this case was specified as

$$r_w(x,t)=\frac{D_0-D_g(t)}{4}\tanh(s)+\frac{D_0+D_g(t)}{4}, \tag{18}$$

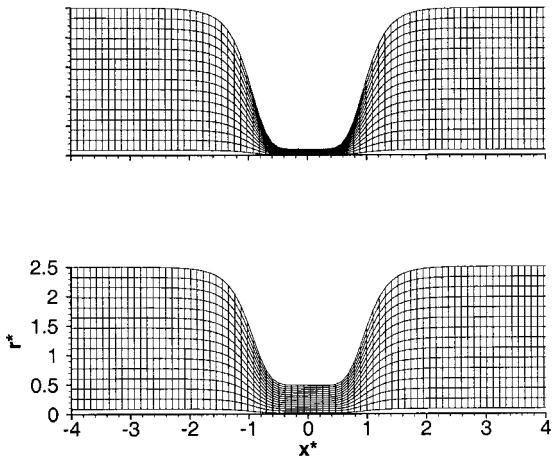


FIG. 3. Geometry and grid in the source region for Case 1 at time $t=T$ and $t=1.5T$.

where

$$s=\frac{b|x|}{D_{\max}}-\frac{bD_{\max}}{|x|}. \tag{19}$$

The maximum diameter of the glottis was $D_{\max}=4$ mm, $D_0=5D_{\max}$ was the duct diameter, $b=1.4$ is a constant, and $D_g(t)$ is the time-dependent glottal diameter, which is given by

$$D_g(t)=\frac{D_{\max}+D_{\min}}{2}+\frac{D_{\max}-D_{\min}}{2}\cos(2\pi f_0 t). \tag{20}$$

The minimum diameter of the glottis was $D_{\min}=0.8$ mm and f_0 was the forcing frequency. Sketches of the wall geometry and the grid at two instants of time are shown in Fig. 3. The variables $x^*=x/D_{\max}$, and $r^*=r/D_{\max}$ are nondimensional coordinates. The computational domain extended from $x^*=-40$ to $x^*=40$. The grid size was 800×80. The flow inside the duct was initially at rest with uniform flow variables.

B. Cases 2–4: With mean flow

Three cases with a uniform mean flow resulting from the imposition of a mean pressure gradient across the glottal orifice were also investigated. For Case 2, the motion of the vocal folds was designed to mimic the oscillatory nature of the human vocal cords described in the Introduction. The radius of the vocal tract was defined as

$$r_w(x,t)=\frac{D_0-D_{\min}}{4}\tanh(s)+\frac{D_0+D_{\min}}{4}+\frac{1}{2}[1-\tanh(s)]\times[\beta_1(x+cD_{\max})-\beta_2(x-cD_{\max})]. \tag{21}$$

Here the constant c was equal to 0.42 and other variables were defined as in the first case. The first two terms in Eq. (21) define the basic shape of the vocal tract for a minimal glottal orifice. The last term was a correction that defines the

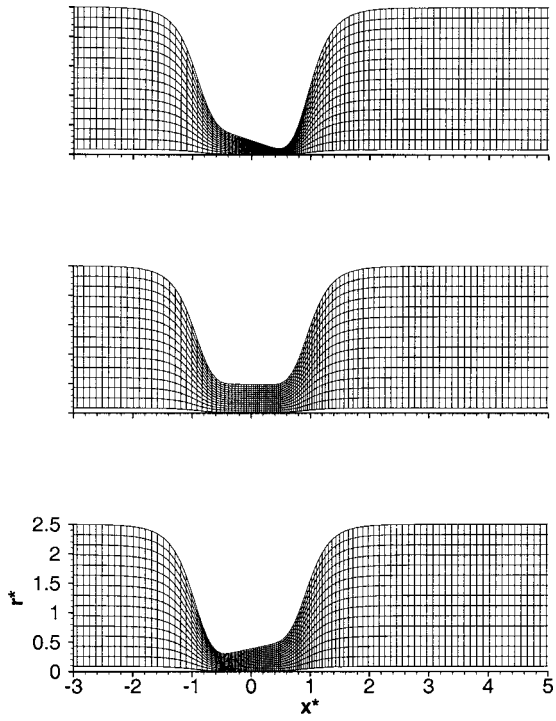


FIG. 4. Geometry and grid in the source region for Case 2 at times $t=1.2$ T, $t=1.5$ T and $t=1.7$ T.

instantaneous shape of the glottal orifice itself. The variables β_1 and β_2 define the included angle of the glottis, α:

$$\alpha=\tan^{-1}(\beta_2)-\tan^{-1}(\beta_1), \tag{22}$$

where β_1 and β_2 are functions of time during one cycle of the glottis motion given by

$$\beta_1(t) = \begin{cases} 0, & \text{for } t\leqslant \frac{1}{9}T, \\ 0.244\left\{1-\cos\left[\frac{9\pi}{4}\left(\frac{t}{T}-\frac{1}{9}\right)\right]\right\}, & \text{for } \frac{1}{9}T<t\leqslant\frac{5}{9}T, \\ 0.488, & \text{for } \frac{5}{9}T<t\leqslant\frac{6}{9}T, \\ 0.244\left\{1+\cos\left[3\pi\left(\frac{t}{T}-\frac{6}{9}\right)\right]\right\}, & \text{for } \frac{6}{9}T<t\leqslant T. \end{cases} \tag{23}$$

The time period was $T=1/f_0$ and

$$\beta_2(t)=\beta_1(t+T/9). \tag{24}$$

The model geometry and typical grids at several times for Case 2 are shown in Fig. 4.

Initially (at $t=0$), air in the vocal tract was at rest, with uniform density, and the following pressure distribution was enforced:

$$p(x,r)=p_\infty+\frac{\Delta p}{2}\left[1-\tanh\left(\frac{10x}{D_{\max}}\right)\right], \tag{25}$$

where $\Delta p=912$ Pa is the initial difference between the upstream and downstream pressures. The maximum velocity through the glottis, estimated using Δp and Bernoulli's equation, $U_M=\sqrt{2\,\Delta p/\rho_\infty}$, was about 40 m/s. The actual maximum velocity inside the vocal tract may be slightly smaller than U_M when the orifice is nearly closed because of viscous losses through the glottis. The frequency of the vocal fold motion was $f_0=125$ Hz. The computational domain extended from $8D_{\max}$ upstream of the glottis and $86D_{\max}$ downstream of the glottis. The grid size used in these simulations was 800×120. A grid convergence study for a similar case with identical geometry but with rigid walls was performed, as discussed in Zhao *et al.*[11] The Mach number corresponding to U_M was artificially increased to 0.2 to improve numerical efficiency.

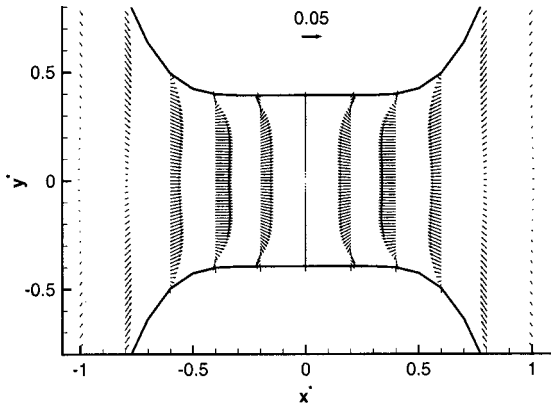


FIG. 5. Velocity vectors near the glottal region at $t=1.25T$ for Case 1.

The integration domain (sound source region V) for the Ffowcs Williams–Hawkings equation was delimited by the region between the inlet and $x^*=30$ for Cases 1 and $x^*=70$ for Case 2, respectively. The results were not sensitive to these boundaries as long as the entire source region was included. Any vortical flow was assumed to contribute to sound generation at the onset, and thus all regions where significant vorticity appeared in the solution were considered for the calculation of the farfield sound using the acoustic analogy.

Two other cases, Case 3 and Case 4, were also investigated. These additional cases were identical to Case 2, except that the rotational component of the wall motion was suppressed. The time-dependent glottal diameter $D_g(t)$ was given by Eq. (20). In Case 3, the values for $D_{\max}$ and $D_{\min}$ were the same as those in Case 1. For Case 4, $D_{\max}=4.464$ mm and $D_{\min}=0.8$ mm.

IV. SOUND GENERATION MECHANISMS

A. The role of displacement flow

Case 1 was investigated first. The glottal geometry featured parallel walls with no mean pressure gradient imposed. The computed velocity field near the glottal region is shown in Fig. 5. Fluid flow is induced purely by the vocal fold motion in this case since there was no imposed mean pressure gradient. Note that the flow is symmetric about the center of the glottis. The velocity is not zero on the surface of the vocal tract due to the vocal fold motion. The magnitude

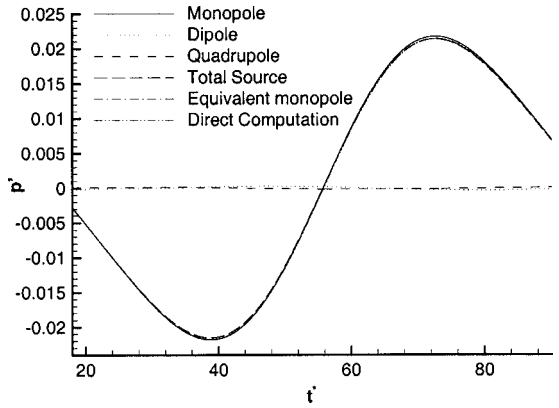


FIG. 6. Predictions of the radiated sound obtained from the acoustic analogy and direct computations for Case 1. The contributions of the different terms in Eq. (17) are also shown.

of the velocity everywhere is very small. The directly computed acoustic pressure, p', at $x^*=40$ is shown in Fig. 6. Also shown in this figure are the predictions from the Ffowcs Williams–Hawkings equation using the computed near-field flow information to calculate the radiated pressure. The directly simulated acoustic pressure and that predicted from the acoustic analogy using either the total source term or only the monopole source, term (III), are in excellent agreement.

The nondimensional acoustic wavelength corresponding to the forcing frequency f_0 is 400, which is much larger than the spatial extent of the source region in this case. Thus, all the acoustic sources can be treated as compact. Since the model geometry is symmetric about the center of the glottis and no pressure gradient is enforced, the net force exerted on the fluid by the moving surfaces of the vocal fold is zero. This results in a zero dipole source contribution. Note that the fluctuations u_1^2, p', and ρ', as well as the viscous stress, are very small. The quadrupole source radiation is thus negligible as well. The motion of the walls displaces a certain amount of fluid in and out of a fixed control region surrounding the glottal orifice. This leads to the injection of mass into the tube at a rate directly proportional to the rate of change of the volume occupied by the fluid within the orifice with time. It would be straightforward in Case 1 to calculate the radiated sound pressure utilizing an ideal equivalent monopole of strength $Q=dV/dt$. Figure 6 shows that the radiated sound pressure computed using the equivalent monopole agrees very well with the directly simulated acoustic signal. Hence, the monopole source contribution due to the wall motion is the dominant sound source for this case, in contrast with results from Case 2, described in the next section.

B. The role of the pressure drop across the orifice

The imposition of a mean flow causes a pulsating jet through the glottal orifice. The necessity of abruptly starting the flow at the initial time step resulted in transients in the numerical solution. It was presumed that the flow near the glottis region reached a periodic state after one cycle of the vocal fold motion. The initial transients in the solution were thus ignored. A time sequence of vorticity contours in the idealized vocal tract during the second cycle of the glottis

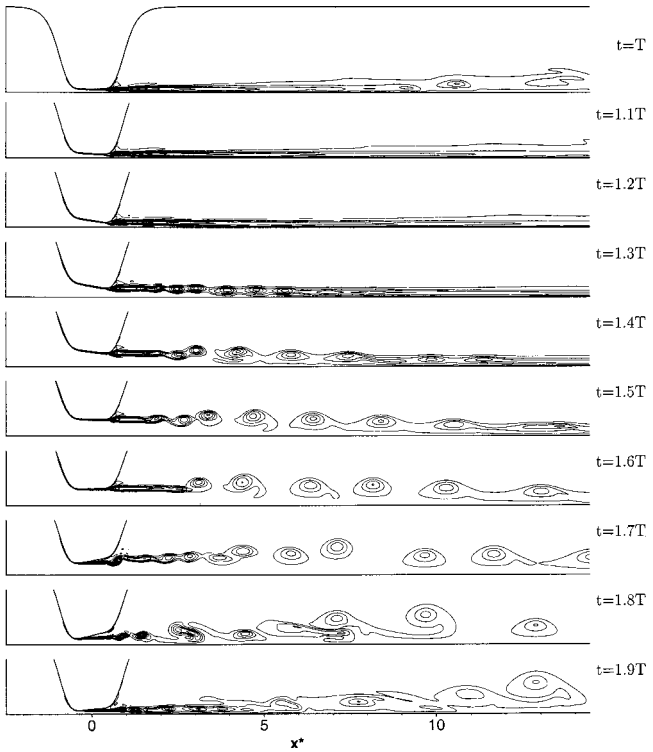
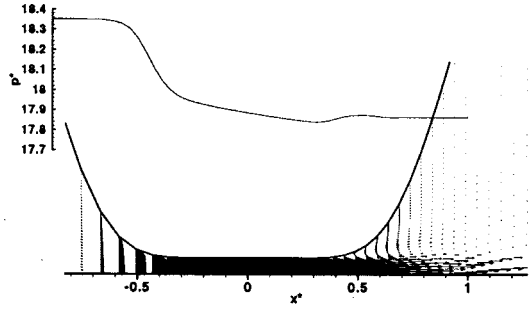


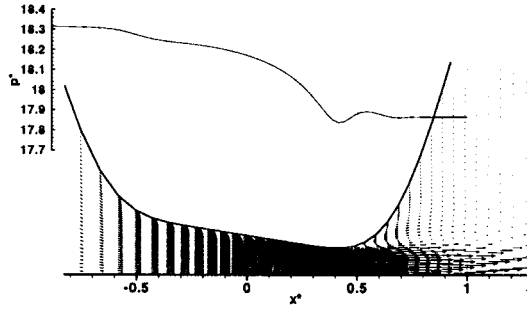
FIG. 7. A time sequence showing contour plots of vorticity during one cycle of the glottis motion for Case 2.

motion, presumed to be representative of steady-state periodic conditions, is shown in Fig. 7 for Case 2. At the beginning of the cycle, the glottal diameter is small and the jet through the glottis is stable. The vortices observed far downstream of the glottis at this time were generated during the previous cycle. As the walls move apart from each other and the glottal diameter becomes larger, the orifice shape gradually changes until it forms a converging nozzle. The flow structure inside the glottis is rather simple, since the mean pressure gradient everywhere within the converging walls is favorable. As the glottal diameter is increased, the jet Reynolds number is increased and consequently the jet is more unstable. This results in secondary vortices generated several glottal diameters downstream of the glottis. After the glottal diameter reaches its maximum value, the glottal walls rotate and begin to adopt a diverging nozzle shape. When the included angle of the glottis exceeds a certain threshold value, the flow separates at the entry of the glottis. Such flow separation in a diverging diffuser is sometimes referred to as a ''diffuser stall.'' Strong vortices are then generated near the outlet due to the interaction between the jet and the glottis walls. This unsteady vortex shedding results in high-frequency unsteady forces on the surface of the vocal folds. Also, the vortices themselves and the interactions between them (vortex pairing, etc.) result in high-frequency Reynolds stress fluctuations. Note that the frequency characteristic of vortex shedding is much greater than the forced oscillation frequency.

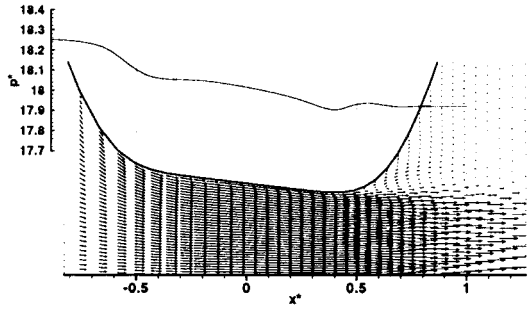
A temporal sequence of velocity vector plots near the glottal region is shown together with the pressure distribution along the glottis surface for Case 2 in Fig. 8. The wall pressure along the glottis is important in speech production



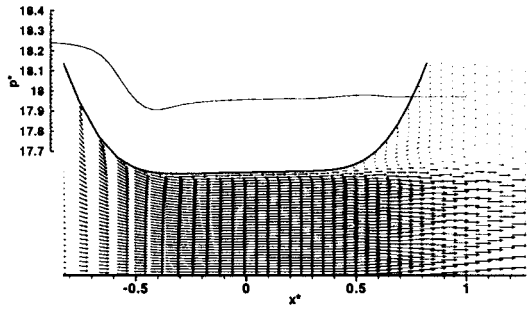
t=T



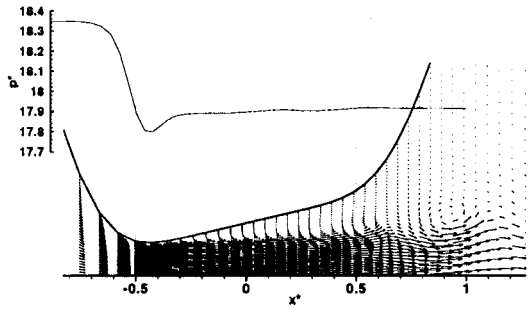
t=1.2T



t=1.4T



t=1.6T



t=1.8T

FIG. 8. A time sequence of velocity vector plots near the glottal region together with pressure profiles along the surface of the glottis for Case 2. For the velocity vector plots, only half the grid points in the axial direction and one-third those in the radial direction are shown for clarity.

because it contributes to the force driving the oscillations of the vocal cords. It also determines in part the dipole strength. At time $t=T$, the glottis takes a uniform shape and the diameter is at its minimum value. Most of the pressure drop occurs upstream of the glottis due to the flow acceleration. Inside the glottis the pressure decreases more gradually due to friction along the surface. The pressure recovers slightly at the end of the glottis due to the expansion and remains almost constant after the flow separates. The total pressure drop through the glottis is greatest at this time. At time $t=1.2T$ and $t=1.4T$ the glottis takes on a converging shape. The pressure decreases as the flow accelerates inside the

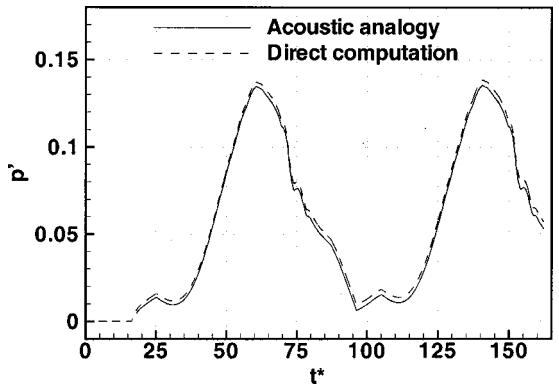


FIG. 9. A comparison of the directly computed acoustic signal and the prediction from the acoustic analogy for Case 2.

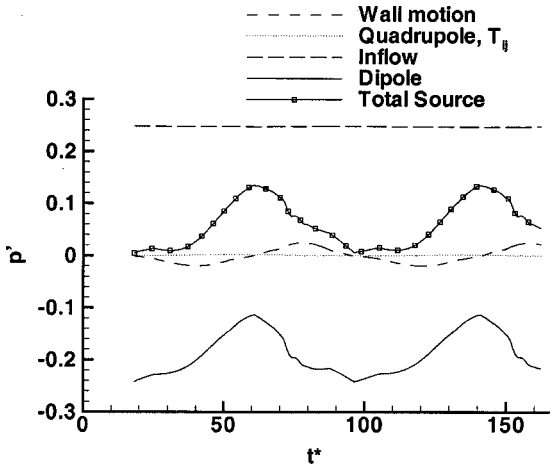


FIG. 10. Contributions to sound radiation from different sound source terms for Case 2.

glottis. A slight recovery of the pressure is observed at the trailing edge of the glottis. The increase of the minimum diameter of the glottis results in a decreased pressure drop through the glottis. At $t=1.6T$, the glottis resembles a slightly diverging nozzle. The flow does not separate inside the glottis. The pressure increases slowly along the glottis due to the flow deceleration. At $t=1.8T$, the included angle of the glottis becomes larger. The pressure decreases sharply as the glottis diameter approaches its minimum value. Then it recovers due to the expansion and remains almost constant after the flow separates inside the glottis. The complex flow field inside the glottis, which features flow separation and reattachment, triggers a Kelvin–Helmholtz-type instability of the jet and secondary vortices are formed near the outlet of the glottis. The pressure drop through the glottis increases as the diameter of the glottis decreases.

The directly calculated acoustic pressure signal at a downstream location of $x^*=82.6$, near the wall, is shown in Fig. 9. Also shown in the figure are predictions obtained using the Ffowcs Williams–Hawkings equation. The agreement between the two results is excellent. The reason for the slight difference has been discussed in Zhao *et al.*[11] As the acoustic wave propagates downstream, a thin boundary layer forms near the the vocal tract wall. The boundary layer reduces the effective cross-sectional area of the tract, and, thus, the characteristic mechanical impedance of the one-dimensional waves in the tube. Therefore, the use of the original area, A_0, in the acoustic analogy results in underestimated acoustic pressure predictions. Other possible sources of differences might be sound speed variations in the farfield (where the observer is located), as well as refraction and dissipation effects on acoustic waves by the mean flow field, which are neglected in the solution of the Ffowcs Williams–Hawkings equation.[21,22]

The validity of the Ffowcs Williams–Hawkings equation to predict acoustic pressure makes it possible to investigate the importance of the different sound sources. The contributions to the radiated sound from each of the four source terms in Eq. (17) are shown in Fig. 10. The monopole source at the inlet does not make any acoustic contribution because of the nonreflecting boundary conditions specified there. It does, however, yield a constant pressure because the mean pressure upstream of the glottis is higher than the downstream pressure, p_∞. The dipole source due to the unsteady force exerted on the surface of the glottis dominates all other contributions in magnitude. As shown in Fig. 8, the pressure drop through the glottis varies with time as the shape and the diameter of the glottis change. The resulting unsteady net force dominates the radiated sound. The strength of the dipole is the net force in the streamwise direction imparted by the moving surfaces onto the surrounding fluid. Because the stationary duct walls are normal to the radiation direction they do not contribute. The monopole source contribution due to the displacement flow here is of the same order of magnitude as for Case 1, although the waveform is different due to the different volume velocity time history.

The magnitude of the quadrupole source contribution in this case is about two orders smaller than that of the dipole source. In order to better appreciate its characteristics, the quadrupole source contribution is replotted on a smaller scale in Fig. 11. The low-frequency component is in phase with the dipole source contribution. The time period for the high-frequency component is approximately 1/20 of the forcing period, T_0. The sound source for low-frequency radiation can be treated as compact, as discussed for Case 1. Neglecting the entropy variation and the viscous stresses, the

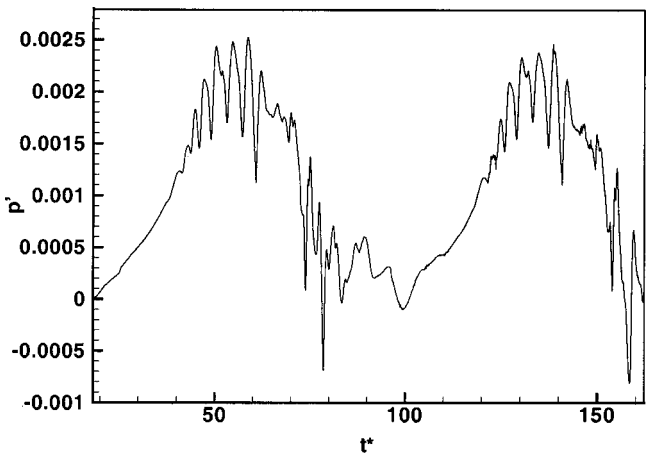


FIG. 11. Sound radiation from the quadrupole source for Case 2.

low-frequency contribution from the quadrupole source can be expressed as

$$(I)_{\text{low}} = \frac{1}{2c_\infty A_0} \frac{\partial}{\partial t} \left[\int \int \int_V \rho u_1^2 dV(\mathbf{y}) \right]_{t^*}, \tag{26}$$

which is associated with the first-order derivative of the total axial kinetic energy of the flow inside the source volume. The temporally changing rate of the kinetic energy is related to the net force exerted by the vocal folds. This explains why the low-frequency contribution of the quadrupole source is in phase with that of the dipole source. The high-frequency contribution from the quadrupole source is produced by strong vortex interactions near the glottis. During the early stage of the glottis opening, the walls assume a converging shape and the diameter is then small. The jet downstream of the glottis is stable and the vortices generated are very weak. The high-frequency oscillations of the pressure signal are small. When the diameter of the glottis is near its maximum value and when the glottis assumes a diverging shape, the high-frequency sound becomes more important due to the strong vortices observed downstream of the glottis.

C. Equivalent monopole or dipole strength

It is sometimes convenient to replace the actual sources of sound within the glottis by one equivalent monopole radiating in a uniform, rigid duct. The concept of an equivalent monopole refers to a source analogous to a snuggly fitted moving piston with the same area as the duct, injecting an acoustic volume velocity or flow rate into the duct equal to the acoustic volume velocity measured downstream. This ideal sound source combines all the source mechanisms discussed so far and itemized in Eq. (17). It may be obtained by inverse filtering the radiated sound. Such an equivalent source is convenient to predict the sound field analytically, when the details of the flow are not important. The dual of an equivalent volume velocity source, a dipole, can also be used. In the ideal case of an infinitely long duct, monopoles and dipoles are interchangeable, and their strengths are related.

The time-dependent volume flux through the glottis has been used to approximate the acoustic source strength in source–filter models of voice production.[2] The validity of this approximation was verified. Figure 12 shows the time variation of volume flux through the cross section at $x^* = -0.47$ and $x^* = 0.49$ for Case 2. These two locations correspond to the leading edge and the trailing edge of the glottis, respectively. Slight differences are observed between them because of the compressibility of the air flow and the motion of walls. The acoustic signals predicted using the volume velocity through the glottis are shown in Fig. 13, and are compared with the directly computed acoustic pressure. The wave form of the local volume flux source prediction is quite different from that of the direct computation. The volume flux through the glottis is governed by the total contribution from all of the sources upstream of the glottis. When it is used to predict the acoustic radiation, the sound sources downstream of the selected control surface are missing. These include the unsteady forces on the downstream side of the vocal folds as a dipole source and the unsteady flow downstream of the glottis as a quadrupole source. Similar observations were also reported by McGowan,[3] Hirschberg, and co-workers.[23,24]

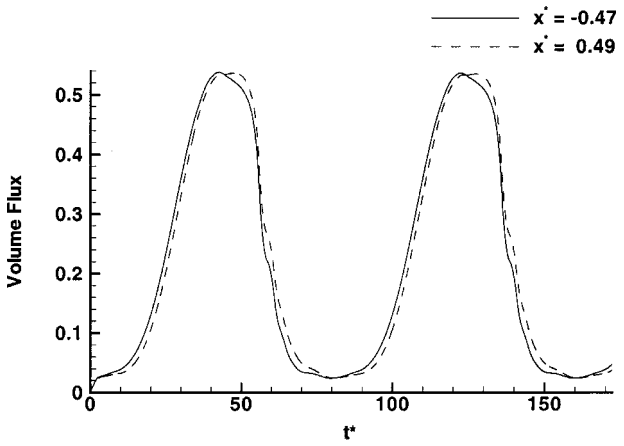


FIG. 12. Volume flux through the glottis at two selected cross-sections for Case 2.

It is therefore important to select a location downstream of the source region, outside the near field, to calculate the equivalent monopole source strength. Predictions for a control surface located after the duct walls become uniform (normal to the flow direction) are accurate. The flow rate within the orifice itself, however, does not include information related to the wall pressures downstream of the orifice, which contribute to the dipole and quadrupole components.

D. Effects of frequency

It is interesting to emphasize that while the dipole strength is independent of frequency (it varies with the orifice coefficient and the pressure gradient), the magnitude of the monopole term is directly proportional to frequency. This is shown in Fig. 14, where the amplitude of the monopole contribution was computed for three different simulations at different frequencies for the geometry and flow conditions of Case 2. The monopole source begins to be significant with

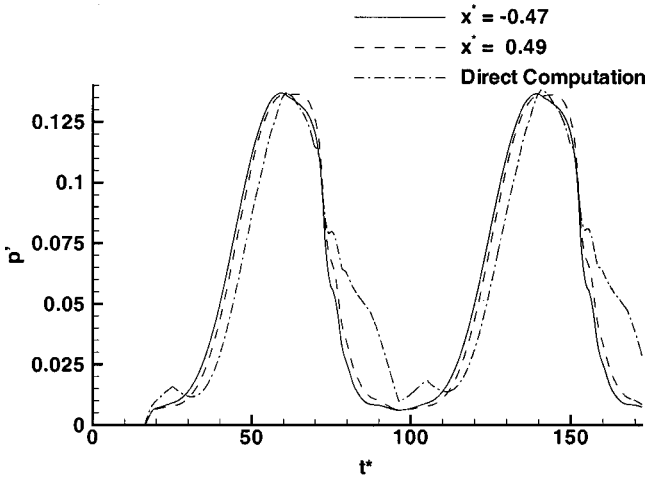


FIG. 13. Sound pressure predicted using an ideal monopole at $x^* = 0$, with the volume flux through the glottis as the sound source strength compared with the direct computation.

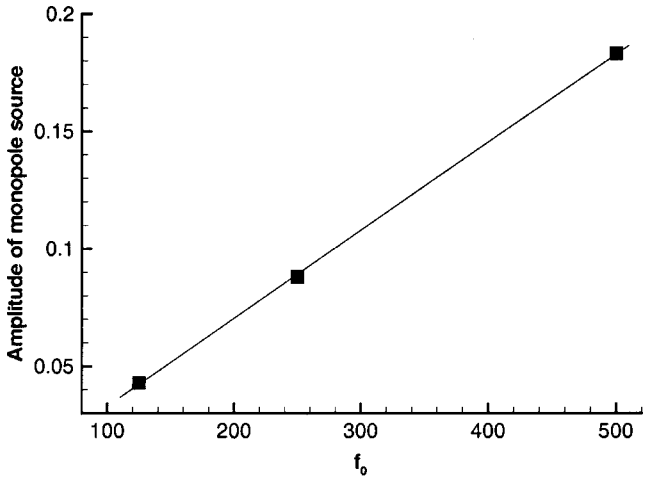


FIG. 14. Effect of frequency on the amplitude of the monopole source.

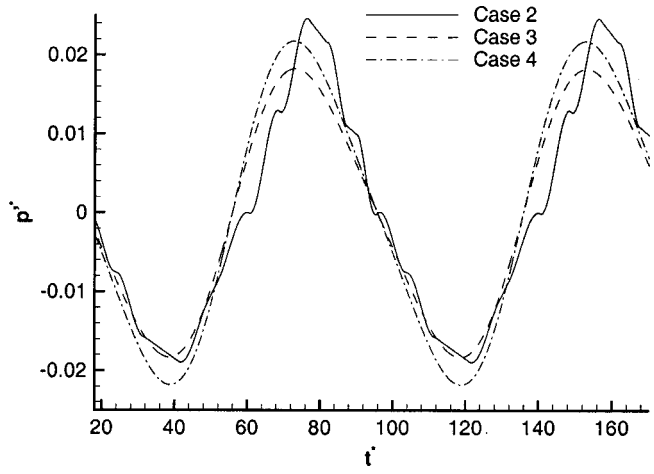


FIG. 15. Sound radiation from the monopole source for Case 2, Case 3, and Case 4.

respect to the dipole source when the frequency exceeds 400 Hz in the particular case considered here. Although it is risky to generalize this observation to actual laryngeal flows, the possible implications of this observation for voice and speech are interesting. The relative contributions of monopole and dipole components are influenced by the geometry of the orifice, and wall motion. In general, the monopole strength relative to the dipole strength should increase with frequency. Therefore, it may be that voice production is predominantly dipole-like at low frequencies (up to the third or fourth formant frequency), and monopole-like at higher frequencies. The dipole mechanism is inherently quasisteady, while the monopole mechanism is not (the monopole amplitude is directly proportional to frequency). The dipole mechanism is nonlinear, while the monopole mechanism is linear (when the reaction of the walls to the fluid excitation is not taken into account). Straightforward adjustments in the source parameters of source–filter multimass models could readily be made to take these distinctions into account. Perhaps the addition of the monopole based on displacement flow evaluation (which requires only the calculation of the instantaneous volume occupied by the folds) could result in improved voice synthesis. This is proposed as a topic of future investigations.

E. Effects of wall rotation

As discussed earlier, the time-varying geometry of the orifice results in different wall pressure distributions. The aerodynamic forces and the elastic forces of the tissue of the vocal folds together constitute the driving force of the self-sustained motion of human vocal folds. However, the question of whether the converging/diverging geometry of the glottis affects voice production has been of some interest to the speech production community.[25] This question was investigated by comparing the radiated sound for Case 2 with that for two additional cases, Cases 3 and 4, whereby the wall motion was modified to eliminate the rotational component. The orifice walls moved with the same amplitude and phase along the center of the orifice, such that the orifice walls were always parallel to the x axis. The amplitude of the parallel wall motion was first established in order for the maximum and minimum areas to match those for Case 2 at the instants where the walls were straight ($t=T$ and $t=1.5T$) (Case 3). Then, the same case was simulated again (Case 4) but this time adjusting the minimum and maximum areas to match the monopole strength of Case 2 (rather than the area). The monopole and dipole contributions for these two cases are shown with the same quantities for Case 2 in Fig. 15 and Fig. 16, respectively. The results show that the monopole amplitude is generally much smaller than the dipole amplitude in all cases (at least at this frequency, which was 125 Hz). Furthermore, the dipole strength for all three cases with flow is comparable.

It appears that the converging/diverging variation of the orifice geometry does not influence the amplitude of the radiated sound significantly at low frequencies. The rotation of the walls, which may induce additional angular momentum to the flow, producing vorticity, causes substantial fluctuations at high frequencies, but they are of small amplitude relative to the fundamental component. It could well be that the area ratio between the duct and the orifice is so large that the orifice coefficient depends more on the area itself than it does on the actual shape of the glottis. The impact of the converging/diverging component of the wall motion cer-

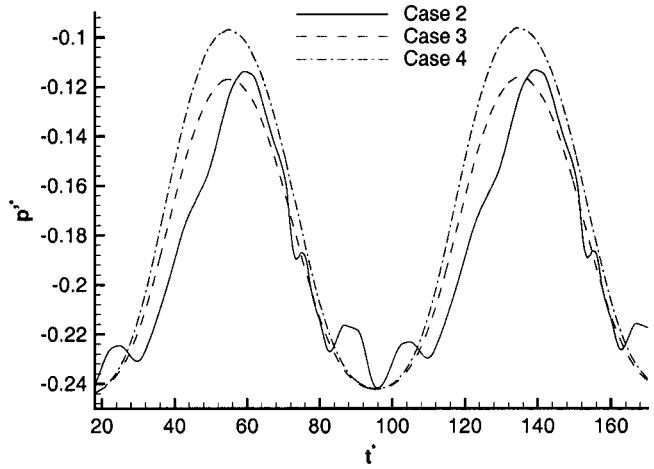


FIG. 16. Sound radiation from the dipole source for Case 2, Case 3, and Case 4.

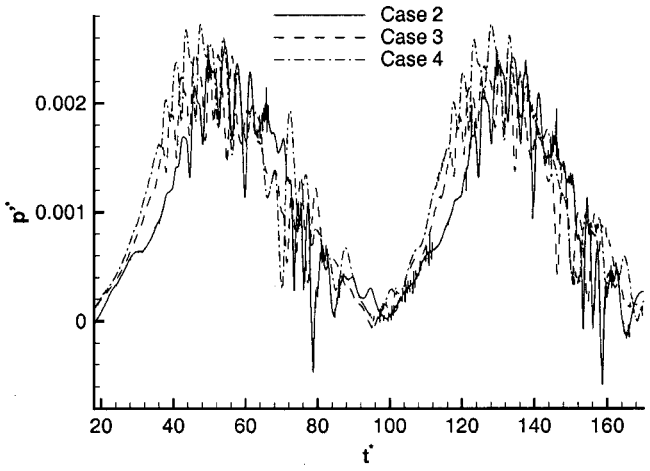


FIG. 17. Sound radiation from the quadrupole source for Case 2, Case 3, and Case 4.

tainly affects the structure and the strength of the shed vorticity. The high-frequency contributions of the quadrupole, shown in Fig. 17 are significantly different for the three cases. But this is again a rather small effect, considering that two-dimensional flow simulations tend to exaggerate the contributions of discrete vortices in the flow field. The presence of turbulence, which occurs in reality, would most certainly dissipate the vortices shed from the orifice edges much faster, reducing the importance of the quadrupole terms. Based on the results of the simulations performed so far, as shown in Fig. 18, it seems that the rotational component of the wall motion does not inherently enhance the direct production of sound from the flow.

V. CONCLUSIONS

Numerical simulations of the flow and acoustic field in an idealized axisymmetric model of the human vocal tract were performed. The compressible Navier–Stokes equations in a moving coordinate system were integrated using a high-order numerical method. An acoustic analogy was employed to identify specific sound generation mechanisms within the flow. The simulated flows featured complex vortical patterns with separation and reattachment occurring during different stages of the vocal fold motion. The main sound source, at a frequency of 125 Hz, was identified as a dipole source due to the unsteady force exerted on the surface of the glottis.

The results of this numerical investigation suggest that while the dipole component of the source strength may dominate at low frequencies, the monopole component related to displacement flow can be significant at higher frequencies above approximately 400 Hz. Furthermore, it appeared that the rotational mode of the fold vibrations, causing the glottis to adopt in succession converging and diverging shapes, does not enhance the voice production process. The area ratio regulated the dipole strength.

The vocal tract model described in this paper, including the geometry and the motion of vocal folds, represents a highly idealized concept of human voice production, due to constraints related to the affordability and simplicity of the simulation. The relative importance of the contributions from different sources may not necessarily be representative of human phonation. The axisymmetric assumption prevents the generation of turbulence, which is expected to dramatically change the flow pattern inside and downstream of the glottis, as well as the subsequent acoustic radiation. The high-frequency components from the quadrupole sources are expected to be at least one order of magnitude higher according to Zhao *et al.*,[26] where sound radiation from both a laminar and a turbulent confined jet with the same flow conditions were simulated using the large-eddy simulation technique. The magnitude of the contribution from the monopole source due to wall motion in the current study is about one-fifth that of the dominant dipole source. It is believed to be negligible in speech production.[2] Nevertheless, the model possesses some of the basic features of human vocal tract. The vocal fold motion style designed is even more realistic than the mechanical models used in recent experiments,[8–10] where parallel vocal folds motion was applied.

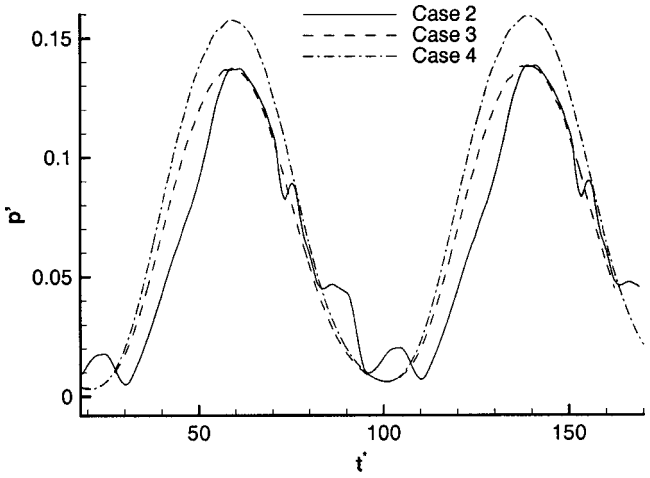


FIG. 18. A comparison of farfield sound for Case 2, Case 3, and Case 4.

ACKNOWLEDGMENTS

This research was supported by Grant No. RO1 DC03577-02 from the National Institutes of Deafness and Other Communication Disorders, National Institute of Health.

[1] C. K. W. Tam, "Computational aeroacoustics: Issues and methods," AIAA J. **33**, 1788–1796 (1995).
[2] J. L. Flanagan, *Speech Analysis, Synthesis and Perception* (Springer-Verlag, New York, 1972).
[3] R. S. McGowan, "An aeroacoustic approach to phonation," J. Acoust. Soc. Am. **83**, 696–704 (1988).
[4] G. C. J. Hofmans, "Vortex sound in confined flows," Ph.D. thesis, Eindhoven University of Technology, Eindhoven, Netherlands, 1998.
[5] H. Iijima, N. Miki, and N. Nagai, "Glottal impedance based on a finite element analysis of two-dimensional unsteady viscous flow in a static glottis," IEEE Trans. Signal Process. **40**, 2125–2135 (1992).
[6] F. Alipour, C. Fan, and R. C. Scherer, "A numerical simulation of laryngeal flow in a forced oscillation glottal model," Comput. Speech Lang. **10**, 75–93 (1996).
[7] W. Zhao, "A numerical investigation of sound radiation from subsonic jets with application to human phonation," Ph.D. thesis, Purdue University, West Lafayette, IN, 2000.
[8] L. Mongeau, N. Franchek, C. H. Coker, and R. A. Kubli, "Characteristics of a pulsating jet through a small modulated orifice, with application to

voice production,'' J. Acoust. Soc. Am. **102**, 1121–1133 (1997).
[9] C. H. Shadle, A. Barney, and P. O. A. L. Davies, ''Fluid flow in a dynamic mechanical model of the vocal folds and tract. II. Implications for speech production studies,'' J. Acoust. Soc. Am. **105**, 456–466 (1999).
[10] A. Barney, C. H. Shadle, and P. O. A. L. Davies, ''Fluid flow in a dynamic mechanical model of the vocal folds and tract. I. Measurements and theory,'' J. Acoust. Soc. Am. **105**, 444–455 (1999).
[11] W. Zhao, S. H. Frankel, and L. Mongeau, ''Numerical simulation of sounds from confined pulsating axisymmetric jets,'' AIAA J. **39**, 1868–1874 (2001).
[12] J. E. Ffowcs Williams and D. L. Hawkings, ''Sound generation by turbulence and surfaces in arbitrary motion,'' Philos. Trans. R. Soc. London, Ser. A **264**, 321–342 (1969).
[13] C. Zhang, W. Zhao, S. H. Frankel, and L. Mongeau, ''Computational aeroacoustics of phonation, Part II: effects of flow parameters and ventricular folds,'' J. Acoust. Soc. Am. **112**, 2147–2154 (2002).
[14] J. C. Tannehill, D. A. Anderson, and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer* (Taylor & Francis, Washington, DC, 1997).
[15] I. R. Titze, *Principles of Voice Production* (Prentice–Hall, Englewood Cliffs, NJ, 1994).
[16] S. K. Lele, ''Compact finite difference schemes with spectral-like resolution,'' J. Comput. Phys. **103**, 16–42 (1992).
[17] K. W. Thompson, ''Time dependent boundary conditions for hyperbolic systems,'' J. Comput. Phys. **68**, 1–24 (1987).
[18] T. J. Poinsot and S. K. Lele, ''Boundary conditions for direct simulations of compressible viscous flows,'' J. Comput. Phys. **101**, 104–129 (1992).
[19] M. J. Lighthill, ''On sound generated aerodynamically, I. General theory,'' Proc. R. Soc. London, Ser. A **211**, 564–587 (1952).
[20] M. E. Goldstein, *Aeroacoustics* (McGraw-Hill, New York, 1976).
[21] B. E. Mitchell, S. K. Lele, and P. Moin, ''Direct computation of the sound generated by vortex pairing in an axisymmetric jet,'' J. Fluid Mech. **33**, 113–142 (1999).
[22] W. Zhao, S. H. Frankel, and L. Mongeau, ''Effect of spatial filtering on sound radiation from a subsonic axisymmetric jet,'' AIAA J. **38**, 2032–2039 (2000).
[23] A. Hirschberg, ''Some fluid dynamic aspects of speech,'' Bull. Commun. Parlée. **2**, 7–30 (1992).
[24] A. Hirschberg, X. Pelorson, G. C. J. Hofmans, R. R. van Hassel, and A. P. J. Wijnands, ''Starting transient of the flow through an in-vitro model of the vocal folds,'' in Vocal Fold Physiology, edited by P. Davis and N. Fletcher (Singular, San Diego, 1996), pp. 31–46.
[25] R. C. Scherer, D. Shinwari, K. J. De Witt, C. Zhang, B. R. Kucinschi, and A. A. Afjeh, ''Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees,'' J. Acoust. Soc. Am. **109**, 1616–1630 (2001).
[26] W. Zhao, S. H. Frankel, and L. Mongeau, ''Large eddy simulation of sound radiation from subsonic turbulent jets,'' AIAA J. **39**, 1469–1477 (2001).

Computational aeroacoustics of phonation, Part II: Effects of flow parameters and ventricular folds

Cheng Zhang,[a] Wei Zhao,[b] Steven H. Frankel,[c] and Luc Mongeau
School of Mechanical Engineering, Purdue University, West Lafayette, Indiana 47907

(Received 31 October 2001; revised 3 July 2002; accepted 19 July 2002)

The results are described of the second part of an ongoing study aimed at performing direct numerical simulations of translaryngeal flows during phonation. The use of accurate numerical schemes allows the radiated sound to be calculated directly, without the need for acoustic analogy models. The goal is to develop a better understanding of this class of flow, and of the basic sound generation mechanisms involved in phonation. In the present study, the effects of subglottal pressure and of glottal oscillation frequency on the near-field flow and farfield sound were investigated. The effects of the presence of the ventricular folds downstream of the oscillating glottal region were also examined. The results highlighted the effects of subglottal pressure and oscillation frequency on the jet vortical structure, wall pressure and shear stress, and sound radiation. Jet impingement on the ventricular folds introduced additional dipole sources similar to those observed in problems involving grazing flows over cavities. © *2002 Acoustical Society of America.*
[DOI: 10.1121/1.1506694]

PACS numbers: 43.70.Bk [AL]

I. INTRODUCTION

In this paper we describe the results from the second part of an ongoing study aimed at performing direct numerical simulations of translaryngeal flows during phonation. The use of accurate numerical schemes allows the radiated sound to be calculated directly, without the need for acoustic analogy models. The goal is to develop a better understanding of this class of flows, and of the basic sound generation mechanisms involved in phonation.

In the first part of this investigation,[1] it was found that the shape of the glottis has a significant effect on the stability and vortical structure of the pulsating jet produced in the vocal tract. Interactions between the pulsating jet and the vocal tract when the glottal shape assumed the form of a conical diffuser resulted in a complex separated flow pattern and introduced additional acoustic sources at the high end of the frequency range. However, the low-frequency radiated sound pressure was found to depend mainly on the pressure drop across the orifice, in turn, primarily regulated by the orifice area. At higher frequencies, the relative importance of ''displacement flow,'' i.e., the mass of fluid displaced by the motion of the glottis walls as they vibrate, was found to be increased.

In the present study, a few more configurations were investigated in order to gain insight on the effects of oscillation frequency and transglottal pressure. The same methods as those employed in the previous investigation were used. Attempts were then made to elucidate the fundamental role of the ventricular folds on sound generation. The geometry of the vocal tract model was modified to add this feature. The interaction of the glottal jet with the ventricular folds downstream created additional dipole sources of sound at frequencies corresponding roughly to the resonance frequency of the cavity formed between the vocal folds and the ventricular folds.

II. SIMULATION DETAILS

Four cases were studied as part of this parametric study to investigate the effects of glottal oscillation frequency (Case 5 versus Case 6), subglottal pressure (Case 6 versus Case 7), and of the presence of the ventricular folds (Case 6 versus Case 8). For Cases 5, 6, and 7, the time-dependent shape of the vocal tract model was also described by Eq. (21) in Part I of this study,[1] with different expressions for functions β_1 and β_2:

$$\beta_1(t)=\begin{cases} 0, & \text{for } t\leqslant \frac{1}{9}T, \\ 0.488\phi\left[2.25\left(\frac{t}{T}-\frac{1}{9}\right)\right], & \text{for } \frac{1}{9}T<t\leqslant\frac{5}{9}T, \\ 0.488, & \text{for } \frac{5}{9}T<t\leqslant\frac{6}{9}T, \\ 0.488\phi\left[1-3\left(\frac{t}{T}-\frac{6}{9}\right)\right], & \text{for } \frac{6}{9}T<t\leqslant T. \end{cases} \tag{1}$$

The time period is $T=1/f_0$ (f_0 is the vocal fold oscillation frequency) and

$$\beta_2(t)=\beta_1(t+T/9). \tag{2}$$

[a]Graduate student.
[b]Currently employed at Reaction Engineering International, Salt Lake City, Utah.
[c]Author to whom correspondence should be addressed. Electronic mail: frankel@ecn.purdue.edu

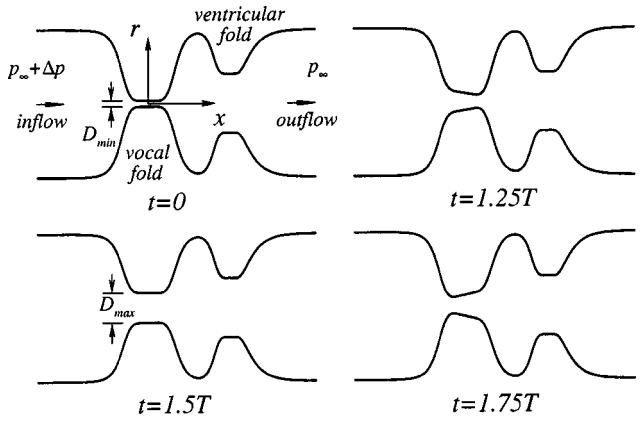


FIG. 1. Sketch of the vocal tract model geometry.

The function $\phi(t)$ is defined as

$$\phi(t)=\begin{cases}\frac{2}{3}\left[t-\frac{1}{5\pi}\sin(5\pi t)\right], & \text{for } 0\leqslant t\leqslant 0.2,\\ \frac{4}{3}(t-0.1), & \text{for } 0.2<t\leqslant 0.7,\\ \frac{2}{3}\left[t+0.5+\frac{3}{10\pi}\sin\left(\frac{10\pi}{3}(t-0.7)\right)\right], & \text{for } 0.7<t\leqslant 1.\end{cases} \tag{3}$$

The ventricular folds were taken into account in Case 8 by modifying the shape of the vocal tract model as follows:

$$r_{w,\mathrm{vf}}(x,t)=r_w(x,t)+\frac{D_0-D_{\mathrm{vf}}}{4}[\tanh(s_{\mathrm{vf}})-1], \tag{4}$$

where

$$s_{\mathrm{vf}}=\frac{b_{\mathrm{vf}}|x-x_{\mathrm{vf}}|}{D_{\max}}-\frac{b_{\mathrm{vf}}D_{\max}}{|x-x_{\mathrm{vf}}|}, \tag{5}$$

and $D_{\mathrm{vf}}=2D_{\max}$ is the diameter of the vocal tract at the location of the ventricular folds. $x_{\mathrm{vf}}=3D_{\max}$ is the x coordinate of the ventricular folds. The constant b_{vf} is equal to 2.0 for $x\leqslant x_{\mathrm{vf}}$ and 1.0 for $x>x_{\mathrm{vf}}$. A sketch of the vocal tract geometry for Case 8 is shown in Fig. 1.

For Cases 5, 6, and 7, the grid size was 800×120. For Case 8, the grid size was 1000×120 and the grid near the vocal fold and ventricular fold regions was refined. A summary of the input conditions for all four cases is shown in Table I.

For the results and discussions presented below, all the parameters and variables are in nondimensional form (indicated by an asterisk), unless otherwise mentioned. The characteristic scales used to nondimensionalize the dependent and independent variables were the diameter of the glottis, $D_{\max}$, the estimated peak velocity U_M, and the initial quiescent values of the density and temperature for Case 5.

TABLE I. Details of the four cases.

Case	Δp	f_0	Ventricular folds
5	9.8 cm H_2O	125 Hz	No
6	9.8 cm H_2O	250 Hz	No
7	19.6 cm H_2O	250 Hz	No
8	9.8 cm H_2O	250 Hz	Yes

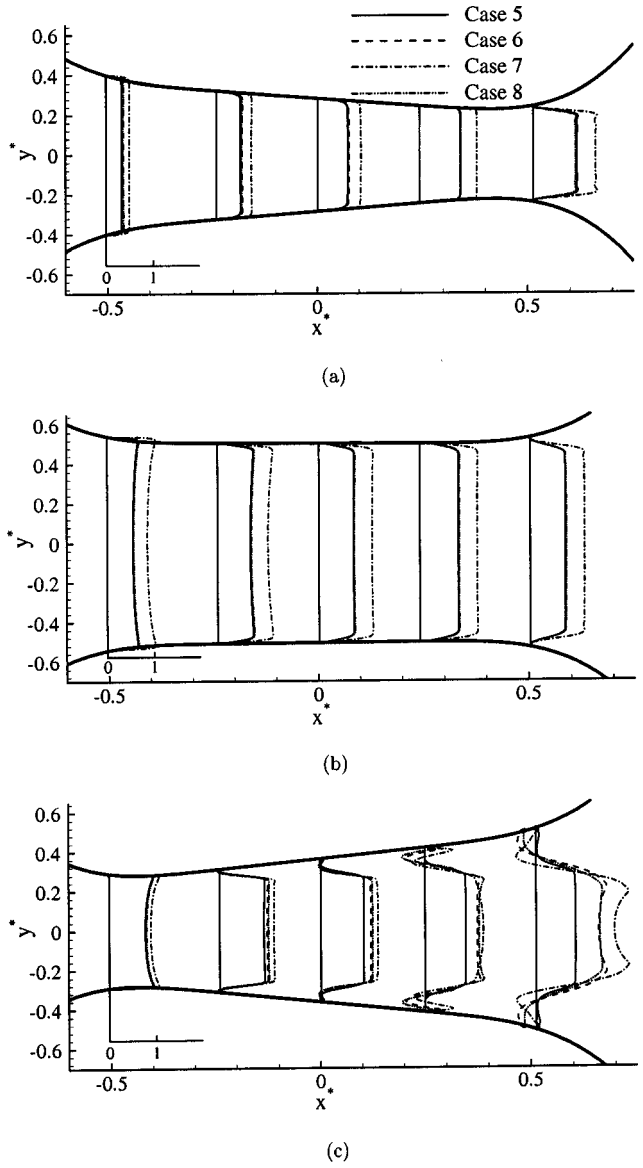


FIG. 2. Velocity profiles for (a) $t=1.25T$, (b) $t=1.5T$, (c) $t=1.75T$.

III. RESULTS

A. Intraglottal velocity profile

Instantaneous nondimensional velocity profiles at several axial locations within the glottis are shown for all four cases at times $t=1.25T$, $t=1.5T$, and $t=1.75T$, where T is the flow period, in Figs. 2(a)–(c), respectively. In the early and middle parts of the cycle, when the glottis assumes a converging or a uniform shape, respectively, the velocity profiles are fairly uniform and no evidence of flow separation (flow reversal near the walls) is observed in the glottis. In the latter part of the cycle, when the glottis assumes a diverging shape, the velocity profiles in all four cases show varying degrees of flow separation. This phenomenon is often termed

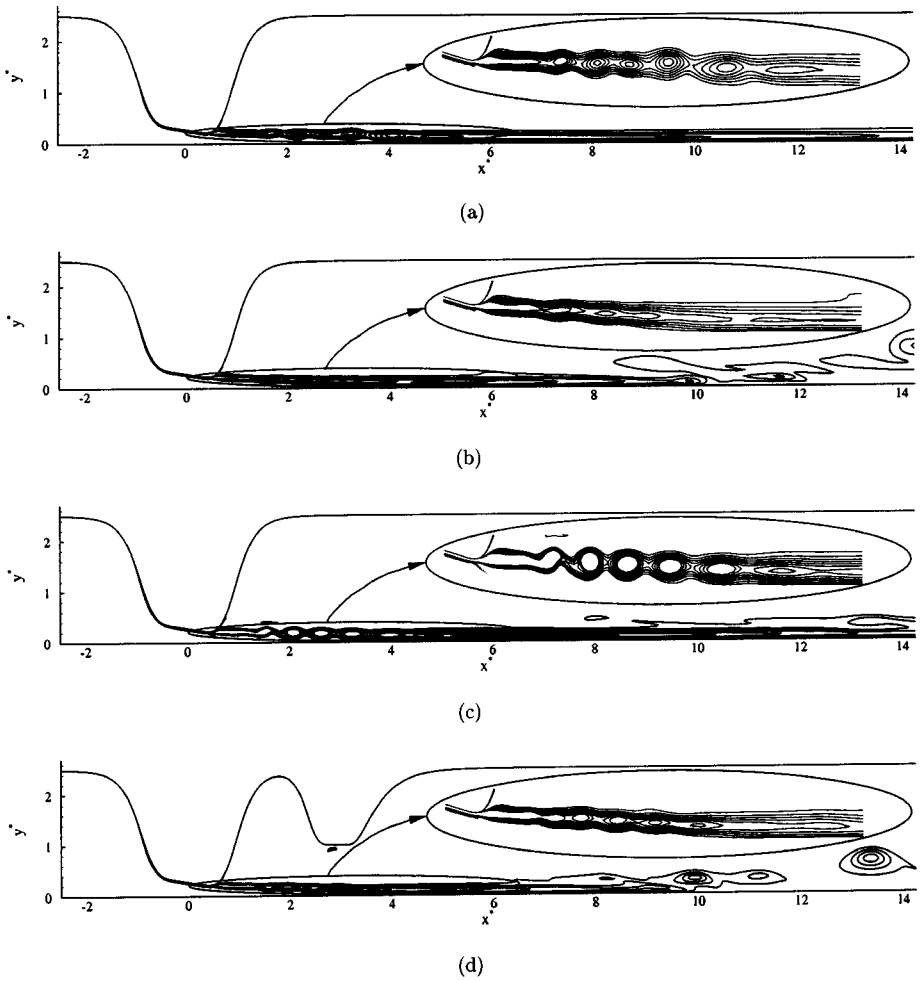


FIG. 3. Vorticity contours at $t = 1.25T$ for all the cases (a) Case 5, (b) Case 6, (c) Case 7, (d) Case 8. The nondimensional strength of the vorticity is between 0.8 and 12.

a ''diffuser stall.'' In many instances of actual flows through diverging orifices, the flow may attach preferentially on one wall, causing the jet centerline to be skewed with respect to the axial direction. This is termed a ''Coanda'' effect. In the present numerical simulations, the assumption of an axisymmetric flow precludes the reproduction of the Coanda effect, which is an intrinsically three-dimensional phenomenon.

Doubling the glottal oscillation frequency, as revealed by a comparison between Case 5 and Case 6, resulted in a stronger recirculation zone. Increasing the subglottal pressure, Case 6 and Case 7, increased the velocities during the entire cycle. When no flow separation occurs, the velocity gradients near the wall were apparently larger due to the increase of the subglottal pressure. The presence of the ventricular folds did not appear to have a significant effect on the velocity profiles within the glottis itself.

B. Instantaneous vortical flow structure

To visualize the flow fields, contour plots of instantaneous vorticity are shown for the four cases at three different times during one cycle in Figs. 3–5. The insert in Fig. 3 shows a magnified view of the jet near the glottal exit. The vorticity shed downstream of the orifice is related to the pressure drop across the orifice. Vortex formation involves the conversion of fluid linear momentum into angular mo-

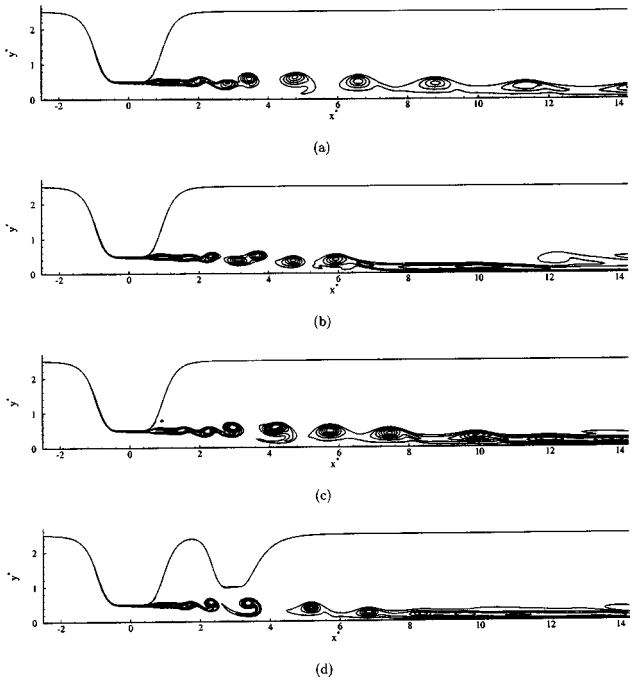


FIG. 4. Vorticity contours at $t = 1.5T$ for all the cases (a) Case 5, (b) Case 6, (c) Case 7, (d) Case 8. The nondimensional strength of the vorticity is between 0.8 and 12.

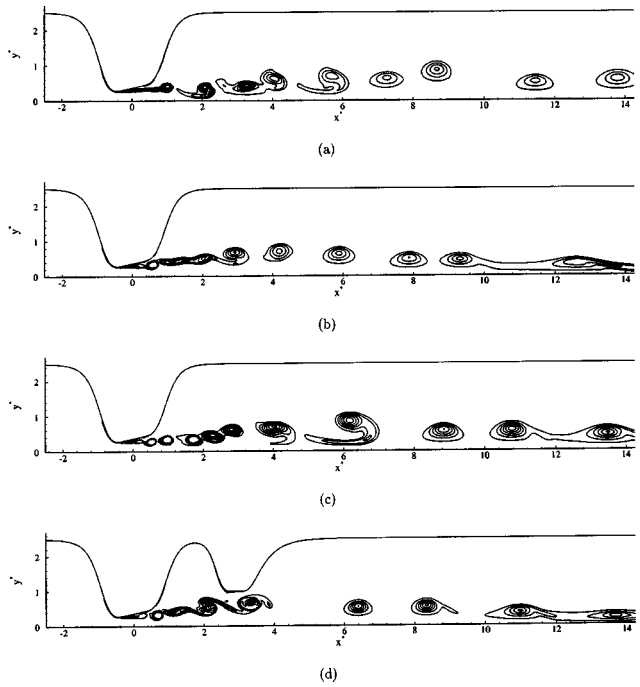


FIG. 5. Vorticity contours at $t=1.75T$ for all the cases (a) Case 5, (b) Case 6, (c) Case 7, (d) Case 8. The nondimensional strength of the vorticity is between 0.8 and 12.

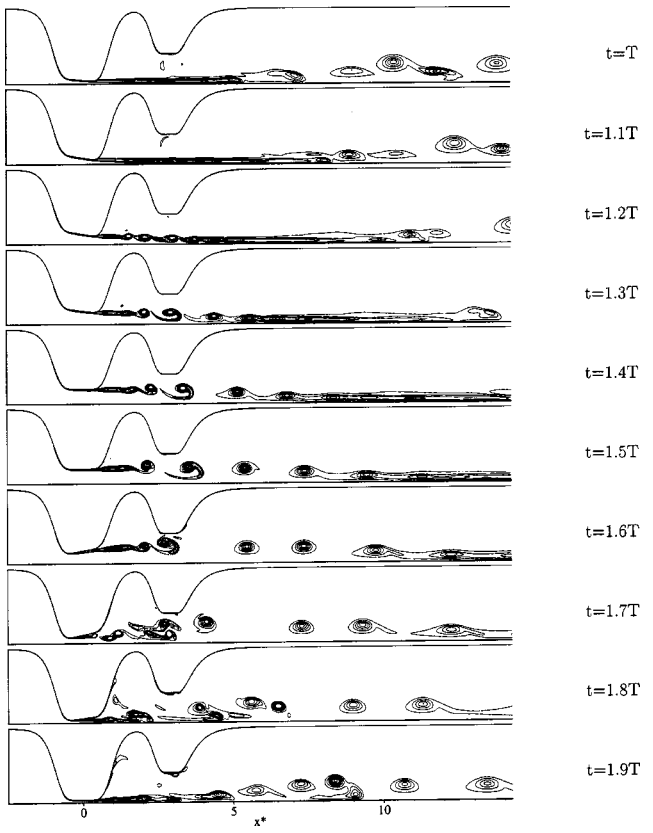


FIG. 6. A time sequence showing vorticity contours during one cycle of the glottis motion for Case 8. The nondimensional strength of the vorticity is between 0.8 and 12.

mentum. Once formed, vortices tend to persist for a long time as they are convected downstream, at a velocity approximately one-half the centerline velocity. Once shed, vortices tend to attract each other because of the low pressure at their core. Toroidal vortices such as those created in the simulations may ''leapfrog,'' or coalesce in complex ways. Such flow structures may be observed, of course, in real flows, but usually the onset of turbulence as well as anisotropy and asymmetry in the boundary conditions for real flows tend to smudge or disperse coherent vortical structures far downstream at a much faster rate than in numerical simulations.

Increasing the glottal frequency was observed to shorten the distance between neighboring shed vortices, increasing vortex interactions. Increasing the subglottal pressure increases the vorticity. A strong interaction occurs when the vortical structures of the pulsating jet impinge on the ventricular folds when the glottis takes on a diverging shape. The vocal folds tend to direct the jet toward the ventricular folds. In order to study how the flow develops during an entire cycle, a time sequence of vorticity contour plots is shown in Fig. 6 for the case with the ventricular folds (Case 8). The evolution of the pulsating jet shows evidence of Kelvin–Helmholtz shear layer instabilities, featuring unsteady vortex roll-up and pairing processes. During the latter part of the cycle a complex interaction between the vortical jet and the ventricular folds can be observed. A complex unsteady flow pattern develops within the cavity, e.g., laryngeal ventricle, between the vocal and ventricular folds, as can be seen in the streamline plots of Fig. 7. The implications of this on the acoustic radiation to the farfield is discussed in Sec. III D.

It is important to note that in all cases the vortices survive and convect downstream of the glottis. This is an artifact of the axisymmetric (two-dimensional) assumption, which precludes vortex stretching and the generation of small-scale turbulence. It would be necessary to conduct three-dimensional simulations to study this effect. Because of the fine grid resolution that would be required and the long simulation time needed for a pulsating flow of this nature, direct numerical simulations (no turbulence model) would not be possible. The more economical large eddy simulations, where only large-scale turbulent structures are resolved and the effects of small-scale (subgrid) turbulence is modeled, would most likely still be prohibitive.[2]

C. Wall pressure and shear stress

The wall pressure and viscous shear stress are of interest as they contribute to the net force on the vocal folds and play a role in establishing their flow-induced oscillations in actual speech. Profiles of the wall pressure and shear stress are shown in Figs. 8 and 9 at the three times under consideration during the cycle for all four cases. To facilitate comparisons, the pressure is nondimensionalized as $p^*=(p-p_\infty)/\Delta p$. In general, the wall pressure decreases in all cases as the flow accelerates when it enters the glottis due to the Bernoulli effect. At $t=1.5T$ and especially at $t=1.75T$, when the vocal folds take on a diverging shape, the wall pressure decreases more rapidly than during the early part of

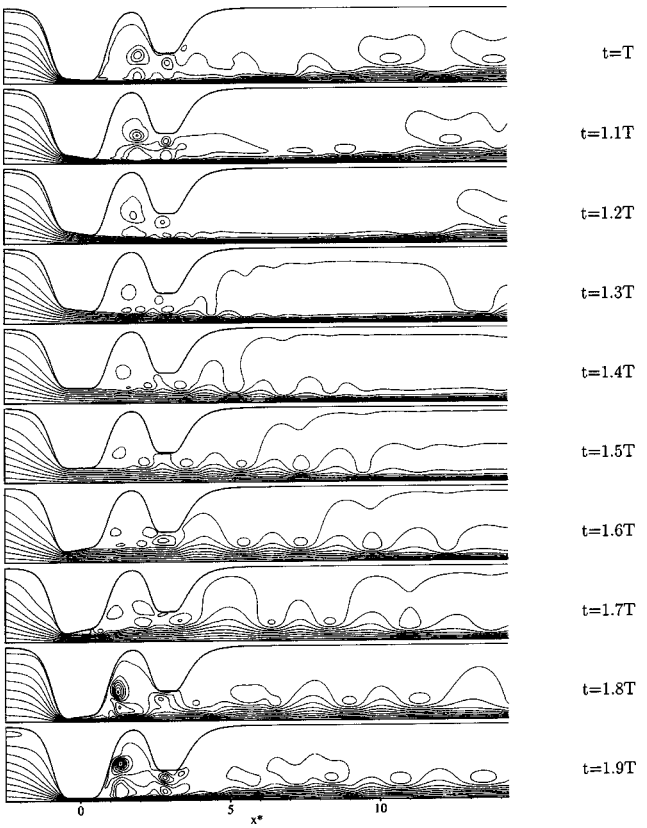


FIG. 7. A time sequence showing streamlines during one cycle of the glottis motion for Case 8.

the cycle andreaches a minimum near the position, where the cross-sectional area becomes smallest. The slight pressure recovery at the glottal exit is mainly caused by the increase of the cross-sectional area. Wall pressure profiles can also be found in Iijima *et al.*,[3] Guo and Scherer,[4] and Scherer *et al.*[5]

The wall shear stress increases when it enters the glottis due to flow acceleration caused by the decrease of cross-sectional area, then it drops due to flow deceleration caused by viscous loss and/or the increase of cross-sectional area. For $t=1.25T$ and $t=1.75T$, the wall shear stress reaches a maximum just upstream of the position where the cross-sectional area is smallest. At the glottal exit and a few other positions, negative wall shear stress can be found, which indicates a recirculation zone. The wall shear stress is highest for the case with the higher subglottal pressure and hence Reynolds number. This is due to larger velocity gradients near the wall for this case. For $t=1.5T$ and $t=1.75T$, flow impingement on the ventricular folds for Case 8 has significant effects on the pressure distribution and wall shear stress. Flow separation due to the presence of the ventricular folds can also be seen at these two times. Any force produced by a wall onto the fluid in the axial direction (either from a normal pressure stress or a shear stress) is expected to radiate sound through a dipole mechanism. The ventricular folds are thus expected to contribute to sound generation to a certain extent, in addition to affect the dipole strength of the glottal orifice itself.

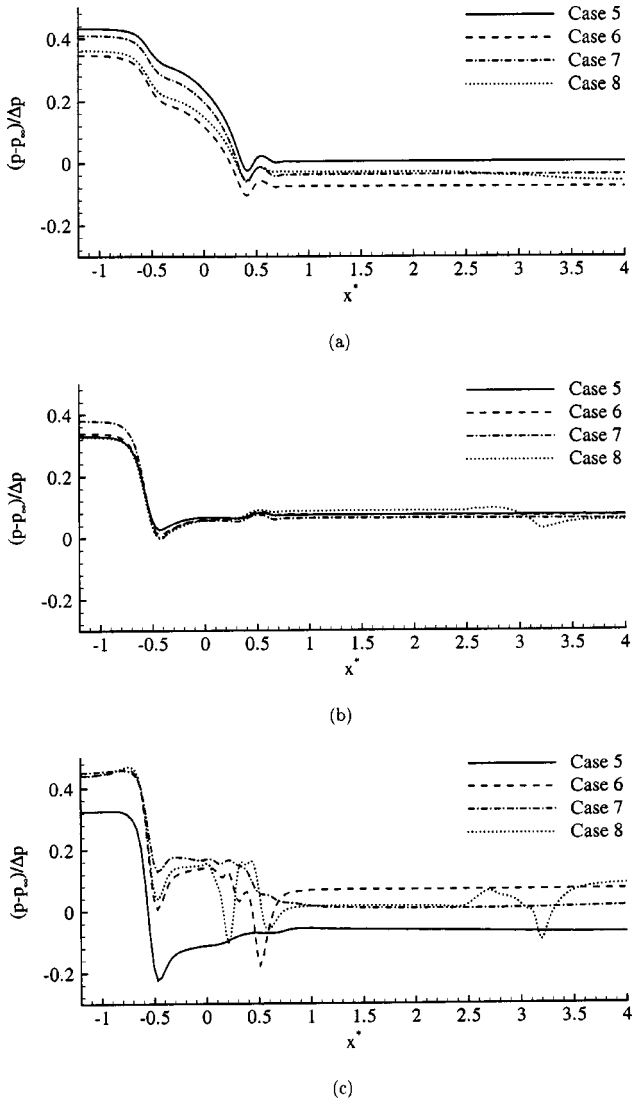


FIG. 8. Wall pressure distribution for (a) $t=1.25T$, (b) $t=1.5T$, (c) $t=1.75T$.

D. Farfield acoustic pressure

The farfield acoustic pressure is plotted versus time for each of the four cases in Fig. 10. To facilitate comparisons the time axis was redefined as $(t-t_0)/T$, where t_0 is the time required for the generated sound to travel from the source position to the farfield. The similar time signatures for all four different cases suggests that the sound radiation is directly related to the shape of the glottis. Doubling the glottal oscillation frequency shows little effect on the farfield sound, while the increase in the subglottal pressure increases the sound level in a predictable fashion. The radiated acoustic pressure is nearly directly proportional to the intraglottal pressure drop, all other things being the same. For the case with the ventricular folds, when the vocal folds take on a diverging shape, additional high-frequency components in the sound pressure data are apparent. Subsequent analysis will demonstrate that these high-frequency components represent dipole source contributions.

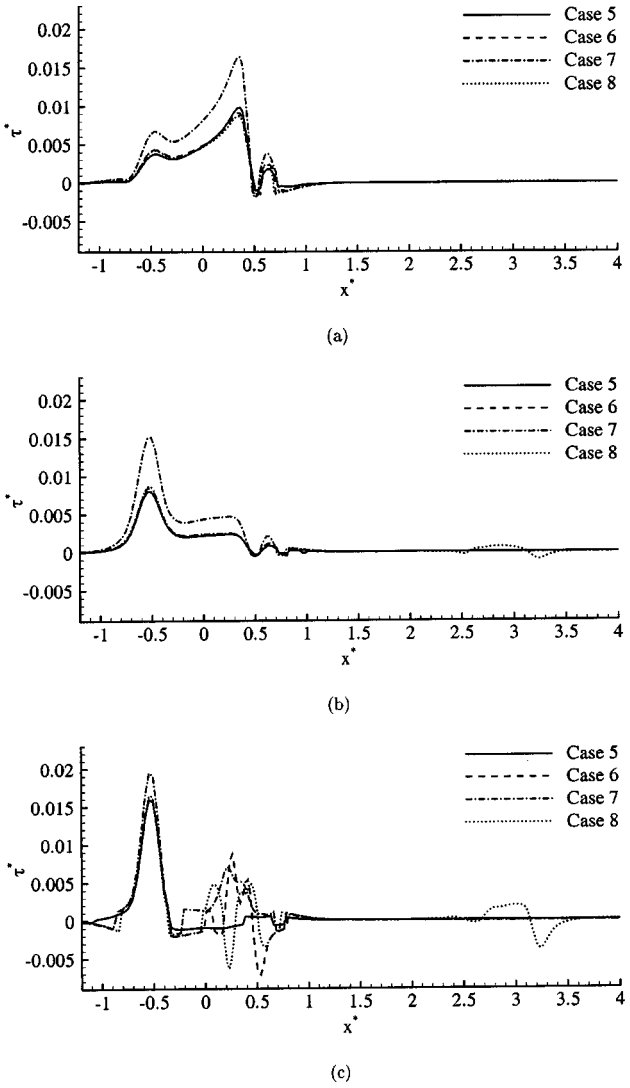


FIG. 9. Wall shear stress for (a) $t=1.25T$, (b) $t=1.5T$, (c) $t=1.75T$.

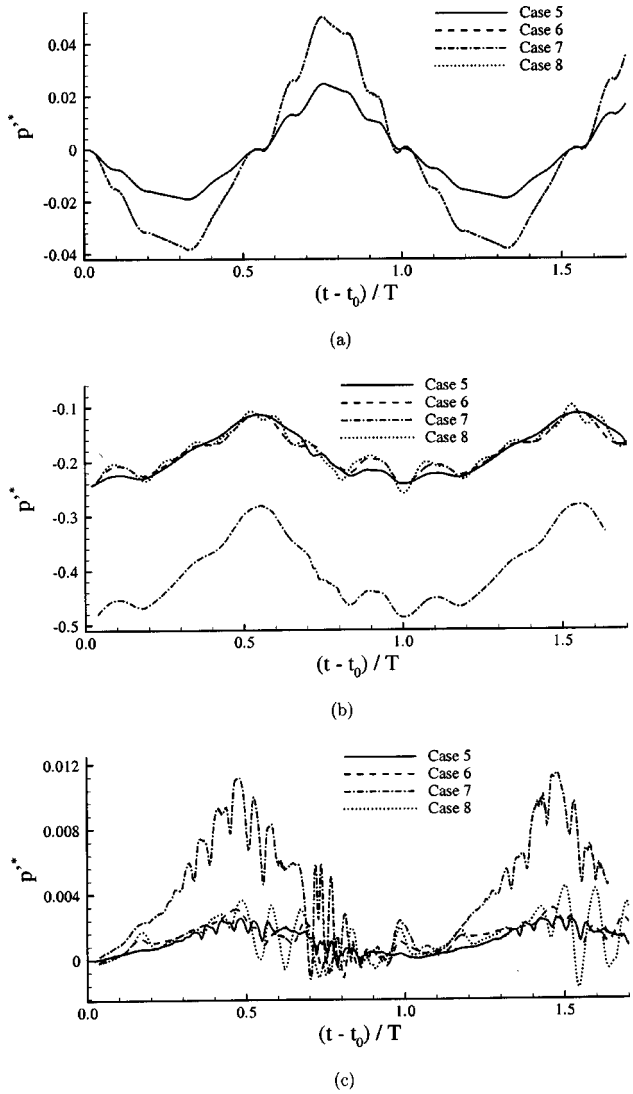


FIG. 11. Contribution to sound radiation from (a) monopole source, (b) dipole source, (c) quadrupole source.

E. Acoustic analogy validation and source contributions

By using the acoustic analogy based on the Ffowcs Williams–Hawkings equation, as described in Part I of this study,[1] the contributions from different sound generation mechanisms were calculated and compared for all four cases. It was found that farfield sound predictions from the acoustic analogy were in excellent agreement with direct computations for all cases. This confirmed the validity of the acoustic analogy.[1] In Figs. 11(a)–(c), the contribution to the total sound radiation from monopole, dipole and quadrupole sources are shown. From Fig. 11(a) it appears that changing the subglottal pressure does not affect the monopole source, but doubling the glottal oscillation frequency nearly doubles the monopole source amplitude. This trend is expected since a doubling of the frequency results in a doubling of the wall velocity when the wall displacement is imposed. Including the ventricular folds did not affect the monopole sources. This can be explained by the fact that the ventricular folds were assumed to be rigid and the monopole amplitude is solely determined by the rate of change of the fluid volume caused by the oscillation of the vocal folds. Hence, it is only related to the vocal fold oscillation frequency.

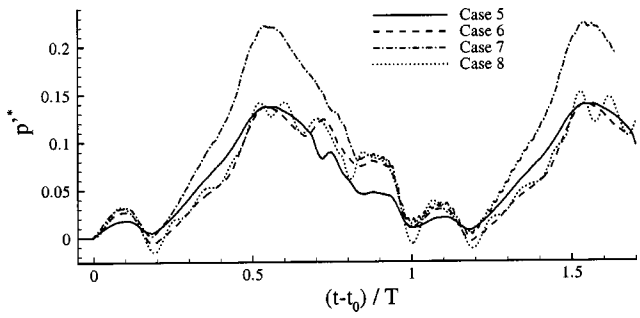


FIG. 10. Time development of acoustic pressure in the farfield.

An examination of Fig. 11(b) reveals that the glottal oscillation frequency has very little effect on the dipole source while increasing the subglottal pressure changes the dipole sources by almost a factor of 2. This is because the dipole source is related to the wall pressure and wall shear stress, which are mainly determined by the Reynolds number, which is directly related to the subglottal pressure. The weak dependence of the dipole strength on frequency implies

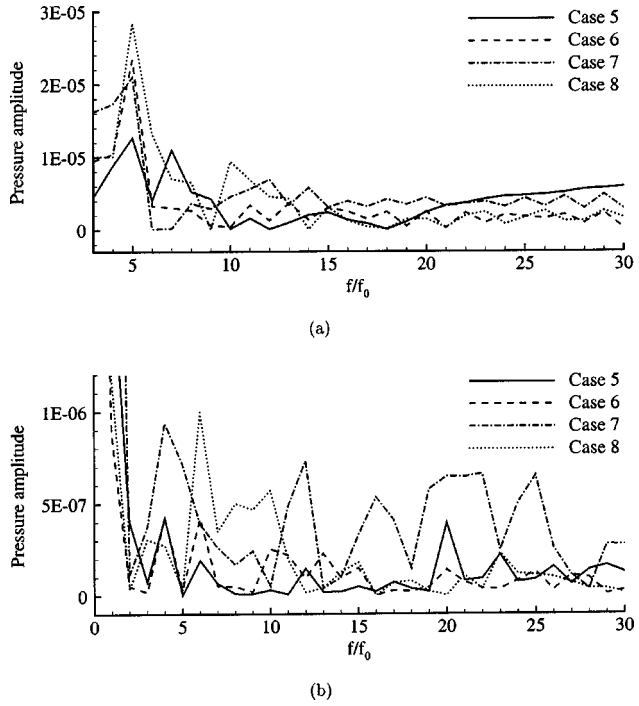


FIG. 12. Fourier components of sound sources: (a) dipole source, (b) quadrupole source.

that the quasisteady approximation may be used to infer the source strength for a given wall geometry and intraglottal pressure drop from a steady flow analysis.

Recall in the previous section that additional high-frequency components in the total sound signal were observed for the case with the ventricular folds when the glottis assumed a diverging shape. These high-frequency components can also be observed at the same time in the cycle from the dipole source contribution. The results of a Fourier analysis, shown in Fig. 12(a), reveal the frequency of these oscillations to be approximately 2500 Hz. This is very close to the resonant frequency predicted by the Rossiter[6] mode 1 of noise produced by flow past an open cavity. This shear layer mode of cavity noise features the roll-up of vorticity in the shear layer and the impingement of this shear layer on the downstream edge of the cavity and was recently studied using CAA techniques by Colonius *et al.*[7] The unstable shear layer that is generated by the glottis is directed past the open cavity formed by the vocal folds and the ventricular folds and impinges on the ventricular folds at the downstream edge of the cavity when the glottis assumes a diverging shape. The Strouhal number of the shear layer mode of cavity noise can be predicted by the modified Rossiter formula:[7]

$$\mathrm{St}=\frac{fL}{U}=\frac{1-\alpha}{M\Big/\sqrt{1+\dfrac{\gamma-1}{2}M^2}+\dfrac{1}{\kappa}}, \tag{6}$$

and the estimated frequency is about 2000 Hz. In Eq. (6), L is the ventricular fold height, the values $\alpha=0.25$ and $\kappa=0.57$ were used following Colonius *et al.* This suggests that the impingement of the pulsating glottal jet on the ventricular folds introduce an additional dipole sound source.[8] The magnitude of this dipole source is significant, but smaller than the dipole source component from the glottal orifice itself. The frequency of the added dipole contributions corresponds roughly to the acoustic resonance frequency of the cavity formed between the vocal folds and the ventricular folds. It is possible that such a higher-frequency component may color the voice spectrum and affect voice quality.

An examination of Fig. 11(c) suggests that an increase in subglottal pressure results in an increase in the quadrupole source contribution, while glottal oscillation frequency has a minimal effect. The presence of the ventricular folds has a minor effect on the quadrupole source, most likely due to a modification of the vortical structure of the jet due to impingement. The quadrupole source is related to vortex shedding. For all four cases the average vortex shedding frequency was close to $20f_0$, although the details of the vortex shedding and vortex interactions were different from case to case. In Fig. 12(b), it is possible that the high-frequency components close to $20f_0$ are due to vortex shedding, since the vortex shedding frequency during one cycle is changing when the glottis is taking different shapes.

It is well known that the surgical removal of the ventricular folds does not affect a patient's voice dramatically, but may change timbre slightly. The results obtained in this investigation tend to confirm this observation, and provide insight on the fundamental nature of these changes. The removal of the ventricular folds suppresses the ''acoustic cavity'' immediately downstream of the glottis. This results in a quenching of the flow-induced cavity oscillations responsible for higher-order harmonic components of the dipole strength. The highly idealized geometries and flow programs assumed in the present numerical simulations, however, do not allow an accurate assessment of the magnitude of these effects in real speech.

IV. CONCLUSIONS

In summary, numerical simulations of the flow and acoustic field in an idealized axisymmetric model of the human vocal tract were performed. The compressible Navier–Stokes equations in a moving coordinate system were integrated using a high-order numerical method. An acoustic analogy was employed successfully to identify specific sound source contributions within the flow. The flow features complex vortical patterns with separation and reattachment occurring during different stages of the vocal fold motion. The directly computed acoustic signal is in excellent agreement with acoustic analogy predictions. The main sound source was identified as a dipole source due to the unsteady force exerted on the surface of the glottis.

The effects of subglottal pressure and glottal oscillation frequency on the velocity, vorticity, wall pressure and shear stress, and acoustic signal of the pulsating jet were investigated. It was found that glottal oscillation frequency influences the rate of change of the fluid volume exiting the glottis and the vortex shedding frequency, which affected the monopole and quadrupole sources. The effect of subglottal pressure, and hence Reynolds number, changed the wall pressure and shear stress, which influenced the dipole source, but had little effect on the monopole and quadrupole sources.

The interaction of the pulsating jet with the ventricular folds was found to be similar to the shear layer mode of cavity noise. This effect introduced additional dipole sources and requires further investigation by numerical and experimental means.

A significant contribution of this study to the understanding of speech production is related to the role of the ventricular folds. The presence of the ventricular folds actually reduces the flow resistance of the glottis. This suggests that the ventricular folds may have an impact on voice production. It would appear, however, that the impact is rather small. The most notable contribution is related to the effects of the grazing flow over the cavity formed by the glottis and the ventricular folds. This contribution is small, and at higher frequencies, at least for the flows considered here.

ACKNOWLEDGMENTS

This research was supported by grant No. RO1 DC03577-05 from the National Institutes of Deafness and Other Communication Disorders, National Institute of Health.

[1] W. Zhao, C. Zhang, S. H. Frankel, and L. Mongeau, ''Computational aeroacoustics of phonation, Part I: computational methods and sound generation mechanisms,'' J. Acoust. Soc. Am. **112**, 2134–2146 (2002).

[2] W. Zhao, S. H. Frankel, and L. Mongeau, ''Large eddy simulation of sound radiation from subsonic turbulent jets,'' AIAA J. **39**, 1469–1477 (2001).

[3] H. Iijima, N. Miki, and N. Nagai, ''Glottal flow analysis based on a finite element simulation of a two-dimensional unsteady viscous fluid,'' in *ICSLP Proceedings*, 1990 International Conference on Spoken Language (The Acoustical Society of Japan, Japan, 1990), Vol. 1.

[4] C.-G. Guo and R. C. Scherer, ''Finite element simulation of glottal flow and pressure,'' J. Acoust. Soc. Am. **94**, 688–700 (1993).

[5] R. C. Scherer, D. Shinwari, K. J. De Witt, C. Zhang, B. R. Kucinschi, and A. A. Afjeh, ''Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees,'' J. Acoust. Soc. Am. **109**, 1616–1630 (2001).

[6] J. E. Rossiter, ''Wind-tunnel experiments on the flow over rectangular cavities at subsonic and transonic speeds,'' Aeronautical Research Council Reports and Memoranda 3438, 1964.

[7] T. Colonius, A. J. Basu, and C. W. Rowley, ''Numerical investigation of the flow past a cavity,'' AIAA Paper AIAA-99-1912, 1999.

[8] P. O. A. L. Davies, R. S. McGowan, and C. H. Shadle, ''Practical flow duct acoustics applied to the vocal tract,'' in *Vocal Fold Physiology: Frontiers in Basic Science*, edited by I. R. Titze (Singular, San Diego, 1993), pp. 93–143.

Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants

Tim Green,[a] Andrew Faulkner, and Stuart Rosen
Department of Phonetics and Linguistics, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, United Kingdom

(Received 18 December 2001; revised 3 July 2002; accepted 10 July 2002)

Four-band and single-band noise-excited vocoders were used in acoustic simulations to investigate spectral and temporal cues to melodic pitch in the output of a cochlear implant speech processor. Noise carriers were modulated by amplitude envelopes extracted by half-wave rectification and low-pass filtering at 32 or 400 Hz. The four-band, but not the single-band processors, may preserve spectral correlates of fundamental frequency ($F0$). Envelope smoothing at 400 Hz preserves temporal correlates of $F0$, which are eliminated with 32-Hz smoothing. Inputs to the processors were sawtooth frequency glides, in which spectral variation is completely determined by $F0$, or synthetic diphthongal vowel glides, whose spectral shape is dominated by varying formant resonances. Normal listeners labeled the direction of pitch movement of the processed stimuli. For processed sawtooth waves, purely temporal cues led to decreasing performance with increasing $F0$. With purely spectral cues, performance was above chance despite the limited spectral resolution of the processors. For processed diphthongs, performance with purely spectral cues was at chance, showing that spectral envelope changes due to formant movement obscured spectral cues to $F0$. Performance with temporal cues was poorer for diphthongs than for sawtooths, with very limited discrimination at higher $F0$. These data suggest that, for speech signals through a typical cochlear implant processor, spectral cues to pitch are likely to have limited utility, while temporal envelope cues may be useful only at low $F0$.
[DOI: 10.1121/1.1506688]

PACS numbers: 43.71.Bp, 43.66.Ts, 43.66.Hg [CWT]

I. INTRODUCTION

Voice pitch information plays an important role in the perception of speech, providing cues to linguistic features such as word emphasis and question–statement contrasts (Highnam and Morris, 1987; Nooteboom, 1997; Wells, Peppe, and Vance, 1995), and also to paralinguistic features such as the age, sex, identity, and emotional state of the speaker (Abberton and Fourcin, 1978; Lieberman and Michaels, 1962). Both spectral and temporal cues to voice pitch are available to normally hearing listeners under normal conditions. In continuous interleaved sampling (CIS), a widely used processing strategy for cochlear implants (Wilson *et al.*, 1991), the electrical stimulation delivered to the auditory nerve represents amplitude envelopes extracted from only a small number of spectral bands. These amplitude envelopes are low-pass filtered, typically at 400 Hz, and imposed on biphasic pulse carriers that generally have a rate of 1–2 kHz. The limited spectral resolution of cochlear implant systems means that the lower harmonics of speech that give normal listeners spectral cues to pitch are not resolved, and the low-pass filtering of the envelopes eliminates fine-structure temporal cues. However, both physiological and psychophysical evidence suggest that the stimulus envelope is unambiguously represented, provided that the carrier pulse rate is 4 to 5 times greater than the maximum modulation frequency (Busby, Tong, and Clark, 1993; McKay, McDermott, and Clark, 1994; Wilson, 1997). Therefore, temporal envelope cues to pitch are, in principle, available in CIS processed speech, as long as the voice fundamental frequency ($F0$) range is passed by the envelope smoothing filter and the pulse rate is high enough.

Direct evidence of this was provided by a study involving users of the LAURA cochlear implant (Geurts and Wouters, 2001). An initial criterion of sensitivity to changes of 20% in the modulation frequency of a sinusoidally amplitude-modulated pulse train for modulation rates around 150 Hz was met by four of eight subjects. The ability of the selected subjects to discriminate differences in the $F0$ of synthesized steady-state vowels was compared in conditions in which the low-pass cutoff frequency of the envelope filter was either 400 or 50 Hz. Listeners used either 7 or 8 electrodes, with pulse rates of 1429 or 1250 pulses per second on each electrode. With the 400-Hz envelope filter, performance was relatively good: with a standard $F0$ of 150 Hz the difference in $F0$ that could be reliably discriminated varied between 6 and 20 Hz across listeners. With the 50-Hz envelope filter, which eliminated temporal envelope fluctuations related to $F0$, performance was generally very much worse. However, there was evidence that stimuli with different $F0$s could occasionally be discriminated on the basis of loudness cues. Although stimuli were balanced in root-mean-square level, the average amplitude of the pulses in individual channels was different for different $F0$s, due to the associated

[a]Electronic mail: tim@phon.ucl.ac.uk

changes in the frequencies of the harmonics of $F0$. For one particular combination of vowel (/i/) and $F0$ region (around 250 Hz), it appeared that the pattern of average amplitude changes across channels with $F0$ was sufficiently regular to allow reliable discrimination.

In several recent studies CIS processing has been simulated using noise-excited vocoder processing (Dorman, Loizou, and Rainey, 1997; Rosen, Faulkner, and Wilkinson, 1999; Fu and Shannon, 1999; Shannon *et al.*, 1995; Shannon, Zeng, and Wygonski, 1998). Bandpass filters divide the spectrum into broad frequency bands and half-wave rectification and low-pass filtering are employed to extract the temporal envelope in each band. In contrast to the traditional vocoder there is no estimation of voicing and voice pitch. Instead, the extracted temporal envelopes modulate white-noise carriers with the same bandwidths as the initial bandpass filters. By using only a small number of frequency bands, this form of processing mimics the limited information on the spectral distribution of energy available through cochlear implant systems. Temporal information carried by these simulations also mimics that carried by CIS implant processors, with temporal fine structure being eliminated by envelope smoothing, and lower rate temporal envelope information being preserved. Provided that there is sufficient envelope bandwidth, such noise-excited vocoding is capable of conveying pitch information for modulation rates up to a few hundred Hz, as indicated by studies using amplitude-modulated noise (e.g., Burns and Viemeister, 1976, 1981; Pollack, 1969). This suggests that the temporal envelope cues to voice pitch available with noise-excited vocoding are broadly similar to those available from CIS processing.

The utility of such temporal envelope pitch cues has thus far received little attention in simulation studies. However, Fu *et al.* (1998) compared the ability of Chinese-speaking listeners to identify Mandarin Chinese's four lexical tonal patterns, characterized by different $F0$ contours, under various noise-excited vocoding processing conditions. On each trial listeners were presented with a processed version of a single syllable consisting of an initial consonant and a following vowel with a particular tone, recorded from a single adult male speaker. They were required to identify the tone from a list containing four alternative syllables, each of which had the same consonant and vowel. Performance was significantly better when the envelope filter cutoff frequency was 500 Hz rather than 50 Hz, suggesting that listeners were able to make use of the temporal envelope cues to voice pitch that were available when the envelope filter covered the $F0$ range. However, performance was well above chance even with the 50-Hz filter, consistent with evidence that cues other than the $F0$ pattern, such as amplitude contour and duration, also contribute to tone recognition (Fu and Zeng, 2000; Whalen and Xu, 1992).

Further evidence regarding temporal envelope pitch cues in simulations of CIS processing was provided by Faulkner, Rosen, and Smith (2000). In the context of an investigation of the influence of pitch and periodicity information on speech intelligibility, they constructed a number of four-channel vocoder simulations in which the salience of such information was varied. There were two noise-carrier processors, in which the low-pass cutoff frequency of the envelope filters was either 400 or 32 Hz, the latter eliminating temporal information in the voice pitch range. Other processors used a 32-Hz cutoff frequency and controlled the carrier rather than the envelope to encode pitch and periodicity information. For example, in one processor, designated *FxNx*, the carrier during voiced speech was a pulse train whose frequency followed the fundamental frequency of the speech input, while the carrier during voiceless speech was random noise.

Although speech intelligibility, as assessed in tests of consonant and vowel identification, BKB sentence perception, and connected discourse tracking, was found to differ only slightly across the different processing conditions, it is likely that this is because such tasks lack sensitivity to aspects of speech such as intonation, which nevertheless are likely to be important factors in speech communication. That the processing conditions did differ in the extent to which they conveyed pitch information was confirmed by performance in a frequency glide discrimination task. Listeners were required to categorise as either ''rising'' or ''falling'' in pitch sawtooth waves whose fundamental frequency changed smoothly over their 500-ms duration. Performance was very high with the *FxNx* processor, even with small start-to-end frequency ratios, as would be expected since the fundamental frequency was directly represented in the carrier. Performance with the noise-carrier processor with the 400-Hz cutoff envelope filter was substantially worse, being high for large start-to-end frequency ratios but declining to near-chance levels for small ratios. Interestingly, performance with the noise processor with the 32-Hz envelope cutoff frequency was only slightly lower than that with the 400-Hz envelope processor despite the elimination of temporal envelope cues to pitch variation. This was attributed to differences in spectral envelope caused by harmonics of the input signal shifting between analysis bands of the processor.

The current study provides a more detailed investigation of the pitch cues available from noise-carrier processors, allowing a dissociation of the contributions of temporal envelope and spectral cues. In our first experiment, sawtooth glide labeling performance was examined with single-band processors with 32- and 400-Hz envelope bandwidths in addition to the two four-band noise-carrier processors described above. With these single-band processors the output was unaffected by the number of harmonics of the sawtooth signal falling in each analysis band, eliminating spectral cues correlated with $F0$. A second experiment examined listeners' ability to identify the direction of pitch glides when stimuli consisted of synthesized diphthongal vowels, allowing an assessment of the impact of variations in formant structure on cues to voice pitch.

II. EXPERIMENT 1: SAWTOOTH GLIDES

A. Participants

Seven normally hearing listeners aged between 25 and 50 years participated, including the three coauthors, three members of departmental staff, and one postgraduate student who was paid £5 per hour.

TABLE I. Start and end $F0$ (in Hz) for each "rise" glide in experiment 1. Start and end $F0$ were reversed for "fall" glides.

	146-Hz center $F0$		208-Hz center $F0$		292 Hz center $F0$	
Ratio	Start	End	Start	End	Start	End
0.50	103	207	147	293	207	413
0.57	110	194	157	275	220	388
0.64	117	182	167	258	235	364
0.73	125	171	177	243	250	342
0.83	133	161	189	228	266	321
0.93	141	151	201	214	283	302

B. Stimuli

Stimuli were generated off-line with a 20-kHz sample rate using MATLAB. They consisted of noise-carrier processed sawtooth-wave frequency glides. The sawtooth waves were synthesized by adding, with the appropriate amplitude and phase relationships, the first ten harmonics of a swept fundamental frequency that changed logarithmically over the 500-ms duration. The ratio of start and end frequencies varied in six equal logarithmic steps from 1:0.5 to 1:0.93. Three ranges of fundamental frequency were used, with the $F0$ at the midpoint in time (i.e., the geometric mean of the start and end $F0$s) of each glide being 146, 208, and 292 Hz, respectively. For each ratio and each $F0$ range there was one ascending and one descending glide, giving a total of 36 different glides. The start and end $F0$s of each glide are shown in Table I.

Four-band noise vocoding comprised the following sequence of steps: analysis bandpass filtering (sixth-order Butterworth) to divide the spectrum into frequency bands; half-wave rectification and low-pass filtering (second-order Butterworth) to extract the amplitude envelope for each band; modulation of a noise carrier by each envelope; output filtering matching the initial analysis filtering; summation across channels. The cutoff frequencies of the analysis and output filter bands were based on equal basilar-membrane distance (Greenwood, 1990). The filter slopes crossed at their −3-dB cutoff frequencies, which were 100, 392, 1005, 2294, and 5000 Hz. In the single-band conditions processing was identical except that the four analysis filters and envelope extractors were replaced with a 50-Hz high-pass filter, followed by a single envelope extractor whose output modulated the level of each of the four output bands. The cutoff frequency of the high-pass filter was set well below the lowest $F0$ used to avoid amplitude cues that might have resulted from attenuation at $F0$. Both four-band and single-band processing were carried out with the cutoff frequency for the low-pass envelope extraction filter at either 32 or 400 Hz, making a total of four processing conditions designated as *Single32, Single400, Four32,* and *Four400*.

Figure 1 shows a spectrogram of an unprocessed rising sawtooth glide from the highest $F0$ range and with the largest start-to-end frequency ratio. Also shown are the −3-dB cutoff frequencies of the analysis filters. The movement of harmonics between analysis bands can be clearly seen. This is expected to provide spectral cues to pitch in four-band noise-vocoding conditions.

C. Procedure

Stimuli were presented through the right earpiece of Sennheiser HD 414 headphones at a comfortable listening level (peak of 85–90-dB SPL measured over an 80-ms window). On each trial subjects heard a single glide and were required to identify it as either "rising" or "falling" in pitch. They responded via computer mouse by clicking on an image of either a rising or falling line. No feedback was given. Before each block of trials subjects were able to listen to a selection of the glides to be presented in that block, visually labeled as rising or falling. To familiarize themselves with the task subjects were first presented with a block of trials in which the glides were unprocessed. Processing condition was then varied across blocks of trials consisting of three repetitions of the 36 glides in random order. The order in which blocks of trials were presented was random with the constraint that each set of four blocks contained one block with each type of processing. Subjects completed five blocks of trials for each processing condition, the first of which was treated as practice.

D. Results and discussion

Mean psychometric functions for the proportion of "fall" responses as a function of the ratio between start and

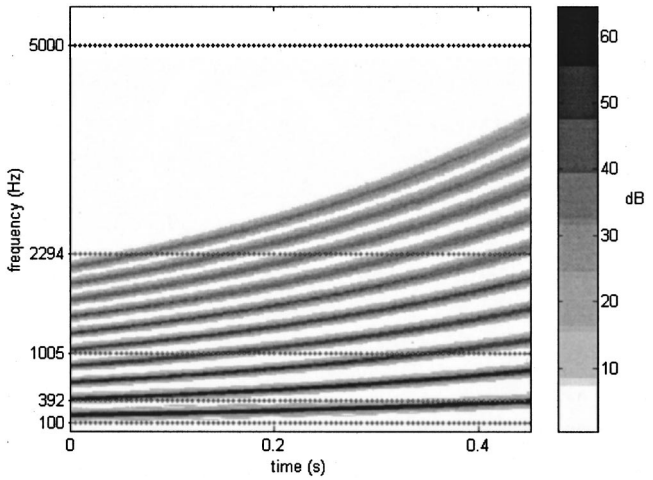


FIG. 1. Spectrogram of an unprocessed rising sawtooth glide (highest $F0$ range, largest frequency ratio) used in experiment 1. The spectrogram was calculated in MATLAB using 10-Hz frequency intervals and 50-ms Hanning windows with overlap of 47.5 ms. Dotted lines show the −3-dB cutoff frequencies of the analysis filters.

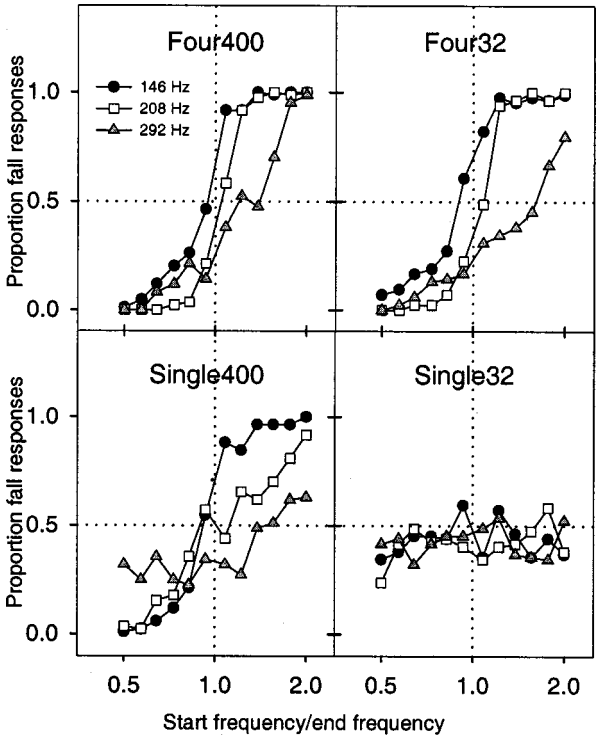


FIG. 2. Proportion of "fall" responses summed across listeners as a function of start-to-end frequency ratio for each processing condition of experiment 1.

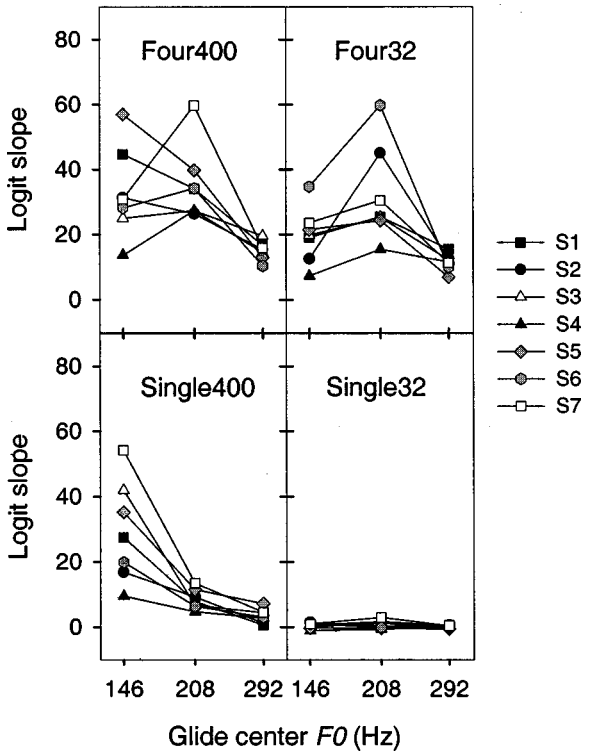


FIG. 3. Slopes of logistic regressions of the proportion of "fall" responses as a function of the log (base 10) of start-to-end frequency ratio for each listener in each processing condition of experiment 1.

end frequencies in each processing condition are shown in Fig. 2. A logistic regression was carried out on the proportion of fall responses as a function of the log (base 10) of the start-to-end frequency ratio for each processing condition and center $F0$ for each subject. None of the fits deviated significantly from the observed data according to χ^2 tests with 10 degrees of freedom. In the *Single32* condition, regression intercept values indicated that most listeners showed a slight bias towards "rise" responses, indicated by negative intercepts. In the other three conditions intercepts varied substantially across listeners, though several listeners showed a bias towards fall responses for glides in the lowest $F0$ range and for a bias towards rise responses for those in the highest $F0$ range.

Figure 3 shows the regression slope estimates, with larger values indicating greater discriminability. For example, for unbiased responding, a slope of 10 corresponds to a proportion correct of 0.58 for glides with the smallest frequency ratio (0.93), while for a slope of 30 the corresponding proportion would be 0.72. A 95% confidence interval was calculated for each slope estimate, but for clarity these are not displayed in Fig. 3. Confidence intervals increased in size with increasing slope. For example, for a slope value of 10 the confidence interval was typically around ± 3, while for a slope of 30 it was around ± 12. Slope estimates were analyzed using a two-way repeated-measures analysis of variance (ANOVA) with factors of processing condition and center $F0$. Since higher slope estimates have a higher standard error, data were logarithmically transformed before analysis. To ensure that there were no negative values, 2 was added to all values before the (base 10) logarithms were taken. The reported F tests used Huynh–Feldt epsilon correction factors. There were significant effects of processing condition [$F(3,18)=178.60$, $p<0.001$] and center $F0$ [$F(2,12)=25.08$, $p<0.001$], and also a significant interaction between these two factors [$F(6,32)=17.29$, $p<0.001$].

The nature of these effects and their interaction can be interpreted from the patterns of performance shown in Figs. 2 and 3. In the *Single32* condition listeners were clearly unable to distinguish between rising and falling glides. For each $F0$ range the mean proportion of fall responses is effectively constant regardless of start-to-end frequency ratio (Fig. 2), and the slopes of the logistic regressions are virtually zero for all listeners and $F0$ ranges (Fig. 3). This is consistent with the expectation that neither spectral nor temporal cues to pitch are available with this form of processing. In the *Single400* condition, in which there are temporal envelope cues to pitch but no spectral information, performance varies substantially according to $F0$ range. For the lowest $F0$ range, the slope of the mean psychometric function is quite steep and performance is near ceiling for several of the larger frequency ratios (Fig. 2), indicating that listeners found it relatively easy to distinguish between rising and falling glides. However, the functions become less steep with increasing glide center $F0$. For the highest $F0$ range there is only a small change in the proportion of fall responses according to start-to-end frequency ratio, and performance is well below ceiling for even the largest ratios. The logistic regressions for individual listeners (Fig. 3) all show a monotonic decrease in slope estimates as glide center $F0$ increases, although there is substantial variation in slope values for the lowest $F0$ range. This decline in the utility of tem-

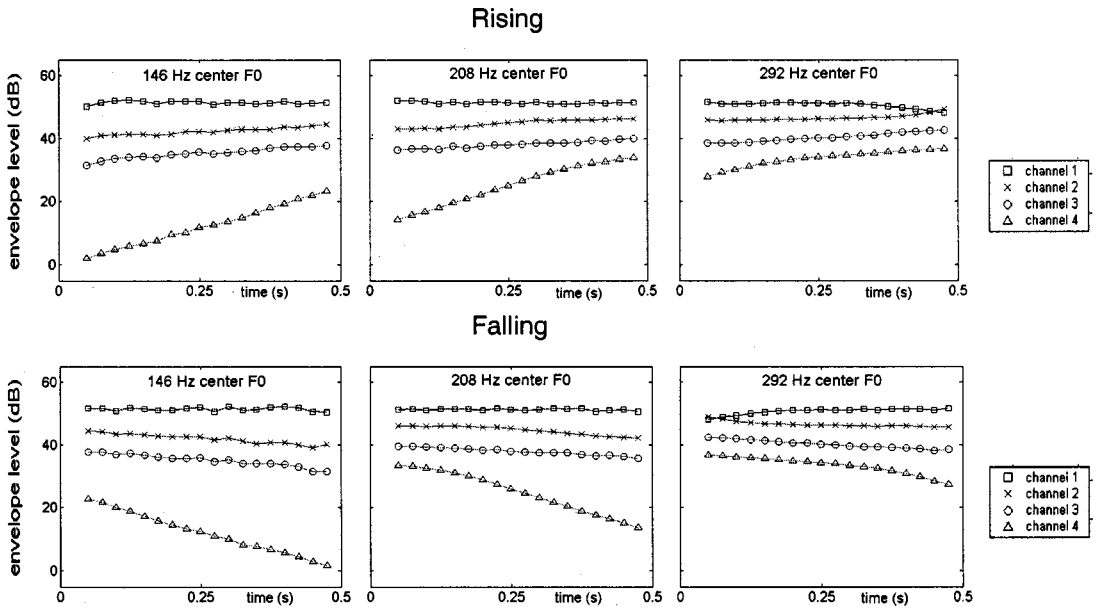


FIG. 4. The level of the envelope in each channel over time for the steepest rising (top row) and steepest falling (bottom row) sawtooth glides in each $F0$ range in the *Four32* condition.

poral envelope cues to pitch with increasing $F0$ is consistent with previous evidence from studies using amplitude-modulated noise that modulation-rate discrimination thresholds increase substantially with increases in modulation rate over the range of center $F0$ values used here (Grant, Summers, and Leek, 1998; Hanna, 1992).

The *Four32* condition isolated the contribution of spectral cues in the absence of temporal envelope information. Performance in this condition again varied according to $F0$ range, but in a markedly different way to that apparent in the *Single400* condition. The mean psychometric functions (Fig. 2) indicate that the highest discriminability occurred with the 208-Hz center $F0$, and the lowest with the 292-Hz center $F0$. For each individual listener the highest regression slope value was observed for the middle $F0$ range (Fig. 3). For five out of seven listeners slopes were greater for the low $F0$ range than the high $F0$ range, though there is considerable variability in slope estimates across listeners for the two lower center $F0$ values. Pitch information, encoded as spectral differences that arise as harmonics of the glide stimuli move between analysis bands, is clearly available for all three $F0$ ranges tested. That this information is more salient with a glide center $F0$ of 208 Hz may be attributable to the fact that the second harmonic of stimuli at this center $F0$ will be changing between values centered at 416 Hz, close to the point at which the slopes of the first two analysis filters cross (392 Hz). Another possible explanation concerns the amount of energy in the highest channel. As illustrated in Fig. 4, which plots the level of the envelopes extracted from each analysis band for the steepest rising and steepest falling glide in each $F0$ range, for glides in the lowest $F0$ range there is always only a small amount of energy in this channel. Conversely, for glides in the highest $F0$ range, the envelope level in the highest channel is always relatively high. However, for the higher start-to-end frequency ratio glides with a center $F0$ of 208 Hz, the amount of energy in the highest channel changes from relatively high to very low (or vice versa) over the course of the glide. This may provide a useful cue to the direction of change in $F0$.

In the *Four400* condition both spectral and temporal envelope cues to pitch variation were available. The mean psychometric functions (Fig. 2) are very similar to those obtained in the *Four32* condition for the two lower $F0$ ranges, while a slightly steeper slope is apparent with the 292-Hz center $F0$. In contrast to the *Single400* and *Four32* conditions, the slope values for individual listeners (Fig. 3) do not conform to a single pattern. For three listeners (S1, S2, and S5) the pattern of slope values is similar to that in the *Single400* condition, with slope values declining consistently as glide center $F0$ increases, while for the other four listeners the pattern is similar to that in the *Four32* condition, with the highest slope value occurring with the 208-Hz glide center $F0$. The close similarity in the mean psychometric functions for 146-Hz center $F0$ stimuli across the *Single400*, *Four400*, and *Four32* conditions would suggest that there is little integration of spectral and temporal sources of pitch information.

This is broadly consistent with previous evidence pointing to the independence of temporal and spectral information. For example, for implant users, rate and place of stimulation appear to provide independent cues to pitch (McKay, McDermott, and Carlyon, 2000; Tong *et al.*, 1983). Similarly, Demany and Semal (1993) found that normal-hearing listeners were not consistently better at detecting combined shifts in $F0$ and formant frequency when the formants changed in the same direction as $F0$ than when they changed in opposite directions. There was, though, substantial variability across subjects in the latter study and similarly here, the variation in the pattern of slope values across glide center $F0$ in the *Four400* condition indicate that there are considerable individual differences in the ways in which temporal envelope and spectral cues are combined and weighted.[1]

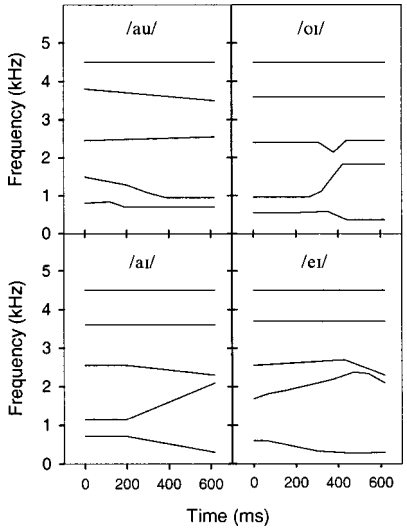


FIG. 5. Trajectories of the formant frequencies of each of the synthesized diphthongs used in experiment 2.

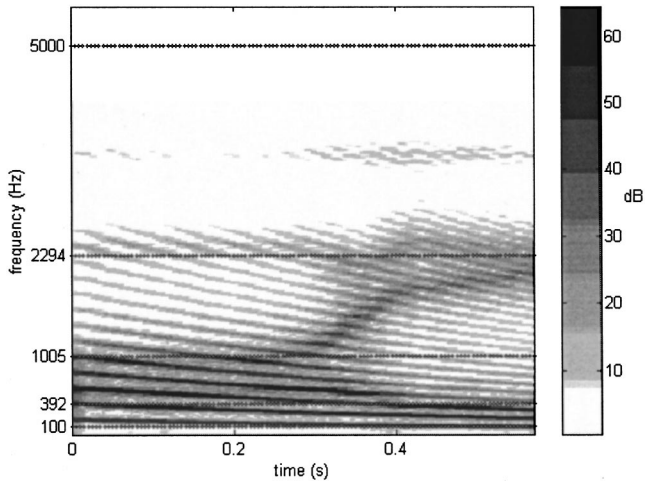


FIG. 6. Spectrogram of an unprocessed falling diphthong (/oɪ/, lowest $F0$ range, largest frequency ratio) used in experiment 2. The spectrogram was calculated as in Fig. 1.

Experiment 1 has demonstrated that, in addition to the temporal envelope cues available at low modulation rates, noise-excited vocoding can also provide spectral cues to pitch due to the movement of harmonics between the processor's analysis bands. However, it appears likely that when a series of carrier bands are modulated by envelopes extracted from speech, rather than from a steady-state periodic sawtooth waveform, spectral cues to pitch are likely to be diminished by the time-varying spectral envelope of speech. This is investigated in experiment 2.

III. EXPERIMENT 2: SYNTHESIZED DIPHTHONGS

A. Stimuli

Diphthongs /aʊ/, /eɪ/, /aɪ/, and /oɪ/ with a duration of 620 ms were created using an implementation of the KLSYN88 Klatt synthesizer in cascade mode with a sample rate of 20 kHz and parameters specified every 5 ms. Formant values were estimated by examining recordings of each diphthong embedded in a cVc context spoken by a male Southern British English speaker. The formant frequency trajectories for each diphthong are portrayed in Fig. 5. For each diphthong, $F0$ was varied in the same way as in experiment 1, i.e., in each of three $F0$ ranges there were six rising and six falling versions, making a total of 144 different stimuli. The start-to-end frequency ratios and center $F0$ values used were identical to those in experiment 1. The same formant values were used regardless of the variation in $F0$.[2] Figure 6 shows a spectrogram of an unprocessed falling $F0$ /oɪ/ from the lowest $F0$ range and with the largest start-to-end frequency ratio. As with the unprocessed sawtooth glide (Fig. 1), the movement of harmonics between analysis bands is apparent. Note, however, that in this case the distribution of energy over frequency is also strongly influenced by the time-varying formant structure of the diphthong. As a consequence, the levels of the envelopes extracted from each analysis band, displayed in Fig. 7, show much greater, and less consistent, variation over time than was the case for the sawtooth glides. The *Single32* condition was omitted, but in all other respects noise-excited vocoder processing was identical to experiment 1.

B. Subjects and procedure

Six of the seven subjects from experiment 1 completed seven blocks of trials for each processing condition, with the first block regarded as practice. Each block consisted of one presentation of each of the 144 combinations of diphthong, $F0$ range, and start-to-end frequency ratio. Other details were as in experiment 1.

C. Results and discussion

Figure 8 shows the proportion of fall responses averaged across the six listeners and the four diphthongs as a function of the ratio between start and end frequencies for each $F0$ range and processing condition. Logistic regressions were carried out as in experiment 1 and again, none of the fits deviated significantly from the observed data. Intercept values showed substantially less variation than in experiment 1 and generally indicated largely unbiased responding, except in the *Four400* condition where most listeners showed a small bias towards fall responses for the lowest $F0$ range and towards rise responses for the highest $F0$ range. Regression slope estimates, displayed in Fig. 9, were transformed and submitted to an ANOVA as in experiment 1. There were significant effects of processing condition [$F(2,10)=33.52$, $p<0.001$] and center $F0$ [$F(2,10)=14.74$, $p=0.001$], and also a significant interaction between these two factors [$F(4,20)=6.17$, $p=0.002$].

For all three $F0$ ranges, mean performance in the *Four32* condition was barely above chance levels even at the largest frequency ratios, and the slopes of the logistic regressions are close to zero for all listeners and $F0$ ranges. This indicates that the variations in spectral envelope associated with diphthongal stimuli obscured the spectral cues to pitch that listeners were able to utilize effectively under the same processing conditions with the sawtooth glides used in experiment 1. Figure 10 shows mean psychometric functions

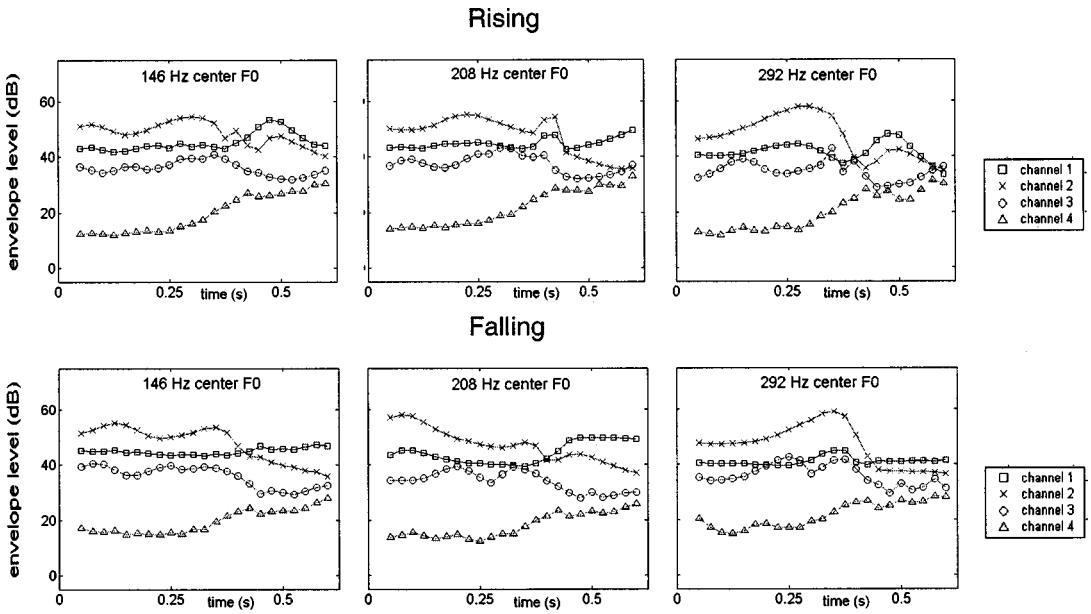


FIG. 7. The level of the envelope in each channel over time for the steepest rising (top row) and steepest falling (bottom row) /oɪ/ glides in each $F0$ range in the *Four32* condition.

for each of the four diphthongs for each processing condition and center $F0$. Consistent with the above argument, it is noticeable that responses in the *Four32* condition appear to be largely determined by the identity of the vowel rather than the start-to-end frequency ratio: /aʊ/ was generally perceived as falling in pitch and /eɪ/ and /oɪ/ as rising in pitch, regardless of the actual direction of $F0$ change.

Comparison of performance across experiments 1 and 2 is complicated somewhat by the fact that the sawtooth waveform stimuli in experiment 1 were of a shorter duration (500 ms) than the synthesized diphthongs used in experiment 2 (620 ms), so that $F0$ was changing slightly more quickly in the former case. However, it seems unlikely that this small difference in stimulus duration would have more than a minor effect on performance. This is supported by the results of extra blocks of trials carried out by subject 3 using sawtooth waveform glides of 620-ms duration in processing conditions *Four400*, *Four32*, and *Single400*. Slopes of logistic regressions carried out on these data differed, according to 95% confidence intervals, from the results obtained with the 500-ms glides of experiment 1 in only two cases (the two higher center $F0$ ranges in the *Single400* condition). In both of these cases slope values were actually higher with the 620-ms glides, suggesting that the longer duration of the stimuli in experiment 2 does not contribute to the poorer performance relative to experiment 1.

In the *Single400* condition mean performance is at or very near chance levels for the two higher $F0$ ranges, while there is some limited pitch information available in the lowest $F0$ range. Regression slopes vary little across listeners and are in line with the pattern of the mean data, being at or near zero for the two highest $F0$ ranges and just above zero for the lowest center $F0$. This is in contrast to the results of experiment 1, in which substantially steeper slopes were apparent for the lowest center $F0$. This reduction in performance can be attributed to the more complex temporal envelope structure associated with diphthongal stimuli compared to sawtooth waveforms. Because of the more variable spectral shape of the diphthongal stimuli, $F0$-related fluctuations in the extracted amplitude envelope are less distinct than is the case with sawtooths.

In the *Four400* condition the slopes of the mean psychometric functions decreased monotonically with increasing center $F0$. With the exception of subject 4, whose performance was at or near chance for all $F0$ ranges, individual listeners' regression slope values follow a similar pattern, being near zero for the 292-Hz center $F0$, slightly above zero for the 208-Hz center $F0$, and substantially above zero for

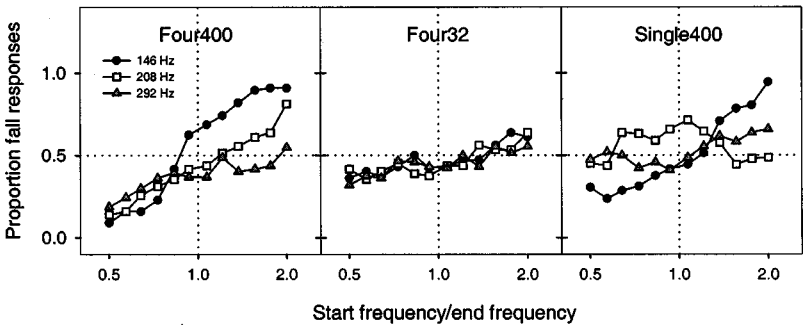


FIG. 8. Proportion of "fall" responses summed across listeners as a function of start-to-end frequency ratio for each processing condition of experiment 2.

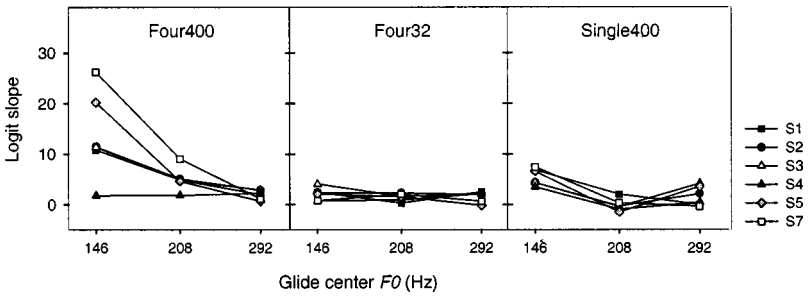


FIG. 9. Slopes of logistic regressions of the proportion of "fall" responses as a function of the log (base 10) of start-to-end frequency ratio for each listener in each processing condition of experiment 2.

the 146-Hz center $F0$. As was the case in the *Single 400* condition of experiment 1, in which only temporal envelope cues were available, there was considerable variability in performance across listeners. In general, the listeners who performed best in the *Single 400* condition of experiment 1 also have the highest slope values here, indicating a substantial degree of consistency in the ability to make use of temporal envelope cues to pitch. Comparing performance in the *Four400* and *Single400* conditions of experiment 2, it appears that, for the limited range of $F0$ values in which temporal cues to pitch are available, such cues are more effective when envelopes are extracted from four separate channels rather than a single broad bandwidth channel. It is possible that this advantage for the four-band condition is attributable to information carried in the envelope extracted from the lowest frequency channel, the cutoff frequencies of which (100–392 Hz) cover the $F0$ range. This may allow the listener to isolate $F0$ related modulation information from other modulations inherent in the diphthong.

IV. GENERAL DISCUSSION AND CONCLUSIONS

A. Pitch cues in noise-excited vocoding

Despite the very limited spectral resolution and the use of a 32-Hz cutoff envelope filter which removed $F0$-related temporal fluctuations, relatively good glide-labeling performance was obtained in the *Four32* condition of experiment 1. This confirms the conclusion of Faulkner *et al.* (2000) that with sawtooth waveform stimuli, spectral envelope shifts arising from harmonics of the input waveform moving between analysis bands provide cues to pitch variation. However, the chance levels of performance apparent with synthesized diphthongal stimuli in the *Four32* condition in experiment 2 demonstrate that such spectral cues were obscured by the presence of variations in spectral structure typical of speech.

The effectiveness of temporal envelope cues was limited to the lower end of the typical range of voice $F0$. When spectral cues were eliminated by using single-band processing or when input consisted of synthesized diphthongs, discrimination of pitch variation was severely limited for stimuli in the two higher $F0$ ranges. In addition, comparison across the *Single400* conditions of experiments 1 and 2 suggests that the utility of temporal cues to pitch is much reduced when stimuli have a more complex temporal envelope typical of spectrally varying speech. With such stimuli, only very limited pitch discrimination was possible in the *Single400* condition. For all but one listener, the utility of temporal envelope cues with diphthongal stimuli in the low $F0$ range was substantially higher with four bands rather than a single band. At first sight this result is in contrast to the finding reported by Fu *et al.* (1998) that recognition of noise-excited vocoder processed Mandarin tones, averaged across the four tones tested, did not differ according to the number of analysis bands (1, 2, 3, or 4). In that study, however, there were unequal numbers of each type of tone in the different analysis band conditions (Fu, 2001, personal communication). The extent to which amplitude contour is correlated to $F0$ pattern, and therefore provides a cue to tone recognition in the absence of pitch information, differs markedly across tones. Therefore, this unequal distribution of tones might have obscured an effect of the number of analysis bands. Indeed, this is suggested by Fu *et al.*'s (1998) Fig. 2, which presents recognition performance for individual tones in the one-band and four-band conditions with 500-Hz envelope filtering, averaged across those subjects who also completed a sentence recognition task. For three of the four tones performance was substantially higher with four bands rather than one, consistent with the better performance in the *Four400* relative to the *Single400* condition observed with the synthesized diphthongal stimuli in the present study.

B. Implications for cochlear implants

In the four-channel noise-excited vocoding simulation of CIS processing used in the current experiments, spectral cues to pitch variation arising from the movement of harmonics

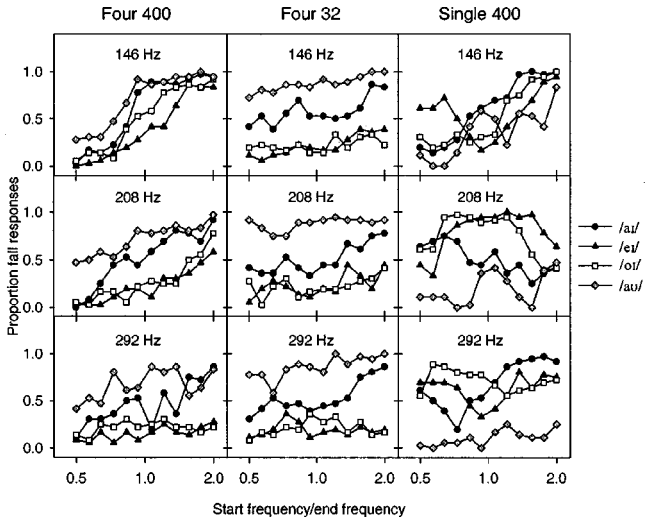


FIG. 10. Proportion of "fall" responses summed across listeners as a function of start-to-end frequency ratio for each vowel in each processing condition of experiment 2.

between analysis bands were rendered ineffective by the spectral variation inherent to synthesized diphthongs. Some caution is necessary in assessing the implications of this finding since current evidence suggests that cochlear implant users typically have six to eight effective channels available (e.g., Fishman, Shannon, and Slattery, 1997; Friesen *et al.*, 2001). With this greater number of channels it is conceivable that some useful cues to pitch variation might be derived from the shifting of harmonics between analysis bands even in the presence of speech-like spectral variation. Also, natural speech may provide implant users with additional cues to pitch not present in our synthesized stimuli, such as the tendency for amplitude contours to covary with $F0$. On the other hand, it is also the case that real speech typically contains greater spectral variation than is present in the stimuli used here. While not conclusive, the current data suggest that, with speech stimuli, reliable perception of pitch variation in noise-excited vocoding simulations of CIS processing will primarily depend upon temporal envelope cues.

With regard to temporal cues, a limitation of noise-excited vocoder simulations must be acknowledged. Because the carrier in a CIS processor is a high rate stream of pulses rather than random noise, the $F0$-related modulation of the carrier is noise-free. Therefore, neural responses to this pulse-carrier stimulation are likely to be more strongly synchronized to the modulation (Wilson, 1997), so it is possible that pitch information derived from temporal envelope cues will be more salient in CIS processing than in noise-carrier simulations. Note though, that the temporal complexity of envelopes derived from typical speech suggests that the neural firing patterns resulting from electrical stimulation are unlikely to represent clearly the period of $F0$. Only very limited information is available regarding implant users' perception of pitch variation signaled by modulation of the amplitude of pulse trains, as in CIS processing. In addition to the data obtained by Geurts and Wouters (2001) using synthesized vowel stimuli, the pitch percepts elicited by sinusoidally amplitude modulated pulse trains have been studied in selected subjects (McDermott and McKay, 1997; Wilson *et al.*, 1997). There is, though, a lack of information on typical performance over a range of modulation rates. There must, therefore, be some uncertainty regarding the extent to which performance in noise-carrier simulation studies accurately reflects the likely performance of implant users. However, on the assumption that simulation data gives at least a reasonably accurate indication of the temporal envelope cues to pitch that are available to implant users, the current results prompt important considerations regarding the implementation of CIS processing strategies.

It would appear that with current processing methods the temporal envelope cues that provide implant users with information on voice pitch variation are severely limited for fundamental frequencies above 200 Hz, and that even for lower $F0$ values, their utility is substantially reduced in the presence of spectral variation typical of speech. Given the important contribution of pitch information to speech understanding in everyday situations, methods of processing that enhance the availability of temporal pitch cues would be expected to be of substantial benefit to implant users, though as Faulkner *et al.* (2000) note, more sensitive speech perception tests than those generally used currently may be required in order to demonstrate such benefits in the laboratory. Geurts and Wouters (2001) devised one CIS processing algorithm designed to enhance the availability of pitch information relative to standard CIS processing. One feature of this algorithm was the elimination of phase distortion from the analysis filters, so that the maxima of $F0$-related fluctuations in each channel of the pulsatile output coincided in time. A second was the use of two envelope filters with cutoff frequencies of 400 and 50 Hz. The output of the 50-Hz filter was subtracted from the output of the 400-Hz filter in order to increase the modulation depth of $F0$-related fluctuations. However, there were no significant differences in the ability to discriminate changes in the $F0$ of synthesized steady-state vowels processed with this algorithm relative to standard CIS processing. Other strategies for enhancing temporal envelope cues to pitch need to be investigated.

That temporal envelope cues appear to be restricted to lower $F0$ ranges is of particular concern given the increasing prevalence of implantation in very young children. Intonation is widely held to play an important role in early language development (Jusczyk, 1997), and is markedly exaggerated in child-directed speech (Fernald *et al.*, 1989). However, the voice pitch range of young children covers around 250 to 400 Hz, while the female voice pitch range typically extends well above 200 Hz. Therefore, with current processing strategies, implanted young children will be unable to perceive much of the pitch information in their own speech and in other speech to which they are exposed during development. Substantial benefits to language development might be expected to result from the availability of information concerning pitch variation over a higher $F0$ range. Since it is the pattern of pitch change that carries most of the relevant information, rather than absolute pitch values *per se*, one potential solution might involve lowering the rate of $F0$-related modulations. A rate-lowering approach has been shown to be effective in the context of an aid to lip-reading consisting of a single electrode external to the cochlea that provided $F0$ information (Fourcin *et al.*, 1984).

ACKNOWLEDGMENTS

This study was supported by the Royal National Institute for Deaf People (UK). We are grateful to Chris Turner, Ed Burns, and two anonymous reviewers for helpful comments on a previous version of this paper.

[1]Given the independence between spectral and temporal envelope sources of information, it might be considered questionable whether the property of spectral cues that allows listeners to make high/low judgements is in fact the same percept that results from temporal envelope cues. Here, though, we regard both types of information as cues to pitch.

[2]This differs from real speech, in which a higher $F0$ is typically accompanied by higher formant frequencies. In addition, because the first formant frequency ($F1$) of three of the four synthesized vowels decreases towards around 300 Hz (Fig. 5), $F0$ was close to or higher than $F1$ towards the end of the duration of several of the stimuli in the highest $F0$ range. Nonetheless, this simplified stimulus set was expected to address adequately the central question of the effects of speech-like spectral variation on spectral and temporal pitch cues.

Abberton, E., and Fourcin, A. (**1978**). ''Intonation and speaker identification,'' Lang Speech **21**, 305–318.
Burns, E. M., and Viemeister, N. F. (**1976**). ''Nonspectral pitch,'' J. Acoust. Soc. Am. **60**, 863–869.
Burns, E. M., and Viemeister, N. F. (**1981**). ''Played-again SAM: Further observations on the pitch of amplitude-modulated noise,'' J. Acoust. Soc. Am. **70**, 1655–1660.
Busby, P. A., Tong, Y. C., and Clark, G. M. (**1993**). ''The perception of temporal modulations by cochlear implant patients,'' J. Acoust. Soc. Am. **94**, 124–131.
Demany, L., and Semal, C. (**1993**). ''Pitch versus brightness of timbre: Detecting combined shifts in fundamental and formant frequency,'' Music Percept. **11**, 1–14.
Dorman, M. F., Loizou, P. C., and Rainey, D. (**1997**). ''Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs,'' J. Acoust. Soc. Am. **102**, 2403–2411.
Faulkner, A., Rosen, S., and Smith, C. (**2000**). ''Effects of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: Implications for cochlear implants,'' J. Acoust. Soc. Am. **108**, 1877–1887.
Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., and Fului, I. (**1989**). ''A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants,'' J. Child. Lang. **16**, 477–501.
Fishman, K., Shannon, R., and Slattery, W. (**1997**). ''Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant,'' J. Speech Lang. Hear. Res. **40**, 1201–1215.
Fourcin, A., Douek, E., Moore, B., Abberton, E., Rosen, S., and Walliker, J. (**1984**). ''Speech pattern element stimulation in electrical hearing,'' Arch. Otolaryngol. **110**, 145–153.
Friesen, L., Shannon, R., Baskent, D., and Wang, X. (**2001**). ''Speech recognition in noise as a function of spectral channels: Comparison of acoustic hearing and cochlear implants,'' J. Acoust. Soc. Am. **110**, 1150–1163.
Fu, Q-J, and Shannon, R. V. (**1999**). ''Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing,'' J. Acoust. Soc. Am. **105**, 1889–1900.
Fu, Q.-J., and Zeng, F.-G. (**2000**). ''Identification of temporal envelope cues in Chinese tone recognition,'' Asia Pac. J. Speech, Lang. Hearing **5**, 45–57.
Fu, Q.-J. (**2001**). Personal communication.
Fu, Q.-J., Zeng, F.-G., Shannon, R. V., and Soli, S. D. (**1998**). ''Importance of tonal envelope cues in Chinese speech recognition,'' J. Acoust. Soc. Am. **104**, 505–510.
Geurts, 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.
Grant, K., Summers, V., and Leek, M. (**1998**). ''Modulation rate detection and discrimination by normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **104**, 1051–1060.
Greenwood, D. D. (**1990**). ''A cochlear frequency-position function for several species—29 years later,'' J. Acoust. Soc. Am. **87**, 2592–2605.
Hanna, T. (**1992**). ''Discrimination and identification of modulation rate using a noise carrier,'' J. Acoust. Soc. Am. **91**, 2122–2128.
Highnam, C., and Morris, V. (**1987**). ''Linguistic stress judgments of language learning disabled students,'' J. Commun. Disord. **20**, 93–103.
Jusczyk, P. (**1997**). *The Discovery of Spoken Language* (MIT, Cambridge, MA).
Lieberman, P., and Michaels, S. B. (**1962**). ''Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech,'' J. Acoust. Soc. Am. **34**, 922–927.
McDermott, H. J., and McKay, C. M. (**1997**). ''Musical perception with electrical stimulation of the cochlea,'' J. Acoust. Soc. Am. **101**, 1622–1631.
McKay, C. M., McDermott, H. J., and Carlyon, R. P. (**2000**). ''Place and temporal cues in pitch perception: are they truly independent?'' ARLO **1**, 25–30.
McKay, C. M., McDermott, H. J., and Clark, G. M. (**1994**). ''Pitch percepts associated with amplitude-modulated current pulse trains by cochlear implantees,'' J. Acoust. Soc. Am. **96**, 2664–2673.
Nooteboom, S. (**1997**). ''The prosody of speech: Melody and rhythm,'' in *The Handbook of Phonetic Sciences*, edited by W. J. Hardcastle and J. Laver (Blackwell, Oxford), pp. 640–673.
Pollack, I. (**1969**). ''Periodicity pitch for white noise—fact or artifact,'' J. Acoust. Soc. Am. **45**, 237–238.
Rosen, S., Faulkner, A., and Wilkinson, L. (**1999**). ''Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants,'' J. Acoust. Soc. Am. **106**, 3629–3636.
Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). ''Speech recognition with primarily temporal cues,'' Science **270**, 303–304.
Shannon, R. V., Zeng, F.-G., and Wygonski, J. (**1998**). ''Speech recognition with altered spectral distribution of envelope cues,'' J. Acoust. Soc. Am. **104**, 2467–2476.
Tong, Y. C., Blamey, P. J., Dowell, R. C., and Clark, G. M. (**1983**). ''Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant,'' J. Acoust. Soc. Am. **74**, 73–80.
Wells, B., Peppe, S., and Vance, M. (**1995**). ''Linguistic assessment of prosody,'' in *Linguistics in Clinical Practice*, edited by K. Grundy (Whurr, London), pp. 234–265.
Whalen, D. H. and Xu, Y. (**1992**). ''Information for Mandarin tones in the amplitude contour and in brief segments,'' Phonetica **49**, 25–47.
Wilson, B., Finley, C., Lawson, D., Wolford, R., Eddington, D., and Rabinowitz, W. (**1991**). ''Better speech recognition with cochlear implants,'' Nature (London) **352**, 236–238.
Wilson, B. (**1997**). ''The future of cochlear implants,'' Br. J. Audiol. **31**, 205–225.
Wilson, B., Zerbi, M., Finley, C., Lawson, D., and van den Honert, C. (**1997**). Eighth Quarterly Progress Report, 1 May through 31 July 1997. NIH Project N01-DC-5-2103: Speech Processors for Auditory Prostheses: Research Triangle Institute.

Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility

Jean C. Krause[a] and Louis D. Braida
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 16 February 2002; accepted for publication 31 July 2002)

Sentences spoken ''clearly'' (and slowly) are significantly more intelligible than those spoken ''conversationally'' for hearing-impaired listeners in a variety of backgrounds [Picheny, Durlach, and Braida, J. Speech Hear. Res. **28**, 96–103 (1985); Uchanski *et al.*, J. Speech Hear. Res. **39**, 494–509 (1996); Payton, Uchanski, and Braida, J. Acoust. Soc. Am. **95**, 1581–1592 (1994)]. However, it is unknown whether slower speaking rates are necessary for highly intelligible speech or whether an alternative form of clear speech exists at faster (i.e., normal) rates. To investigate this question, talkers with significant public speaking experience were asked to produce clear and conversational speech at slow, normal, and quick rates. A method for eliciting clear speech was introduced that ensured the clearest possible speech was obtained at each of these speaking rates. To probe for other highly intelligible speaking modes, talkers also recorded sentences in two other speaking modes: soft and loud. Intelligibility tests indicated that clear speech was the only speaking mode that provided a consistent intelligibility advantage over conversational speech. Moreover, the advantage of clear speech was extended to faster speaking rates than previously reported. These results suggest that clear speech has some inherent acoustic properties that contribute to its high intelligibility without altering rate. Identifying these acoustic properties could lead to improved signal-processing schemes for hearing aids. [DOI: 10.1121/1.1509432]

PACS numbers: 43.71.Bp, 43.71.Gv, 43.71.Ky [CWT]

I. INTRODUCTION

This report is concerned with clear speech, a speaking style that many talkers adopt in order to be understood more easily in difficult communication situations. Previous studies have demonstrated that this altered speaking style is significantly more intelligible (roughly 17 percentage points) than conversational speech for hearing-impaired listeners in a quiet background as well as for both normal-hearing and hearing-impaired listeners in noise and reverberation backgrounds (Picheny *et al.*, 1985; Uchanski *et al.*, 1996; Payton *et al.*, 1994). Furthermore, the intelligibility advantage is independent of listener, presentation level, and frequency-gain characteristic (Picheny *et al.*, 1985). These results suggest that signal-processing schemes that convert conversational speech to a sufficiently close approximation of clear speech could improve speech intelligibility for hearing aid users in many situations.

Before such signal-processing schemes can be developed, however, it is first necessary to determine the extent to which a reduction in speaking rate is responsible for the intelligibility improvement provided by clear speech, since the typical speaking rate for clear speech (100 words per minute) is roughly half that of conversational speech (Picheny, Durlach, and Braida, 1986). This question is particularly important for real-time hearing aid applications, since audio and visual signals must remain synchronized for maximum benefit to the listener. Although the role of speaking rate in highly intelligible speech is not well understood, some linguists hypothesize that clarity is independent of speaking rate (Zwicky, 1972). If so, it should be possible to obtain clear speech and conversational speech at the same speaking rate. Therefore, this paper investigates: (1) whether alternative forms of clear speech can exist at normal speaking rates (or faster), and (2) whether clear speech (spoken slowly) has an intelligibility advantage over typical slow speech, without emphasis on clarity. In either case, it would follow that clear speech has some inherent acoustic properties, independent of rate, that account for its higher intelligibility. Furthermore, the subsequent task of identifying those properties would be greatly simplified, since clear and conversational speech at the same rate could be compared directly.

Most previous attempts to achieve clear speech at normal rates have focused on signal-processing techniques. For example, two studies investigated straightforward time-scale manipulations of clear and conversational speech, compressing clear speech to normal conversational speaking rates and expanding conversational speech to clear speaking rates. In the first of these studies (Picheny, Durlach, and Braida, 1989), the time scale of sentences was altered uniformly, and in a subsequent study (Uchanski *et al.*, 1996), a nonuniform time-scaling method was applied, altering phonetic segments within sentences to reflect the segmental-level durational differences previously measured between clear and conversational speech. Neither time-scaling procedure produced clear speech at normal speaking rates that was more intelligible than unprocessed conversational speech, though nonuniform time scaling was generally less harmful to intelligibility than uniform time scaling. Other work has examined the role of pauses in highly intelligible speech, because the reduced speaking rate found in clear speech is partly a result of more frequent and longer pauses (in conjunction with lengthened

[a]Electronic mail: jeanie@mit.edu

speech sounds) (Picheny *et al.*, 1986). One such study (Uchanski *et al.*, 1996) showed that key words excised from clear and conversational sentences have nearly the same intelligibility as the same words in sentence context, suggesting that differences in pause structure do not necessarily account for differences in intelligibility. This result is supported by another study (Choi, 1987) which found that artificial manipulations of pause structure do not substantially alter the intelligibility of clear and conversational speech; that is, adding pauses to conversational speech did not improve its intelligibility, and deleting pauses from clear speech did not decrease its intelligibility. Such manipulations did increase the effective speaking rate for clear speech, but not nearly enough to achieve normal speaking rates.

In addition to attempts to obtain clear speech at normal speaking rates through signal-processing techniques, one preliminary experiment (Uchanski *et al.*, 1996) sought to elicit clear speech at normal speaking rates naturally. In that experiment, a professional talker attempted to produce clear speech at a variety of rates. Results of intelligibility tests suggested that the talker could not improve his intelligibility without slowing down. However, only one talker was examined, and talkers vary considerably in their ability to produce highly intelligible speech (Bradlow, Toretta, and Pisoni, 1996). Therefore, more work in this area must be completed before any conclusions regarding the existence of naturally produced clear speech at normal speaking rates are justified.

In order to increase the chances of finding at least one talker in this study who could produce a form of clear speech at normal speaking rates, much attention was given to screening talkers, and a method was developed to elicit the clearest possible speech from talkers at a given speaking rate by providing them with training and feedback on intelligibility. In addition, the possibility was also investigated that speaking modes other than clear speech could exist with comparable or even greater intelligibility advantages relative to conversational speech. Some naturally occurring speaking modes differ acoustically from conversational speech [e.g., mothers addressing infants produce more extreme vowels (Kuhl *et al.*, 1997)], but without intelligibility measurements, it is unknown whether these speaking modes also differ significantly from conversational speech in intelligibility. In this study, the intelligibility of loud and soft speech was evaluated relative to conversational speech presented at the same intensity. These two speaking modes were chosen because they occur frequently in natural speech, are relatively easy to elicit from talkers, and have acoustic effects on the speech spectrum that are independent of intensity (Licklider, Hawley, and Walking, 1955). Finally, this study also investigated whether clear speech (spoken slowly) is more intelligible than typical slow speech, without emphasis on clarity. Although other studies have examined acoustic properties of slow speech (Crystal and House, 1982, 1988; Han, 1966), none has considered intelligibility. Such an analysis is straightforward to conduct and lends insight into whether these studies are useful for understanding clear speech.

In this paper, an objective method is introduced for training talkers to produce the clearest possible speech at a variety of speaking rates. Methods are also described for eliciting soft and loud speech. Intelligibility results are reported for all speaking modes, loud and soft as well as clear and conversational speech at slow, normal, and quick rates. The goal of the intelligibility tests was to determine whether normal-hearing listeners with simulated hearing losses, achieved by additive noise, could derive intelligibility benefits (relative to conversational speech) from one or more of the speaking modes elicited at a given speaking rate. If so, further examination of such highly intelligible speaking modes could lend insight into possible signal-processing approaches for hearing aids with the potential to improve speech clarity as well as audibility.

II. SPEECH ELICITATION METHODS

In order to improve the chances of obtaining clear speech at normal speaking rates (i.e., rates comparable to conversational speech) naturally, much attention was given to talker selection and training. Talkers were recruited from the New England area, but only talkers with significant public speaking experience (e.g., students or professionals in television or radio broadcasting, public speaking, or other communications disciplines) were considered, because their experience increased the likelihood that they could respond to training in relatively short amounts of time. Since the training was fairly intensive, however, it was not feasible to train all potential talkers. Therefore, talkers who responded with at least 2 years of public speaking experience were asked to participate in a preliminary screening to evaluate their intelligibility. A description of the 15 talkers who participated in the screening and their speaking experiences is summarized in Table I. Based on the results of the screening, the five participants with the highest potential for producing clear speech at a variety of speaking rates were then selected for training.

A. Screening of talkers

In order to obtain a form of clear speech from each talker with minimal training, the talkers were familiarized with the characteristics of previously obtained clear speech. The talkers listened to samples of both conversational and clear speech materials recorded for an earlier study of clear speech (Picheny *et al.*, 1985), and differences between the two speaking modes were discussed. The talkers were asked to mimic the clear speech that had been presented, and they were given feedback by the experimenter on both rate and clarity. The goal of obtaining clear speech at normal speaking rates was explained, but each talker was instructed not to increase speaking rate at the expense of clarity. Each talker was then given 1 hour to practice producing clear speech. At the conclusion of the practice period, each talker was asked to record a unique set of 100 sentences, 50 spoken clearly and 50 spoken conversationally.

1. Recording procedure and sentence materials

All recording sessions took place with the talker seated in a sound-treated room. A Sennheiser MD 421 cardioid microphone was positioned approximately 6 in. in front of the talker's mouth. The roll-off filter on the microphone was adjusted to the speech setting, and the microphone output

TABLE I. The following data were used to evaluate talkers during the preliminary intelligibility screening: (1) speaking experience; (2) intelligibility (I), based on % correct key-word scores; and (3) speaking rate (r) in words per minute, excluding pauses. Conversational and clear data are differentiated with subscripts. The change in intelligibility as a function of rate is represented by the slope of a line that would connect the two data points, calculated by $m=-(I_{\text{clear}}-I_{\text{conv}})/(r_{\text{clear}}-r_{\text{conv}})$. Individual results are listed here for the five talkers who were ultimately selected to participate in formal training (T1–T5). Results for the remaining talkers who participated in the screening (T6–T15) are reported in aggregate.

Talker	Sex	Speaking experience	Yrs	I_{conv}	r_{conv}	I_{clear}	r_{clear}	m
T1	F	College television, radio, public speaking	5	42	307	48	169	0.04
T2	F	Professional speaker	5	42	213	70	97	0.24
T3	F	Broadcasting student	2	25	315	64	61	0.15
T4	F	Debate team	6	52	175	75	57	0.20
T5	M	Debate team	7	48	164	79	90	0.42
T6–T15	5 M, 5 F	Varied	2–5	20–45	160–285	34–64	55–198	0.10–0.34

was amplified using a Symetrix SX202 dual microphone preamplifier. The speech was recorded at a 48-kHz sampling rate, using a SONY 59ES digital audio tape deck. The files were then downsampled to 20 kHz and normalized for rms level. The sentence materials recorded were obtained from the corpus of nonsense sentences described by Picheny *et al.* (1985). These sentences have five to eight words each and provide no semantic context that could aid listeners in identifying key words, which are defined as all nouns, verbs, and adjectives (e.g., ''The right cane could guard an edge.'').

2. Listeners and testing conditions

Two normal-hearing listeners (one male, one female) were obtained from MIT and the surrounding community. They were native speakers of English who possessed at least a high school education and had no prior experience with the task. Their hearing thresholds were no greater than 20 dB HL at frequencies between 250 and 4000 Hz.

Each talker's speech was presented in a background of speech-shaped noise (Nilsson, Soli, and Sullivan, 1994) to listeners monaurally over TDH-39 headphones. The stimuli were stereo signals with speech on one channel and speech-shaped noise of the same rms level on the other channel. The speech was attenuated by 4 dB and added to the speech-shaped noise, and the resulting signal (SNR$=-4$ dB) was presented to the listeners from a PC through a DAL card. Listeners were given the opportunity to choose a comfortable listening level and to select which ear would receive the stimuli. They were also encouraged to switch the stimulus to the other ear when fatigued. Listeners responded by writing their answers on paper. They were given as much time as needed to respond, but were presented each sentence only once. Intelligibility scores for key words (nouns, verbs, and adjectives) were determined using the scoring rules described by Picheny *et al.* (1985). Under these testing conditions, normal-hearing listeners have been shown to receive benefits from clear speech that are consistent with hearing-impaired listeners in quiet (Payton *et al.*, 1994).

3. Evaluation and selection of talkers

Speaking rates in words per minute (wpm) and intelligibility scores for each talker, averaged across listener, are summarized in Table I. Since the goal of the screening was to identify talkers who could not only produce clear speech but also demonstrate a high potential for producing clear speech at normal speaking rates, pauses (silent periods of 10 ms or longer) were excluded from the speaking rate calculation. For clear speech, this procedure provided a rough estimate of the minimum speaking rate that could be achieved by training talkers without sacrificing intelligibility, given that deleting pauses from clear speech does not reduce its intelligibility (Choi, 1987). It also provided an indication of articulation rates, which facilitated analysis of talkers' abilities to produce clear speech at normal rates.

When speaking clearly, all talkers achieved some improvement in intelligibility over speaking conversationally, but none achieved this improvement without slowing down. Since the screening did not provide talkers with training, this result was not unexpected. However, the relationship between intelligibility and speaking rate was considered an important criterion for screening talkers' potential to achieve clear speech at normal rates. In order to quantify this relationship, an equation for each talker was derived for the line passing through the two data points (conversational and clear) achieved by that talker during the screening, with intelligibility as a function of speaking rate. The equation was of the form $I=a-mr$, where I represents intelligibility, r represents speaking rate excluding pauses, and a and m are positive constants. The value of m satisfying this equation for each talker is also summarized in Table I.

Since it was unknown what talker characteristics were most likely to be associated with talkers who were capable of producing clear speech at normal rates, talkers with different characteristics were selected in order to improve the chances of finding at least one talker who could produce clear speech at normal speaking rates. The data for the five selected talkers are listed in Table I. T1 was selected for her ability to speak at a higher rate than most other talkers, both in conversational (307 wpm) and clear (169 wpm) modes. In addition, she exhibited an unusually low value of the slope parameter m. It was hoped this value of m could be increased with training. Talkers T2 and T4 were selected because their overall intelligibility in both modes was higher than most of the other talkers at similar speaking rates. Two talkers with this characteristic were selected because it was considered likely to be associated with an ability to speak clearly at

normal rates. T3 was selected because she had the greatest increase in intelligibility between conversational and clear speech. She also demonstrated the ability to change her speaking rate significantly, from 61 wpm in clear mode to 315 wpm in conversational mode. T5 was selected because his clear speech had the overall highest intelligibility at 79 percent as well as the highest value of *m*. These five talkers (four females, one male) participated in the formal training sessions described below.

B. Training and recording procedure

After participating in intensive training, talkers recorded clear and conversational speech from talkers at slow, normal, and quick speaking rates, resulting in six speaking styles (referred to by mode/rate): clear/slow, conv/slow, clear/normal, conv/normal, clear/quick, conv/quick. The novel training procedure, described below, included quantitative feedback on both speaking rate and intelligibility, ensuring that the clearest possible speech for a given speaking rate was obtained from talkers. Whether this speech could provide an intelligibility advantage over conversational speech of the same rate was evaluated with intelligibility tests described in the next section.

1. Regulating speaking rate

Speaking rates were specific to individual talkers and were regulated at the sentence level in order to make the task as natural for talkers as possible. Moreover, regulating sentence rate rather than word or syllable rate allowed the talker freedom to determine the duration of individual speech segments and words, thus maximizing the chance of obtaining natural clear speech at normal and quick rates. After an initial determination of speaking rates, talkers were required to retain their individual slow, normal, and quick rates throughout all training and recording sessions. The appropriate speaking rate was communicated to the talker by presenting over headphones the output of a metronome that was set to the desired sentence rate. Talkers could take as much time as desired between sentences, but were required to fit each sentence between two consecutive metronome clicks. Most talkers adjusted easily to monitoring their speaking rate in this way and did not appear to find the procedure difficult or unnatural.

Individual speaking rates were initially determined by the average rates achieved by each talker when asked to do the following: (1) slow—produce 100 clear sentences, with no constraints on rate, at the conclusion of training; (2) normal—read 200 sentences at a rate appropriate for normal conversation; (3) quick—read 50 sentences as rapidly as possible. Clear speech was used to establish the slow rate for two reasons. First, imposing no rate constraints on the production of clear speech was similar to methods used for eliciting clear (slow) speech in the screening and in previous studies (Picheny *et al.*, 1985; Uchanski *et al.*, 1996) and ensured that highly intelligible speech would be elicited. Second, this method allowed for elicitation of conv/slow speech at rates comparable to clear (i.e., clear/slow) speech, which allowed for a direct comparison of the intelligibility of clear and conversational modes at the same rate. Although speaking rates were not uniform across talkers, Fig. 1 shows that there was very little overlap in the distributions of slow, normal, and quick rates across talkers. Only T2, who varied her rate relatively little, chose a quick rate (193 wpm) that overlapped with the remaining four talkers' normal rates, which averaged 187 wpm and ranged from 171 wpm (T5) to 198 wpm (T3).

2. Regulating intelligibility

The procedure for regulating intelligibility required the talker to repeat a sentence with increased emphasis on articulation until it was perceived correctly by a listener. This procedure was derived from an existing method for eliciting clear speech with syllables (Chen, 1980) in which a talker repeated a syllable until the listener perceived it correctly in the presence of masking noise. While such a method could be used with nonsense sentences, its disadvantage is that repetition of sentences increases their intelligibility (Uchanski *et al.*, 1996). To avoid the intelligibility benefit of repetition, four normal-hearing listeners were employed successively to provide the talker with feedback on intelligibility. The talker's speech was distorted by multiplicative noise (Schroeder, 1968) and presented to each of the listeners, in turn, monaurally over headphones. Multiplicative noise was used because it maintains a constant SNR, thus preventing the talkers from increasing intelligibility simply by speaking more loudly. At the beginning of each session, the SNR was set to 0 dB, and it was decreased in increments of 0.2 dB until the listeners received on average no more than one key word correctly from the talker's first utterance of the sentence. Throughout the experiment, the SNR was decreased in increments of 0.2 dB as the talker's intelligibility improved.

The talker and listeners were seated separately and did not have visual contact. Each listener could hear the talker only when directly addressed and presented with a sentence. The designated listener then responded verbally with the sentence heard. The talker and the experimenter could both hear the listener's response, and they could communicate freely with each other throughout the session. The experimenter provided instruction, reminding the talker to adhere to the timing cues provided by the metronome (for clear/normal and clear/quick modes) and pointing out patterns of mistakes among the listeners. The experimenter also served as a judge of the listener's responses and decided whether or not the talker should repeat a sentence. The listener's response was regarded as correct if more than half of the key words were correctly identified. If the response was incorrect, the talker repeated the sentence to the next listener; if the response was correct, the talker presented a new sentence to the next listener. The four listeners were not given feedback on whether or not the response was correct. A sentence was not repeated additional times after it had been presented to all four listeners. If three sentences in a row were presented to all four listeners without a correct response, the SNR was increased by 0.2 dB.

This procedure was used for both training and recording. During training sessions, talkers were encouraged to experiment with different speaking strategies and allowed to practice as much as desired. After speaking strategies had been

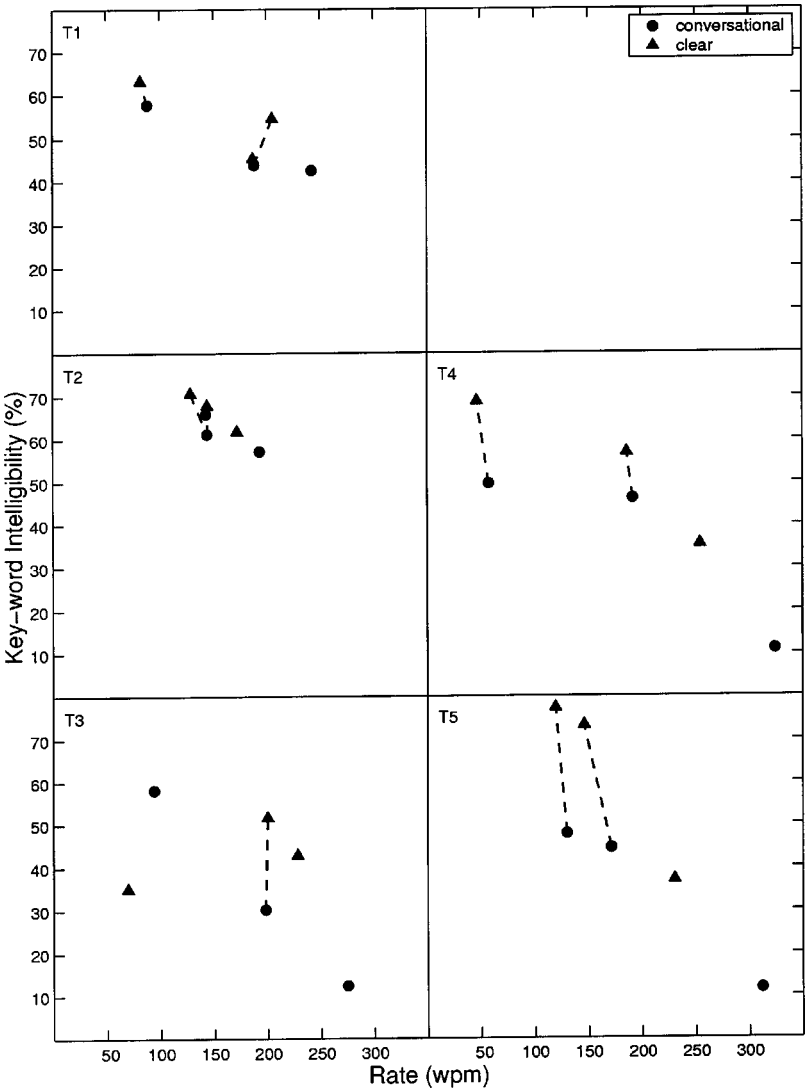


FIG. 1. Average key-word scores for each talker, averaged over listener, versus speaking rate. Dashed lines represent instances where intelligibility was improved without a significant change (no more than 25 wpm) in speaking rate. Two dashed lines emanate from T2's conv/normal data point, because both clear/slow and clear/normal speech provided an intelligibility advantage without a significant change in rate.

explored, training of each condition (clear/slow, clear/normal, clear/quick) was considered complete when the SNR had reached a constant value (within ± 0.2 dB). After a short break, the condition was then recorded.

3. Eliciting other speaking modes

Soft/normal and loud/normal speech were elicited from talkers with the aid of a Realistic digital sound-level meter, located approximately 2 1/2 ft. from the talker's mouth and set to measure the maximum sound pressure level (A-weighted, FAST-acting average: updated every 0.25 s) in a sentence. The talker was instructed to read ten sentences in a conversational manner, and the largest and smallest sound levels were noted. For soft/normal speech, each talker was instructed to speak at sound levels at least 15 dB below the largest level measured for that talker during conversational speech. The level on the meter was reported to the talker after each sentence was read, and the sentence was repeated if necessary. Loud/normal speech was elicited in a similar manner, except the talker was instructed to exceed the smallest level measured during conversational speech by at least 15 dB. It should be noted that although these modes were elicited at various intensities, all sentences were normalized to the same rms level for the purposes of intelligibility tests. Thus, the question addressed in the intelligibility tests was whether these speaking modes provide an intelligibility advantage over conversational speech of the same intensity.

4. Recording sessions

Each talker recorded 700 nonsense sentences over four recording sessions, 2 to 3 h in length. The 600 sentences discussed here consisted of six unique sentence lists, each containing 50 sentences. Every sentence list was recorded in two speaking modes: L1—soft/normal and conv/normal; L2—loud/normal and conv/normal; L3—clear/normal and conv/normal; L4—clear/slow and conv/normal; L5—clear/quick and conv/quick; L6—clear/slow and conv/slow). One mode was the test condition, and the other was always conversational, to establish baseline intelligibility scores. In most cases, the two speaking modes were recorded at the same speaking rate in order to facilitate intelligibility comparisons without speaking rate as a factor. One list (L4), however, was recorded at two different rates: once in conversational mode at the normal rate and once in clear mode at the slow rate. The recording procedure was identical to that used for the screening, except that (1) the amplifier output

was recorded directly to a PC disk, using a DAL card with a 20-kHz sampling rate, and (2) pauses were excluded for normalization purposes in order to ensure that the rms level reflected the level of the speech only, since pause durations varied greatly across the three speaking rates. Sentences were checked for errors, and in a few cases, mispronounced words were noted so that responses could be graded accordingly during intelligibility tests.

III. INTELLIGIBILITY TESTS

Eight normal-hearing listeners (four males, four females; aged 18 to 29 years) participated in the final intelligibility tests. These listeners met the same requirements for normal-hearing listeners as those who participated in the screening. None had participated in the screening nor had any prior familiarity with the task. The eight listeners were divided into two separate testing groups.

Listeners participated in 16 2-h sessions over the course of approximately 8 weeks. In each session, listeners were tested on 4–5 sentence lists and were given a 5-min break after the presentation of each list and were also encouraged to rest briefly as necessary. In addition, a 10-min break was given near the halfway point of each session. The stimuli were prepared and presented in the same manner as in the screening, except that the SNR was set to −1.8 dB. This SNR was chosen to avoid ceiling and floor effects on all conditions, which varied considerably in difficulty. It was determined through informal listening tests that were administered to normal-hearing volunteers from the laboratory.

Although the groups met at different times, both groups heard the same lists in the same order. It was not necessary to randomize the order of presentation of the lists, because learning effects over the course of an experiment are minimal with these speech materials (Picheny *et al.*, 1985). However, because each list was presented in two different modes in this experiment (conversational, for reference, and a test condition), learning effects within a list were a possibility. In order to minimize these learning effects, the conversational condition was presented prior to the test condition in roughly half of the cases and after the test condition in the remaining half. In addition, the amount of time between the first and second presentations of a each list was always at least 2 weeks.

IV. RESULTS

Intelligibility scores, averaged across talkers and listeners, and speaking rates were calculated as in the screening. To simplify discussion, the data discussed here are also averaged across all presentations of a given condition. Although the number of lists presented in each condition varied, the lists have been shown to be of equal difficulty (Rosengard, 2000). Therefore, this averaging method should provide the best estimates of the intelligibility of various conditions.

The clear/slow speaking style was most intelligible at 63 percent key words correct, followed in order of decreasing intelligibility by clear/normal (59%), loud/normal (53%), conv/slow (51%), clear/quick (46%), conv/normal (45%), conv/quick (27%), soft/normal (26%) modes. The 18-percentage-point advantage for clear/slow (63%) relative to conv/normal speech (45%) is consistent with previous studies (Picheny *et al.*, 1985; Uchanski *et al.*, 1996). Moreover, a 14-point advantage was obtained for clear/normal (59%) relative to conv/normal speech, thus extending the benefit of clear speech to normal speaking rates. In addition, this advantage was comparable to the 12-point advantage obtained for clear/slow (63%) relative to conv/slow (51%) speech. Although a 19-point advantage was also obtained for clear/quick (46%) relative to conv/quick speech (27%), this finding does not extend the benefit of clear speech to quick rates, because the speaking rate for clear/quick speech (218 wpm) was significantly slower than the speaking rate for conv/quick speech (269 wpm). Of the speaking modes other than clear speech that were evaluated, soft/normal speech was less intelligible on average than conv/normal speech. On average, loud/normal speech (53%) was more intelligible than conv/normal speech (45%), but the advantage was less for loud/normal speech (8 points) than for clear/normal speech (14 points). Moreover, not all talkers achieved an intelligibility advantage with loud/normal speech. Thus, at normal rates, none of the alternative speaking modes tested provided as large or as consistent of an intelligibility advantage over conversational speech as clear speech.

An analysis of variance was performed on the key-word scores for the data specific to clear and conversational speech (conv/slow, conv/normal, conv/quick, clear/slow, clear/normal, clear/quick), after an arcsine transformation (arcsin $\sqrt{I_j/100}$) to equalize the variances. The analysis of variance used a standard factorial model (nonrepeated measures) and analyzed four factors, assuming one random factor, listener, and three fixed factors: talker, mode, and rate. Although many factors and interactions had statistically significant F-ratios ($p<0.01$), only five of them accounted for substantial portions of the variance: rate (20%), talker (19%), mode (13%), listener (7%), and rate×talker (10%). For the purposes of this study, it is most important to note that a substantial portion of the variance is accounted for by the mode factor alone, and that a relatively small portion of the total variance (5% or less) is accounted for by each of the interactions with mode that were significant. In fact, with the exception of mode×talker and mode×rate×talker, each statistically significant interaction with mode accounted for less than 2% of the variance, including mode×listener and mode ×rate as well as some higher-order interaction terms. Therefore, to a first-order approximation, it can be concluded that the intelligibility benefit obtained by speaking clearly (mode) is statistically significant, and the size of the benefit is largely independent of listener, talker, and speaking rate. In other words, a comparable intelligibility benefit is obtained for a majority of talkers, independent of their overall intelligibility, for a majority of listeners, independent of their overall skill at the listening task, and for a majority of speaking rates (at least slow and normal), independent of the overall decline in intelligibility that comes with increased rate.

Of the interactions with mode that were statistically significant, the two terms accounting for the most variance were mode×talker (5%) and mode×rate×talker (5%). These in-

TABLE II. Slopes $m1$ and $m2$ for each talker in this study, where $m1=(I_{\text{normal}}-I_{\text{slow}})/(r_{\text{normal}}-r_{\text{slow}})$ and $m2=(I_{\text{quick}}-I_{\text{normal}})/(r_{\text{quick}}-r_{\text{normal}})$. I represents intelligibility (%), r represents speaking rate (wpm), excluding pauses, and subscripts indicate nominal speaking rate. Asterisks indicate slopes in this study that could not be calculated because of data points that were eliminated due to the failure of the talker to achieve the desired rate or intelligibility. Comparison of average slopes (calculated from intelligibility and rate values averaged across talker) in this study with those obtained for previous attempts to elicit clear speech at normal rates shows that the present study's speech elicitation methods (talker selection and training) extended the advantage of clear speech to normal rates by increasing the value of $m1_{\text{clear}}$ to that of $m1_{\text{conv}}$.

Talker/Study	Listener(s)	Scaling	$m1_{\text{conv}}$	$m2_{\text{conv}}$	$m1_{\text{clear}}$	$m2_{\text{clear}}$
T1	Normal	Natural+training	−0.14	−0.02	−0.07	*
T2	Normal	Natural+training	*	−0.08	−0.18	−0.21
T3	Normal	Natural+training	−0.27	−0.23	*	−0.31
T4	Normal	Natural+training	−0.03	−0.27	−0.09	−0.32
T5	Normal	Natural+training	−0.08	−0.23	−0.15	−0.43
AVG (present study)	Normal	Natural+training	−0.10	−0.20	−0.10	−0.33
Picheny *et al.* (1989)	Impaired	Uniform	⋯	⋯	−0.30	⋯
Uchanski *et al.* (1996)	Impaired	Nonuniform	⋯	⋯	−0.25	⋯
Uchanski *et al.* (1996)	Normal	Natural	⋯	⋯	−0.25	−0.29
Uchanski *et al.* (1996)	Impaired	Natural	⋯	⋯	−0.21	−0.28

teractions are shown in Fig. 1. For both T4 and T5, conv/slow speech was comparable in intelligibility to conv/normal speech, and the clear speaking mode at these speaking rates was much more intelligible than the conversational mode (average key-word scores were 18 points higher for T4 and 28 points higher for T5). Trends for the other three talkers are less clear. T3 failed to produce highly intelligible speech at the slow speaking rate, although her clear/normal and clear/quick styles were more intelligible than her conversational speech at similar speaking rates. T2 varied her speaking rate the least of all the talkers, and reported having difficulty adhering to the metronome. Her intelligibility drops off quickly at speaking rates above 150 wpm. T1 reported that she preferred speaking quickly, which may partly explain her achieving a higher key-word score for clear/quick speech than for clear/normal speech.

Instances where talkers achieved an intelligibility benefit (at least 5 percentage points) without a substantial change in speaking rate (no more than 25 wpm) are indicated in Fig. 1 with dashed lines. At normal rates, all talkers met this criterion, producing a form of clear speech that was more intelligible than conv/normal speech produced at nearly the same speaking rate (T1 met this criterion with clear/quick speech, which was produced within 25 wpm of her conv/normal speech). In addition, clear/slow speech was also more intelligible than conv/slow speech for all talkers except T3, according to this criterion. Although all talkers obtained an increase in intelligibility at quick rates by speaking clearly, none of them did so without also speaking more slowly. Additional training may have been necessary, or it may not be possible to produce clear speech at such high rates. Regardless, it should be noted that the clear/quick speech obtained for some of the talkers (particularly, T1 and T3) seems likely to be more intelligible than conversational speech at comparable rates, if such conversational speech had been obtained in this study (based on an interpolation of the conversational speech scores that were obtained for each talker). If so, it appears that the clear speech benefit can also extend to faster than normal speaking rates.

The clear speech benefit achieved by talkers in this study was also analyzed by comparing two slope values across speaking modes: $m1$, the decrease in intelligibility from slow to normal rates, and $m2$, the decrease in intelligibility from normal to quick rates. These slopes, shown in Table II, were calculated for each talker and speaking mode from the data pictured in Fig. 1, with the exception of three data points, which were omitted because they failed to exhibit a change in rate or an improvement in intelligibility: T1's clear/normal data (no intelligibility improvement over conv/normal), T2's conv/slow data (not significantly slower than conv/normal), and T3's clear/slow data (no intelligibility improvement over conv/slow). In the first case, T1's clear/quick data point was substituted for the clear/normal omission, because it provided an intelligibility improvement over conv/normal speech and was within 25 wpm of her normal rate. In addition, average slope values were calculated from mean intelligibility and rate values averaged across talker, for this study as well as for previous attempts to elicit clear speech at normal rates. As shown in Table II, the average slope $m1_{\text{clear}}$ in this study was greater than $m1_{\text{clear}}$ for all previous studies and equal to $m1_{\text{conv}}$ (-0.10), indicating that the clear slope has been flattened (relative to previous studies) such that the size of the intelligibility benefit of clear modes over conversational modes remains roughly equal for slow and normal speaking rates. At faster rates, however, the average slope $m2_{\text{clear}}$ in this study became considerably steeper, dropping off more quickly than $m2_{\text{conv}}$ as well as all $m2_{\text{clear}}$ values obtained from previous studies. The point where these $m2$ lines converge may represent a physical limit on articulation resulting from physiological constraints at very high speaking rates.

V. DISCUSSION

The results of this study show that (1) at a given speaking rate, neither of the additional speaking modes examined provided an intelligibility benefit as large as that of clear speech; and (2) with proper training of talkers, the benefits of clear speech can be extended to faster speaking rates than those previously reported. Specifically, a form of clear

speech was obtained at slow (roughly 100 wpm) and normal (roughly 200 wpm) rates. Because the intelligibility advantage of clear/slow over conv/slow speech was comparable to that of clear/normal over conv/normal for nearly all talkers and listeners, it was also shown that over this range of speaking rates (slow through normal), the relative intelligibility benefit of clear speech is largely independent of rate, talker, and listener.

While the intelligibility advantage of clear speech did not extend to quick rates, as shown by the fact that $m2_{\text{clear}} < m2_{\text{conv}}$ for all talkers, it could still exist at faster than normal rates. Assuming physiological constraints on articulation, the intelligibility of clear speech at very high speaking rates must decrease more rapidly than the intelligibility of conversational speech in order to compensate for its higher intelligibility. Consequently, clear speech cannot maintain an intelligibility advantage above a certain ''cutoff'' speaking rate. Yet, with the limited number of rates that were examined in this study, only a lower bound (normal rates of roughly 200 wpm) on the cutoff rate was established. At least some of the talkers (T1 and T3) appear likely to have exceeded this lower bound, producing clear/quick speech that could be more intelligible than conversational speech at comparable rates, if such conversational speech had been elicited. Moreover, since the training provided in this study increased the cutoff rate beyond those previously reported, it is possible that additional training could increase the cutoff rate even further.

Subjective comments from talkers regarding the training procedure indicated that the listener feedback provided during training was very helpful for developing clear speech. In particular, one talker noted that trends in listener responses raised his awareness of common phoneme confusions. He reported that this information was useful in deciding which phonemes to emphasize. Other talkers expressed interest in listening to speech distorted by multiplicative noise in order to gain information on how to speak more clearly. This request suggests that some talkers believe they have natural strategies for speaking clearly in difficult communication situations. Moreover, these strategies may differ depending on the nature of the distortion.

However, even if different types of clear speech exist, depending on the nature of the distortion, the clear speech obtained in this experiment warrants further study because it provided an intelligibility benefit to listeners with simulated hearing loss, without an accompanying reduction in speaking rate. Acoustical analyses of the differences between clear/slow and conv/slow speech and between clear/normal and conv/normal speech should help identify which characteristics of clear speech that have been reported previously (Picheny *et al.*, 1986; Cutler and Butterfield, 1990, 1991) contribute to its high intelligibility without altering rate. In addition, analyzing the intelligibility of clear speech at several rates between 200 and 300 wpm, elicited with varying amounts of training, should help determine the maximum cutoff rate for achieving a sizable clear speech benefit. Such findings could ultimately lead to the development of signal-processing approaches for hearing aids that convert conversational speech to a close approximation of clear speech over as wide a range of speaking rates as possible.

ACKNOWLEDGMENTS

The authors are grateful to the many participants in this study, particularly the talkers for their dedication and enthusiasm during training. In addition, we thank Rosalie M. Uchanski for many helpful technical discussions. Financial support for this work was provided by a grant from the National Institute on Deafness and Other Communication Disorders (NIH Grant Number 5 R01 DC 00117) and a National Defense Science and Engineering Graduate fellowship from the Office of Naval Research.

Bradlow, A., Toretta, G., and Pisoni, D. (**1996**). ''Intelligibility of normal speech I. Global and fine-grained acoustic-phonetic talker characteristics,'' Speech Commun. **20**, 225–272.

Chen, F. R. (**1980**). ''Acoustic characteristics and intelligibility of clear and conversational speech,'' Master's project, Mass. Inst. Tech., Cambridge, MA.

Choi, S. (**1987**). ''The effect of pauses on the intelligibility of sentences,'' Bachelor's project, Mass. Inst. Tech., Cambridge, MA.

Crystal, T. H., and House, A. S. (**1982**). ''Segmental durations in connected-speech signals: Preliminary results,'' J. Acoust. Soc. Am. **72**, 705–716.

Crystal, T. H., and House, A. S. (**1988**). ''Segmental durations in connected-speech signals: Current results,'' J. Acoust. Soc. Am. **83**, 1553–1573.

Cutler, A., and Butterfield, S. (**1990**). ''Durational cues to word boundaries in clear speech,'' Speech Commun. **9**, 485–495.

Cutler, A., and Butterfield, S. (**1991**). ''Word boundary cues in clear speech: A supplementary report,'' Speech Commun. **10**, 335–353.

Han, M. S. (**1966**). ''Acoustic-phonetic study on speech tempo,'' Stud. Sounds **12**, 70–83.

Kuhl, P., Andruski, J., Chistovich, I., Chistovich, L., Kozhevnikova, E., Ryskina, V., Stolyarova, E., Sundberg, U., and Lacerda, F. (**1997**). ''Cross-language analysis of phonetic units in language addressed to infants,'' Science **277**, 684–686.

Licklider, J., Hawley, M., and Walking, R. (**1955**). ''Influences of variations in speech intensity and other factors upon the speech spectrum,'' J. Acoust. Soc. Am. **27**, 207.

Nilsson, M., Soli, S. D., and Sullivan, J. A. (**1994**). ''Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise,'' J. Acoust. Soc. Am. **95**, 1085–1099.

Payton, K. L., Uchanski, R. M., and Braida, L. D. (**1994**). ''Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing,'' J. Acoust. Soc. Am. **95**, 1581–1592.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1985**). ''Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech,'' J. Speech Hear. Res. **28**, 96–103.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1986**). ''Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech,'' J. Speech Hear. Res. **29**, 434–446.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1989**). ''Speaking clearly for the hard of hearing. III. An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech,'' J. Speech Hear. Res. **32**, 600–603.

Rosengard, P. (**2000**). Personal correspondence *re*: Poster. International Hearing Aid Conference, Lake Tahoe.

Schroeder, M. R. (**1968**). ''Reference signal for signal quality studies,'' J. Acoust. Soc. Am. **44**, 1735–1736.

Uchanski, R. M., Choi, S., Braida, L. D., Reed, C. M., and Durlach, N. I. (**1996**). ''Speaking clearly for the hard of hearing. IV. Further studies of the role of speaking rate,'' J. Speech Hear. Res. **39**, 494–509.

Zwicky, A. M. (**1972**). ''On causal speech,'' *Papers from the Eighth Regional Meeting* (Chicago Linguistic Society, Chicago), pp. 607–615.

Underwater audiogram of a Pacific walrus (*Odobenus rosmarus divergens*) measured with narrow-band frequency-modulated signals

R. A. Kastelein,[a] P. Mosterd, B. van Santen, and M. Hagedoorn
Harderwijk Marine Mammal Park, Strandboulevard-oost 1, 3841 AB Harderwijk, The Netherlands[b]

D. de Haan
Netherlands Institute for Fisheries Research (RIVO-DLO), P.O. Box 68, 1970 AB IJmuiden, The Netherlands

(Received 18 March 2002; revised 7 June 2002; accepted 23 July 2002)

The underwater hearing sensitivity of an 18-year-old male Pacific walrus was measured in a pool by using a go/no-go response paradigm and the up-down staircase method. Auditory sensitivity was measured using narrow-band, frequency-modulated signals (1500 ms duration) with center frequencies ranging from 0.125 to 15 kHz. The resulting underwater audiogram (50% detection thresholds) for this individual walrus shows the typical mammalian U-shape. Maximum sensitivity (67 dB *re* 1 μPa) occurred at 12 kHz. The range of best hearing (10 dB from the maximum sensitivity) was from 1 to 12 kHz. Sensitivity fell gradually below 1 kHz and dropped off sharply above 12 kHz. The animal showed a peculiar insensitivity for 2 kHz signals. His much higher sensitivity for 1.5- and 3-kHz signals indicated that this is a narrow-band phenomenon. Walrus hearing is relatively sensitive to low frequency sound, thus the species is likely to be susceptible to anthropogenic noise. The thresholds found during a small test with four frequencies with signal durations of 300 ms did not differ significantly from those obtained with signal durations of 1500 ms. [DOI: 10.1121/1.1508783]

PACS numbers: 43.80.Lb, 43.80.Nd [WA]

I. INTRODUCTION

The walrus is a large pinniped species (adults weigh ca. 1800 kg) found in the arctic and sub-arctic waters of the Northern Hemisphere (Fay *et al.*, 1986). It leads an amphibious life and spends much time both resting on land and in the water during activities such as migration, foraging and courtship (Kastelein, 2001).

Walruses produce both aerial and underwater vocalizations. Their large variety of high amplitude aerial vocalizations are described as ''bellows'' or ''barks,'' ''roars,'' guttural calls and grunts, and seem to have a social function (Ray and Watkins, 1975; Kibal'chic and Lisitsina, 1979; Miller and Boness, 1983; Miller, 1985; Kastelein *et al.*, 1995). Both in the wild and in captivity walruses produce loud whistles, especially, but not exclusively, during the mating season (Ray and Watkins, 1975; Verboom and Kastelein, 1995). Walruses also produce a large variety of underwater vocalizations described as ''bell-like,'' ''clicks,'' ''knocks,'' ''rasps,'' and sounding ''like castanets.'' Some are so loud that people standing on the ice can hear them in the air when the walruses swim underneath. They are assumed to be used mainly by males for advertisement, underwater territory (lek) establishment, and courtship. Patterns of underwater vocalization cycles are often associated with dive cycles (Fay, 1960; Fay *et al.*, 1984; Schevill *et al.*, 1966; Ray and Watkins, 1975; Stirling *et al.*, 1983, 1987; Nowicki *et al.*, 1997). Because of the amphibious nature of walrus vocalizations, amphibious hearing may play an important role in walrus ecology.

Within the geographic range of the Pacific walrus (*Odobenus rosmarus divergens*) there has been a recent increase in anthropogenic noise both in air and under water. Anecdotal information suggests that walruses are disturbed by anthropogenic noise (Fay and Kelly, 1982; Fay *et al.*, 1986; Born and Knutsen, 1990). Such predominantly low-frequency noise may influence the ecology of walruses by deterring them from their feeding, breeding or resting grounds (Born *et al.*, 1995). Anthropogenic noise may also mask communication between conspecifics necessary for successful breeding and rearing of young. It may also interfere with the detection of ambient sounds (including conspecific signals) that might be useful in spatial orientation. Prolonged disturbance may ultimately contribute to a decline in body condition and lead to disease or reduced reproductive output. Due to disturbance, mothers and calves may become separated, making calves vulnerable to predation by polar bears, *Ursus maritimus* (Fay *et al.*, 1984). Calves may also be killed if noise disturbances result in stampedes (Fay *et al.*, 1990). Either very loud or moderately loud and chronic anthropogenic noise exposure may induce temporary hearing threshold shifts (TTSs) or permanent hearing threshold shifts (PTSs) in walruses. Kastak *et al.* (1999) observed TTSs in other pinniped species.

To determine the importance of sound to the Pacific walrus, information is needed on its hearing sensitivity both in air and underwater. Aerial hearing sensitivity was tested by Kastelein *et al.* (1996b). However, information on the under-

[a]Electronic mail: r.kastelein@dolfinarium.nl
[b]Part of Grévin et Compagnie, France.

TABLE I. The mean 50% detection thresholds of an 18-year-old male walrus for 16 signals, mean session threshold range, number of sessions, number of reversals, and Prestimulus response rate over all (signal present and signal absent) trials.

Center frequency (kHz)	Signal type	Frequency modulation range (kHz)	Transducer	Mean 50% detection threshold (dB *re* 1 μPa)	Mean session threshold range (dB*re* 1 μPa)	No. of sessions	Total no. of reversals	Prestimulus response rate (%)[c]
0.125	Pure tone	···	DRS-12	106	101–111	10	48	9
0.2	Pure tone	···	DRS-12&DRS-8	91[a]	86–98	14	65	9
0.25	FM[b]	0.2475–0.2525	DRS-8	85	79–89	10	42	6
0.5	FM	0.495–0.505	DRS-8	81	74–87	10	43	3
1	FM	0.99–1.01	DRS-8	78	72–83	10	44	8
1.5	FM	1.485–1.515	DRS-8	74	70–80	10	47	6
2	FM	1.98–2.02	DRS-8	93	86–96	10	52	7
3	FM	2.97–3.03	DRS-8	77	74–82	10	49	7
4	FM	3.96–4.04	DRS-8	73	69–76	10	60	4
8	FM	7.92–8.08	DRS-8	71	69–75	10	52	3
10	FM	9.9–10.1	DRS-8	72	66–74	10	60	2
12	FM	11.88–12.12	DRS-8	67	63–75	10	48	7
14	FM	13.86–14.14	DRS-8	99	92–104	10	47	5
15	FM	14.85–15.15	DRS-8	122	116–126	10	61	6
16	FM	15.84–16.16	DRS-8	>131	···	3	···	4
32	FM	31.68–32.32	DRS-8	>127	···	3	···	3

[a]Based on four sessions with the DRS-12 transducer and ten sessions with the DRS-8 transducer.
[b]Including one session with pure tones.
[c]Number of prestimulus responses in all (signal present+signal absent trials) trials.

water hearing sensitivity of this species is not available. Underwater audiograms have so far been obtained for only 8 of the 34 pinniped species, and of the 3 families (Phocidae, Otariidae and Odobenidae), only species of the first 2 have been tested. Knowledge of the underwater hearing sensitivity of the Pacific walrus could be used to determine appropriate locations for bottom trawl fishing, drilling platforms and shipping routes to minimize their impact on walrus ecology. Information on the underwater hearing abilities of the walrus is therefore important in the development of a strategy to protect this marine mammal. In addition, knowledge of walrus underwater hearing will increase our understanding of this species' behavior (intraspecific communication, predator avoidance, etc.) and ecology, as well as the evolution of sensory systems in pinnipeds. The walrus is the only representative of the family of Odobenidae, a family which is of interest since it has morphological features observed both in Phocidae (seals) and in Otariidae (sea lions).

Therefore, the aim of this study was to test the underwater hearing (frequency range of hearing and frequencies of best hearing sensitivity) of a captive Pacific walrus.

II. MATERIALS AND METHODS

A. Study animal

The study animal was a male Pacific walrus (Igor, code OrZH 003) which was orphaned on the coast of Alaska at the age of a few months. The animal arrived at the Harderwijk Marine Mammal Park at the age of three years. During the current experiment, the animal was healthy, 18 years old, and his body weight varied between 1557 and 1660 kg. Veterinary records showed that the animal had not been exposed to ototoxic medication. The animal was chosen because he had cooperated well in previous psychophysical sensory tests (Kastelein and van Gaalen, 1988; Kastelein *et al.*, 1990, 1996b). The animal received 10–40 kg of thawed fish (herring, *Clupea harengus* and squid, *Loliga* sp.) per day, divided over five to six meals (Kastelein *et al.*, 2000).

B. Study area

The experiment was conducted in an outdoor concrete kidney-shaped pool [20 m (l)×12 m (w), on average 3 m deep] with an adjacent haul-out space at the Harderwijk Marine Mammal Park, The Netherlands. The water level was kept constant. The average monthly water temperature varied between 9 °C and 19 °C, the salinity was approximately 1.5% NaCl. The water circulation pump and the pump of the neighboring pool were switched off 10 min before and during sessions, so that there was no current in the pool during the experiments. No other animals were present in the pool during the tests. The equipment used to produce stimuli was housed out of sight of the walrus (10 m away from the animal and trainer), in an observation and data collection cabin (Fig. 1).

C. Test stimuli production and calibration

Acoustic signals were produced by a wave form generator (Hewlett Packard, model 33120A). Most stimuli consisted of a narrow-band sinusoidal frequency-modulated (FM) signal. The modulation range of the signal was ±1% of the center frequency (the frequency around which the signal fluctuated symmetrically), and the modulation frequency was 100 Hz. Table I shows the frequency ranges of the signals. For example, if the center frequency was 10 kHz, the frequency fluctuated 100 times per second between 9.9 and 10.1 kHz.

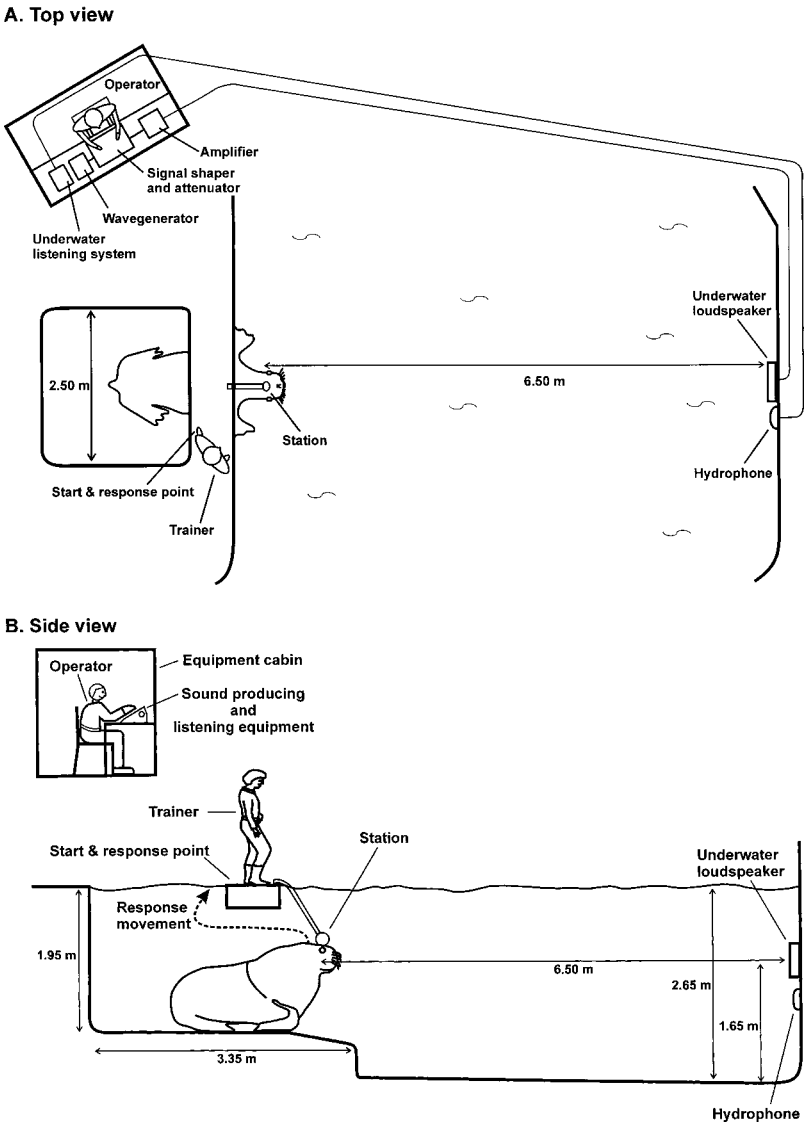


FIG. 1. The study area, showing the walrus under water in the correct position at the station, and its response movement: (a) top view and (b) side view. Both figures are to scale.

In most previous studies of pinniped hearing, pure tones are used as test signals. In the present study, narrow-band FM signals were used because such signals reduce constructive and destructive interference effects (standing waves) on the signals reaching the animal. In humans, FM signals tend to have a slightly higher arousal effect than pure tones, causing slightly lower sensitivity thresholds. To investigate the effect of pure tone versus FM signals, the study animal's hearing thresholds were determined for 250-Hz pure tones and narrow-band frequency-modulated signals with a center frequency of 250 Hz.

No harmonic distortions with sufficient energy in the higher frequencies (i.e., with sound pressure levels (SPLs) near the walrus' hearing thresholds for the frequencies of the harmonics) were found for the test frequencies used, except at 0.125 and 0.2 kHz. Therefore, signals at those frequencies were produced without frequency modulation, which eliminated most harmonic distortion. These low frequency signals had longer wavelengths and were therefore less likely to cause standing waves than higher frequencies.

A modified audiometer for human aerial hearing (Midimate, model 602, s/n 29433) was used to control the signal duration and amplitude (Fig. 2). During the main experiment, the stationary portion of the signal was 1500 ms in duration. Rise and fall times (each 50 ms) preceded and followed the signal to prevent abrupt signal onset and offset transients. This signal duration was expected to be more than the hearing system's integration time. The SPL at the walrus' head while at the station could be varied in 5-dB steps. Before each session, the voltage output level (peak-to-peak) of the system, at the input of the impedance matching transformer (while the attenuator was at the same setting as during calibrations), was checked with an oscilloscope (Philips PM3233; Fig. 2).

The majority of the signals (0.2–32 kHz) were projected by an underwater piezoelectric transducer (Ocean Engineering Enterprise, model DRS-8; 25-cm diameter) with its impedance matching transformer (Fig. 2). A wideband amplifier (Toellner, type TOE 7607) was used for four center frequencies (0.2, 15, 16, and 32 kHz) produced with the DRS-8. The transducer was fixed to the pool wall in a protective stainless steel cage during the entire study period. It was directly in front of the study animal when he was positioned at the station, as determined with a laser beam. During mounting,

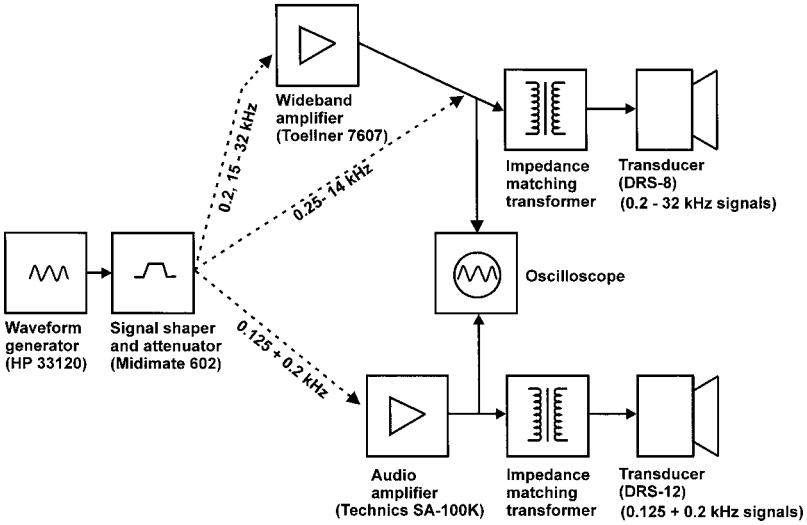


FIG. 2. Block diagram of the signal generation system used in the walrus underwater hearing study.

the front of the transducer was directed towards the walrus' head by means of a mirror on the transducer, so that the animal's head was in the center of the sound beam of the transducer.

The 0.125- and 0.2-kHz signals were projected by another underwater piezoelectric transducer (Ocean Engineering Enterprise, model DRS-12; 30-cm diameter) with its impedance matching transformer. An audio amplifier (Technics, model SA-100k) was used for the two test frequencies produced with the DRS-12. The transducer was suspended with ropes in the water column, just in front of the fixed DRS-8 transducer (Fig. 2). The 0.2-kHz signal was thus produced by and tested with both transducers.

The underwater hearing sensitivity of the walrus was measured for signals with center frequencies of 0.125, 0.2, 0.25, 0.5, 1, 1.5, 2, 3, 4, 8, 10, 12, 14, 15, 16, and 32 kHz. The lowest frequency tested, 0.125 kHz, was selected because the DRS-12 transducer could not produce sounds of <0.125 kHz without distortion. Initially octaves between 0.125 and 32 kHz were tested, but after three sessions with each of these planned frequencies, the 1.5- and 3-kHz signals were added to the tested frequencies to investigate whether the observed relative insensitivity for 2-kHz signals was narrow band or wide band. Also, after the initial three sessions for each frequency, it was determined that the animal was not able to hear the 16- and 32-kHz signals offered at the highest amplitude that could be produced by the equipment. Therefore, the hearing sensitivities for 10-, 12-, 14-, and 15-kHz signals were also tested to determine the high frequency cutoff more accurately, and the 16- and 32-kHz signals were not tested further.

A second experiment to study the effect of signal duration on underwater hearing thresholds was conducted after the main experiment. Four signals (with center frequencies of 0.25, 1, 8, and 14 kHz) were tested with a shorter signal duration (300 ms, excluding 50-ms rise and fall times). Two sessions were conducted for each frequency.

Before each session, the selected signal (usually at a higher amplitude than used during the session) was checked aurally by the signal operator via a hydrophone (LabForce 1BV) next to the transmitting transducer. The hydrophone's output was connected to an amplifier and loudspeaker.

The root-mean-square (rms) SPLs (dB *re* 1 μPa) for each test frequency were measured approximately once each month at the walrus' typical head position when he was at the listening station during the tests (Fig. 1). The walrus was not in the pool during these calibrations. The location of the walrus' head (while at the station) relative to the transducer was very consistent (within a few cm in all six directions); the animal did not adjust its position to a potential spot with a higher signal SPL.

The calibration equipment used for all signals consisted of a broad-band hydrophone (Brüel & Kjaer 8101), with flat frequency response (within 1 dB) in the tested frequency range, a conditioning amplifier (Brüel & Kjaer, Nexus 2690), a computer with an analog input/output card (National Instruments, PCI-MIO-16E-1, 12 bits resolution) and a coaxial module to receive the input signals (National Instruments, model BNC-2090). The system was calibrated with a pistonphone (Brüel & Kjaer, 4223).

During calibration, the SPL was measured for levels that were 10–30 dB above the threshold levels found in the present study. The linearity of the attenuation of the audiometer was checked four times during the study and was precise. The spectra of the signals were monitored to detect potential surface reflections causing cancellations. Multipath arrivals and standing waves introduce both temporal and spatial variations in the observed SPL at the station. Using frequency modulation in generating stimuli will produce smaller amplitude variations. Multi-path arrivals or standing waves are therefore less likely to hamper interpretation of the observations, although they may still occur and introduce a small bias. To assess if multi-path arrivals and standing waves occurred, grid measurements were done during one of the calibration sessions. SPL measurements on a 10-cm grid (up to 50 cm from the station and in all six directions around the station) for 2-kHz stimuli showed variations between 0 and 6 dB.

The signals were digitized at a sample rate 512 kHz and fast Fourier transformed (FFT) into the frequency domain

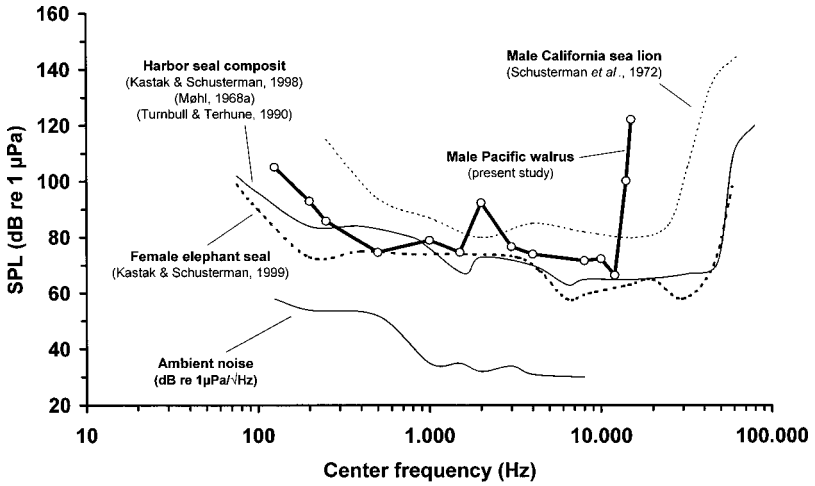


FIG. 3. The mean 50% detection thresholds in dB *re* 1 μPa (rms) for the narrow-band FM signals obtained for the Pacific walrus (for details see Table I). Also shown are the underwater audiograms of a harbor seal [a composite of data from Møhl (1968a), Turnbull and Terhune (1990), and Kastak and Schusterman (1998)], a female northern elephant seal (Kastak and Schusterman, 1999), and a California sea lion (Schusterman *et al.*, 1972). The ambient noise in the pool of the present study up to 8 kHz is plotted in dB *re* 1 μPa/$\sqrt{\text{Hz}}$ (note that this is a different unit than the one along the *Y* axis).

using a Hanning window. The highest peak in the spectrum was selected to determine the SPL and five consecutive 0.2-s time blocks were used to calculate the average SPL. The maximum variation between SPL measurements varied per frequency between 0.2 and 9 dB, but was generally around 5 dB. The average SPL per frequency was calculated based on all calibration sessions. These averages were used to determine the session thresholds.

D. Background noise

No activity took place near the pool during sessions, and the water pumps in a nearby engine room were switched off. The underwater background noise level was measured under the same conditions as during the study. To allow comparison, the recording and analysis methodology of the ambient noise levels was the same as used for the stimuli calibrations. The recorded ambient signal was analyzed by FFT in ten blocks of 0.2 s. The average ambient spectral level was based on 2 s (ten blocks). The ambient noise levels between 0.125 and 8 kHz are plotted in Fig. 3. At higher frequencies the ambient noise could not be measured, as the electronic noise of the recording system was higher than the ambient noise.

E. Experimental procedure and analysis

Training the walrus for the underwater hearing experiment took one month (January 2000). Operant conditioning using positive reinforcement was used to condition the go/no-go response. During preliminary sessions, the rough hearing threshold per test frequency was determined. During the experimental sessions, a trial began when the animal was positioned at the foot of the trainer [start and response point; Fig. 1(a)]. When the trainer gave the animal the vocal command accompanied by a gesture (pointing downwards), the animal descended to the station (a water-filled plastic ball on a water-filled stainless steel tube), so that its external auditory meatus was 6.5 m from the sound source and about 1.0 m below the water surface [Fig. 1(b)]. He was trained to position his nose below the station and his body axis within approximately 5 degrees of the most direct angle to the transducer.

Two types of trial were presented during each test session: signal-present trials and signal-absent trials. In signal-present trials, the stimulus was presented unpredictably between 7 and 12 s after the animal was positioned at the station. A minimum waiting time of 7 s was chosen because it took about 5 s for the small waves created by the animal's descent to dissipate. If the animal detected the sound, it responded by leaving the station (go response) at any time during the signal duration and returning to the start and response point [Fig. 1(b)]. The signal operator then informed the trainer (via head-set radio) that the response was correct (a hit), after which the trainer gave a vocal signal and the walrus received a fish reward. If the animal did not respond to the signal, the signal operator informed the trainer that the animal had failed to detect the signal (a miss). The trainer then signaled to the animal (by tapping on the floor) that the trial had ended, thus calling the animal back to the start and response point. No reward was given following a miss. If the animal moved away before a signal was produced (a prestimulus response), the signal operator told the trainer to end the trial and not provide a reward to the animal.

On signal-absent, or catch, trials, the signal operator told the trainer, after a random time period between 3 and 12 s after the walrus had stationed, to end the trial, which was done by blowing a whistle. This was done as quietly as possible to make the difference in received SPLs from the whistle and underwater signal as small as possible. The rationale behind this was that if the animal would be subjected to a high amplitude whistle during a catch trial, he may be less inclined to react to a successive faint test signal. If the animal responded correctly (i.e., remained at the station; a correct rejection), he returned to the start and response point and received a fish reward. If the walrus left the station before the whistle was blown (a prestimulus response), the signal operator told the trainer to end the trial without rewarding the animal. The same amount of fish was given as a reward for correct go and no-go responses. In both signal-present and signal-absent trials, the trainer was unaware of the trial type until the animal had responded.

One frequency was presented in each session. A modified up/down staircase psychometric technique was used (Robinson and Watkins, 1973). This is a variant of the method of limits, which results in a 50% correct detection threshold (Levitt, 1971). Following each hit, the signal amplitude was reduced by 5 dB. Following each miss, the signal

level was increased by 5 dB. Prestimulus responses did not lead to a change in signal amplitude for the next trial. The starting SPL of a session was set at 10–15 dB above the threshold found in the previous session with each frequency. A session always consisted of 20 trials and lasted for about 15 min. Each session consisted of 50% signal-present and 50% signal-absent trials presented in random order based on a pseudorandom series table (Gellermann, 1933; maximum of three consecutive similar trials), with the modification that the first trial in a session was always a signal-absent trial. Each day one of six data collection sheets with different random series was used. Sessions with more than 15% prestimulus responses (3 out of the 20 trials per session) were eliminated (these occurred only about ten times during the year-long project), because these usually coincided with obvious transient ambient noises. Sometimes other behaviors were solicited from the walrus between trials to occupy the animal during short periods of visible or audible disturbances outside the building (such as passing airplanes, flocks of birds, vehicles, etc.).

In order to prevent the animal's learning curve from affecting the shape of the audiogram, the order in which the frequencies were tested was mixed. The study began by usingthe DRS-8 transducer, and specific frequencies were presented in ascending and descending order (this way the difference in frequency between sessions was limited, reducing potential habituation effects to successive frequencies) until about six sessions had been obtained for each frequency. Then the DRS-12 transducer was used to present the 0.125- and 0.25-kHz frequencies in random order until ten mean session thresholds per frequency had been obtained. Thereafter, the frequencies that were projected with the DRS-8 transducer were tested in ascending and descending order until a total of ten sessions had been obtained for each frequency. The 16- and 32-kHz signals were each tested during only three sessions, as the animal never reacted to a signal of those frequencies at the maximum level that could be produced with the available equipment (131 and 122 dB, respectively).

A switch in the walrus' response from a detected signal (a hit) to a nondetected signal (a miss), or *vice versa*, is called a reversal. Intensities at which the animal reversed its response behavior were taken as data points. The 50% detection threshold per frequency was defined as the mean amplitude of all the reversals obtained in ten sessions after the mean session thresholds leveled off, which occurred usually after two or three sessions.

Generally, one experimental session was conducted daily (5 days/week) between 0800 and 0820 h, when the walrus had not been fed for 15 h, and the park was still closed to visitors. Data were collected between April 2000 and February 2001, with a 3-week break in August. The walrus rutting period between January and February was also avoided, since at this time males reduce their food intake greatly and do not cooperate (Kastelein *et al.*, 2000). In total, 3000 trials were used to calculate the thresholds in the main experiment (1500-ms signal duration), and 160 trials in the second experiment (300-ms signal duration).

All statistical analysis was carried out on Minitab for Windows version 13 (Ryan and Joiner, 1994) with a significance level of 5%. Test assumptions were tested for and data transformed as necessary (Zar, 1984).

TABLE II. Comparison of the underwater hearing threshold for four center frequencies obtained with 1500- and 300-ms signal durations.

Frequency (kHz)	Mean threshold with 1500-ms signals (dB *re* 1 μPa)	Mean session threshold range with 1500-ms signals (dB *re* 1 μPa)	Mean session thresholds with 300-ms signals (dB *re* 1 μPa)	
			Session 1	Session 2
0.25	85	79–89	83	84
1	78	72–83	82	82
8	71	69–75	70	70
14	99	92–104	95	92

III. RESULTS

The mean 50% detection thresholds for the 16 narrow-band FM signals of the study animal are listed in Table I. The resulting underwater audiogram for this individual walrus shows the typical mammalian U-shape (Fay, 1988). Maximum sensitivity (67 dB *re* 1 μPa) occurred at 12 kHz. The range of best hearing (10 dB from the maximum sensitivity) was from 1 to 12 kHz (Fig. 3). Sensitivity fell gradually below 1 kHz and dropped off sharply above 12 kHz. At 16 kHz, the maximum available level of 131 dB was inaudible to the walrus. The animal showed a peculiar insensitivity for 2-kHz signals. His much higher sensitivity for 1.5- and 3-kHz signals indicated that this is a narrow-band phenomenon (Table I).

The four sessions using the DRS-12 transducer at 0.2 kHz resulted in similar thresholds (mean 91.0 ± 3.6 dB) as those obtained during ten sessions with the DRS-8 transducer (mean 91.4 ± 4.0 dB) for that frequency. No statistical difference was found between the threshold means (two-sample *T*-Test; $T = -0.18$, degrees of freedom $= 6$, $P = 0.861$). Therefore all reversals of the 14 sessions were used in the calculation of mean 50% detection threshold for the 0.2 kHz signal (Table I and Fig. 3). The match between the 0.2-kHz thresholds obtained with the two transducers suggests that the shape of the audiogram is not influenced by differences in transducer characteristics. After the initial two to three sessions (which were not included in the analysis), the animal's sensitivity for each test frequency was stable over the one-year study period.

The animal's average prestimulus response rate (for both signal present and signal absent trials) varied between 2% and 9%, depending on the frequency (Table I). The overall average prestimulus response rate was 5.5%.

The mean threshold during one session with 250-Hz pure tones was 84 dB *re* 1 μPa while the mean threshold of the nine sessions with FM signals with a center frequency of 250 Hz was 85 dB *re* 1 μPa.

The thresholds obtained for the four frequencies tested with the 300-ms signal durations in the second experiment are shown in Table II. The thresholds obtained with 300-ms signal durations do not differ significantly from the thresholds obtained with 1500-ms signal durations (two-way

TABLE III. Two-way ANOVA for effects of frequency (250, 1000, 8000, and 14000 Hz) and duration (300 and 1500 ms) on log transformed walrus hearing thresholds [$\log(x-1)$].

Source of variation	Degrees of freedom	Adjusted means square	Test statistic (F)	Probability (P)
Duration	1	0.000 128	0.49	Not sign.
Frequency	3	0.020 737	79.44	<0.0001
Interaction (Duration×freq.)	3	0.000 630	2.41	Not sign.
Error	40	0.000 261		
Total	47			

ANOVA, Zar, 1984; Table III). Since the purpose of this ANOVA was to test for differences in threshold related to duration of signal, the expected significant result due to frequency was not followed by *posthoc* testing to explore where the significant differences lay.

IV. DISCUSSION AND CONCLUSIONS

A. Evaluation of the data

The use of FM signals instead of pure tones probably had no effect on the thresholds found in the present study. This assumption is based on the evidence from the small test with the 250-Hz signals in the present study, suggesting no difference in thresholds between FM and pure tone signals, and on a study with California sea lions by Bullock *et al.* (1971). They used much wider band FM signals than in the present study (1 octave sweeps in 10–50 ms: 3.2→6.4 kHz and 6.4→3.2 kHz; 5-ms rise and fall time), and found similar evoked potential thresholds to those obtained with pure tones.

The influence of ambient noise on the walrus audiogram found in the present study cannot be neglected. If the critical bands of the walrus' hearing are of the same order of magnitude as in humans (that is more than 10%), then at a center frequency of 10 kHz the critical band is more than 1000 Hz. Thus, the spectrum power-density level of 30 dB *re* 1 μPa/$\sqrt{}$Hz results in the noise intensity in the critical band more than 60 dB *re* 1 μPa. At a center frequency of 1 kHz (supposed critical ratio of 100 Hz), the spectrum power-density level near 60 dB *re* 1 μPa/$\sqrt{}$Hz results in the noise intensity in the critical band of 80 dB *re* 1 μPa. This is close to the found threshold values. Therefore the audiogram in the present study may be a conservative estimate of walrus hearing sensitivity.

The unusually sharp bend in the audiogram at the high-frequency cutoff found in the present study is probably related to the small frequency steps that were added once the rough audiogram was determined (after the three initial sessions). Such small (1 kHz) steps were not used in the measurement of the underwater hearing sensitivity of other pinnipeds.

The walrus displayed a sharp (narrow-band) insensitivity to 2-kHz signals. This was not due to variation in SPL of the test signal; the maximum variation between calibrations for this frequency was 4 dB. It could either be a species characteristic, or due to a hearing loss of this individual animal. The 2-kHz insensitivity was not found in the aerial hearing study with the same animal (Kastelein *et al.*, 1996b). A similar insensitivity "notch" at 4 kHz was reported for aerial hearing in a harbor seal (Møhl, 1968a,b), aerial hearing in Northern fur seals and a California sea lion (Moore and Schusterman, 1987) and underwater hearing of a harp seal (Terhune and Ronald, 1972). Notches occurring in the aerial audiogram are not always observed in both the in-air and underwater audiograms of the same animal (Møhl, 1968a,b). It is possible that these "notches" are an effect of wavelength relationships relative to the size of the ear (middle ear cavity and ossicles) of the study animals, causing resonance. The 2-kHz insensitivity found in the present study could also have resulted from a TTS or PTS due to noise in the walrus' present or former environment. The spectra of the following three sound sources which frequently occurred in the animal's environment were measured:

(1) Noise from the water circulation pumps that was radiated into the pool 24 h per day during most of the animal's life. However, with the pumps running, the majority of the energy was below 1 kHz, and no peak in the ambient noise level was found around 2 kHz.
(2) The trainer's whistle which was used during training sessions in some of the years that the study animal was at the park. However, the whistle produced a 7.5 kHz tone.
(3) The animal's own, very loud whistle produced in air (120 dB *re* 1 pW; Verboom and Kastelein, 1995). Since his adolescence 9 years before the present study, the animal had produced this whistle dozens of times per hour, almost 24 h per day, during the 2-month rutting season (and less frequently during the rest of the year). The animal typically whistled when he was within 50 cm of tight curve in the concrete wall around the pool, thus maximizing the amplitude of the reflected sound at his ears. The peak frequency of the walrus' whistle was found to be 1100 Hz, with an almost equally strong first harmonic at around 2200 Hz. The 2-kHz insensitivity was not found in the aerial hearing study with the same animal (Kastelein *et al.*, 1996b). However, at the time of the aerial study, the animal was only 9 years old, the age at which he began to whistle frequently during the rutting period. With age, mammals usually lose high-frequency sensitivity, not low-frequency sensitivity. Replication of the earlier aerial hearing test with the study animal, with the kinds of training and controls used in the present study, could provide an explanation for the 2-kHz reduced hearing sensitivity. This would provide some insight into whether the 2-kHz insensitivity was due to some physical idiosyncrasy of this frequency under water or represents sensorineural hearing loss resulting from exposure to the walrus' own vocalizations.

The thresholds obtained in the second experiment with the 300-ms signals did not differ significantly from the thresholds obtained with the 1500-ms signals in the primary experiment. In retrospect, it may have been better to select a signal duration of around 500 ms. Such a signal is evidently

long enough to be detectable at maximum sensitivity, but reduces the chance of the animal producing prestimulus responses during the signal presentation. With a shorter signal duration researchers may have a better chance of identifying prestimulus responses. In humans, aerial hearing thresholds begin to increase when signal durations drop below 300 ms (Watson and Gengel, 1969). Another advantage of shorter signals is that they reduce the chance of the animal being able to adjust the position of its head and body in order to find regions of slightly higher SPL due to multipath effects. Terhune (1988, 1989), studying the effects of signal duration (0.5, 1, 5, 10, 50, 100, and 500 ms) on underwater hearing thresholds of harbor seals, found that sharp increases in hearing thresholds only occurred when signal duration became less than 50 ms. The absence of an effect on the walrus' thresholds in the present study of shortening the signal duration from 1500 to 300 ms agrees therefore with the findings of Terhune in the harbor seal. The integration time of the walrus is probably also shorter than 300 ms.

To be able to communicate both in air and underwater, the walrus has amphibious hearing. Preliminary free field measurements of aerial hearing were done on male Atlantic walruses (Kastelein *et al.*, 1993). Headphone and free field testing was conducted on a captive 10-year-old male Pacific walrus, the present study animal (Kastelein *et al.*, 1996b). Although these aerial hearing studies did not test a large frequency range, they indicate that walrus hearing in air is fairly flat between 0.25 and 4 kHz (free field test in a non-sound-isolated, echoic, chamber) and between 1 and 8 kHz (headphone test in which the animal probably closed off its external auditory meata). The headphone study showed a drop in sensitivity below 0.25 kHz, which is also seen in the present underwater hearing study.

B. Comparison with other pinniped and sirenian hearing studies

Comparing the findings of the present study to those of other pinniped hearing studies is difficult, as the absolute accuracy of the studies often cannot be established due to a lack of information on the calibration methodology, threshold calculation, and variation in the threshold data between sessions. In addition, most studies have used different methodology and stimuli parameters such as signal type (pure tone versus FM signal) and signal duration. Also, the effort an animal has to perform probably determines the threshold level. In some cases the animals only had to press a paddle next to them as a response, while in the present study the walrus had to move its entire body several meters. The latter situation has probably led to a conservative audiogram, but most other audiograms on pinniped underwater hearing are probably conservative as well.

Despite these limitations, general comparisons can be made between the underwater audiogram of the walrus and the underwater audiograms of the tested pinnipeds. These are the California sea lion, *Zalophus californianus* (Schusterman *et al.*, 1972; Schusterman and Moore, 1978a; Kastak and Schusterman, 1998), harbor seal, *Phoca vitulina* (Møhl, 1968a, b; Terhune, 1988; Terhune and Turnbull, 1995; Kastak and Schusterman, 1998), ringed seal, *Pusa hispida* (Terhune and Ronald, 1975), harp seal, *Pagophilus groenlandicus* (Terhune and Ronald, 1972), northern fur seal, *Callorhinus ursinus* (Schusterman and Moore, 1978b; Moore and Schusterman, 1987), gray seal, *Halichoerus grypus* (Ridgway and Joyce, 1975), Hawaiian monk seal, *Monachus schauinslandi* (Thomas *et al.*, 1990), and northern elephant seal, *Mirounga angustirostris* (Kastak and Schusterman, 1998, 1999). The high frequency cutoff of walrus hearing is much lower than that of other pinnipeds tested so far (Fig. 3). The hearing sensitivity of the walrus between 500 Hz and 12 kHz is similar to that of some phocids. The walrus is much more sensitive to frequencies below 1 kHz than sea lion species tested. In three studies with seals (ringed seal, harp seal and a female northern elephant seal), the underwater hearing sensitivity was tested for frequencies below 1 kHz. In those cases, the hearing sensitivity was similar to that of the walrus in the present study (Terhune and Ronald, 1972, 1975; Kastak and Schusterman, 1998). The present study shows that walrus underwater hearing is adapted to low frequency sounds. The walrus has large middle and inner ear structures compared to the other tested pinnipeds (Kastelein and Gerrits, 1990; Kastelein *et al.*, 1996a). The physics of sounds relate to the size of ear structures: the larger the size, the more suitable to low frequency sound reception. It would therefore be of interest to test the underwater hearing sensitivity of an adult male elephant seal, which is the only pinniped that is as big as, or bigger, than an adult walrus.

An underwater audiogram has been obtained for only one sirenian species: the West Indian manatee, *Trichechus manatus* (Gerstein *et al.*, 1999). The walrus audiogram is more similar to that of the West Indian manatee than it is to the audiograms of the other pinnipeds, but the walrus underwater hearing range is shifted still further towards low frequencies. The similarity between the underwater hearing sensitivities of the walrus and the manatee might be due to their similarity in body weight (which probably relates to the size of the ear structures), or because they are both relatively shallow divers.

C. Ecological significance and suggestions for future research

The maximum underwater hearing sensitivity range (1–12 kHz) obtained in the present study and those found in the aerial studies (ca. 0.5–8 kHz) correspond with the frequency range of the majority of the vocalizations produced by Pacific walruses (ca. 200 Hz to 10 kHz; Fay, 1960; Schevill *et al.*, 1966; Ray and Watkins, 1975; Kibal'chic and Lisitsina, 1979; Miller and Boness, 1983; Stirling *et al.*, 1983, 1987; Fay *et al.*, 1984; Miller, 1985; Verboom and Kastelein, 1995; Kastelein *et al.*, 1995; Nowicki *et al.*, 1997). Therefore, the aerial and underwater auditory sensitivities of the walrus seem well tuned to the majority of its vocalizations. Relatively acute underwater sensitivity to low frequency sounds and waterborne high amplitude, low frequency social vocalizations suggests that walruses can hear each other over relatively great distances. The shape of social signals (i.e., click, sweep, pure tone) influences the detection range as was shown for harbor seals (Turnbull and Terhune, 1994; Terhune, 1989).

The hearing range found in the present study also suggests that walruses can hear many of the social calls of one of their two main predators, the killer whale (*Orcinus orca*; 500 Hz to 35 kHz; Schevill and Watkins, 1966; Awbrey *et al.*, 1982; Ford, 1989). Walruses can probably also hear some of the killer whale's echolocation clicks (4 to $\geqslant$50 kHz; Schevill and Watkins, 1966; Diercks *et al.*, 1971; Awbrey *et al.*, 1982; Barrett-Lennard *et al.*, 1996). In addition, the frequency of the click repetition rate (PRF) of killer whale's echolocation signals (10–330 Hz) is sometimes within the walrus range.

The results from the present study show that walrus hearing is sensitive to most underwater anthropogenic noise (in which usually the energy is below 1 kHz). Audible man-made noise can have many types and degrees of effect on marine mammals (Severinsen, 1990; Cosens and Dueck, 1993; Richardson *et al.*, 1995). Therefore, bottom trawl fishing, tanker routes and drilling platforms should be planned far enough away from areas that are important in walrus ecology. In addition to the hearing sensitivity of the walrus, the distance depends on several other factors such as the general ambient noise level, water depth, ocean floor sediment and the spectrum, source level and duration of the anthropogenic noise. Pinniped hearing is generally more sensitive to anthropogenic noise than odontocete hearing, which is more sensitive to higher frequencies due to need for echolocation.

To estimate at what distances walruses can hear conspecifics, echolocation clicks of killer whales and anthropogenic noise, additional information is needed. Information needs to be obtained on how walruses hear in the presence of masking noise (critical ratios, critical bands), how they hear sounds of different durations (especially pulsed sounds), and how they perceive the spatial positioning of sounds coming from different directions (directivity index).

ACKNOWLEDGMENTS

We thank Rick Wieland, Addie Kappers and Jan Bredeweg for technical assistance and Rob Triesscheijn for making the figures. We also thank Peter van de Sman (Shell, The Netherlands), Herman Walstra (Veenhuis Medical Audio, The Netherlands), Wim Verboom (TNO, Institute of Applied Physics, Delft, The Netherlands), Dana Seagars (U.S. Fish and Wildlife Service, Anchorage, USA), Jack Terhune (University of New Brunswick, Canada), Nancy Vaughan (University of Bristol, UK), Colleen Reichmuth Kastak (Long Marine Laboratory, UCSC, USA), Jen Philips (Hawaii Institute of Marine Biology, University of Hawaii, USA), Brandon Lee Southhall (Long Marine Laboratory, UCSC, USA), Edmund Gerstein (Florida Atlantic University, USA), Whitlow Au (Hawaii Institute of Marine Biology, University of Hawaii, USA) and two anonymous reviewers for their valuable constructive comments on this manuscript. This study was funded by the Harderwijk Marine Mammal Park, The Netherlands, and Veenhuis Medical Audio, Gouda, The Netherlands (special thanks to Marco Veenhuis, Age van Dijk and Herman Walstra). Thanks also to Philippe & Co. (Henk and Teun Kloet) who provided the stainless steel equipment for this study. The first author would like to thank Dana Seagars (U.S. Fish & Wildlife Service, Anchorage, USA) for asking him to conduct a walrus underwater hearing sensitivity study at the Harderwijk Marine Mammal Park at each of the last eight Marine Mammal Society Conferences. This study is dedicated to late Francis (''Bud'') Fay, whose valuable and thorough book on Pacific walruses inspired the first author to study walruses.

Awbrey, F. T., Thomas, J. A., Evans, W. E., and Leatherwood, S. (**1982**). ''Ross Sea Killer Whale Vocalizations: Preliminary Description and Comparison with those of some Northern Hemisphere Killer Whales,'' Rep. Int. Whal. Commn. **32**, 667–670.

Barrett-Lennard, L. G., Ford, J. F. B., and Heise, K. A. (**1996**). ''The mixed blessing of echolocation: differences in sonar use by fish-eating and mammal-eating killer whales,'' Anim. Behav. **51**, 553–565.

Born, E. W., and Knutsen, L. Ø. (**1990**). ''Satellite tracking and behavioural observations of Atlantic walrus (*Odobenus rosmarus rosmarus*) in NE Greenland in 1989,'' Teknisk rapport-Grønlands Hjemmestyre. Afdelingen for Levende Resources, Vol. 20.

Born, E. W., Gjertz, I., and Reeves, R. R. (**1995**). ''Population assessment of Atlantic walrus (*Odobenus rosmarus rosmarus* L.),'' Meddelelser Norsk Polarinstitut **138**, 74–76.

Bullock, T. H., Ridgway, S. H., and Suga, N. (**1971**). ''Acoustically evoked potentials in midbrain auditory structures in sea lions,'' Z. Vergl. Physiol. **74**, 372–387.

Cosens, S. E., and Dueck, L. P. (**1993**). ''Icebreaker noise in Lancester sound, N.W.T., Canada: implications for marine mammal behavior,'' Marine Mammal Sci. **9**, 285–300.

Diercks, K. J., Trochta, R. T., Greenlaw, C. F., and Evans, W. E. (**1971**). ''Recordings and Analysis of dolphin echolocation signals,'' J. Acoust. Soc. Am. **49**, 1729–1732.

Fay, F. H. (**1960**). ''Structure and function of the pharyngeal pouches of the walrus (*Odobenus rosmarus*),'' Mammalia (Paris) **24**, 361–371.

Fay, R. R. (**1988**). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay, Winnetka).

Fay, F. H., and Kelly, B. P. (**1982**). ''Herd composition and response to disturbance of walruses in the Chukohi Sea,'' Cruise Report, KIS Entuziast 1982, NOAA/OCSEAP, POB1808, Juneau, Alaska.

Fay, F. H., Ray, G. C., and Kibal'chich, A. A. (**1984**). ''Time and location of mating and associated behavior of the Pacific walrus, *Odobenus rosmarus divergens* Illiger,'' in *Soviet-American Cooperative Research on Marine Mammals. Pinnipeds*, edited by F. H. Fay and G. A. Fedoseev, NOAA Tech. Rep. NMFS, Vol. 12, pp. 81–88.

Fay, F. H., Kelly, B. P., and Fay, B. A. (eds). (**1990**). ''The ecology and management of walrus populations,'' Report of an international workshop, 26–30 March 1990, Seattle, WA, NTIS PB91-100479.

Fay, F. H., Kelly, B. P., Gehnrich, P. H., Sease, J. L., and Hoover, A. A. (**1986**). ''Modern populations, migrations, demography, trophics, and historical status of the Pacific walrus,'' NOAA/OCSEAP, Envir. Assess. Alaskan Cont. Shelf, Final Rep. Prin. Invest. 37, 231–376. NTIS PB87-107546.

Ford, J. K. B. (**1989**). ''Acoustic behavior of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia,'' Can. J. Zool. **67**, 727–745.

Gellermann, L. W. (**1933**). ''Chance orders of alternating stimuli in visual discrimination experiments,'' J. Gen. Psychol. **42**, 206–208.

Gerstein, E. R., Gerstein, L., Forsythe, S. E., and Blue, J. E. (**1999**). ''The underwater Audiogram of the West Indian manatee (*Trichechus manatus*),'' J. Acoust. Soc. Am. **105**, 3575–3583.

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., and Schusterman, R. J. (**1999**). ''In-air and underwater hearing of a northern elephant seal (*Mirounga angustirostris*),'' Can. J. Zool. **77**, 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**, 1142–1148.

Kastelein, R. A. (**2001**). ''Walrus (*Odobenus rosmarus*),'' in *Encyclopedia of Marine Mammals*, edited by W. F. Perrin, B. Würsig, and J. G. M. Thewissen (Academic, San Diego), pp. 1294–1300.

Kastelein, R. A., and Gerrits, N. M. (**1990**). ''The anatomy of the Walrus

head (*Odobenus rosmarus*). Part 1: The Skull,'' Aquatic Mammals **16**, 101–119.
Kastelein, R. A., and van Gaalen, M. A. (**1988**). ''The sensitivity of the vibrissae of a Pacific Walrus (*Odobenus rosmarus divergens*). Part 1,'' Aquatic Mammals **14**, 123–133.
Kastelein, R. A., Postma, J., and Verboom, W. C. (**1995**). ''Airborne vocalizations of pacific walrus pups (*Odobenus rosmarus divergens*),'' in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil, Woerden, The Netherlands), pp. 265–286.
Kastelein, R. A., Schooneman, N. M., and Wiepkema, P. R. (**2000**). ''Food consumption and Body weight of captive Pacific walruses (*Odobenus rosmarus divergens*),'' Aquatic Mammals **26**, 175–190.
Kastelein, R. A., Stevens, S., and Mosterd, P. (**1990**). ''The tactile sensitivity of the mystacial vibrissae of a Pacific Walrus (*Odobenus rosmarus divergens*). Part 2: Masking,'' Aquatic Mammals **16**, 78–87.
Kastelein, R. A., Dubbeldam, J. L., de Bakker, M. A. G., and Gerrits, N. M. (**1996a**). ''The anatomy of the walrus head (*Odobenus rosmarus*). Part 4: The ears and their function in aerial and underwater hearing,'' Aquatic Mammals **22**, 95–125.
Kastelein, R. A., Mosterd, P., van Ligtenberg, C. L., and Verboom, W. C. (**1996b**). ''Aerial hearing sensitivity tests with a male Pacific walrus (*Odobenus rosmarus divergens*), in the free field and with headphones,'' Aquatic Mammals **22**, 81–93.
Kastelein, R. A., Van Ligtenberg, C. L., Gjertz, I., and Verboom, W. C. (**1993**). ''Free field hearing tests on wild Atlantic walruses (*Odobenus rosmarus rosmarus*) in air,'' Aquatic Mammals **19**, 143–148.
Kibal'chich, A. A., and Lisitsina, T. Yu. (**1979**). ''Some acoustical signals of calves of the Pacific walrus,'' Zool. Zhur. **58**, 1247–1249 (translated by F. H. Fay, 1983).
Miller, E. H. (**1985**). ''Airborne acoustic communication in the Walrus *Odobenus rosmarus*,'' Nat. Geogr. Res., winter, 124–145.
Miller, E. H., and Boness, D. J. (**1983**). ''Summer behavior of Atlantic Walruses *Odobenus rosmarus rosmarus* (L.) at Coats Island, N.W.T. (Canada),'' Z. Säugetierkunde **48**, 298–313.
Levitt, H. (**1970**). ''Transformed Up-down methods in Psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
Møhl, B. (**1968a**). ''Auditory sensitivity of the common seal in air and water,'' J. Aud Res. **8**, 27–38.
Møhl, B. (**1968b**). ''Hearing in seals,'' in *The Behavior and Physiology of Pinnipeds*, edited by R. J. Harrison, R. C. Hubbard, R. S. Peterson, C. E. Rice, and R. J. Schusterman (Appleton-Century-Crofts, New York), pp. 172–195.
Moore, P. W. B., and Schusterman, R. J. (**1987**). ''Audiometric assessment of Northern fur seals, *Callorhinus ursinus*,'' Marine Mammal Sci. **3**, 31–53.
Nowicki, S. N., Stirling, I., and Sjare, B. (**1997**). ''Duration of stereotyped underwater vocal displays by male Atlantic walruses in relation to aerobic dive limit,'' Marine Mammal Sci. **13**, 566–575.
Ray, G. C., and Watkins, W. A. (**1975**). ''Social function of underwater sounds in the walrus (*Odobenus rosmarus*),'' in *Biology of the Seal*, edited by K. Ronald and A. W. Mansfield (Rapp. P.-v. Réun. Cons. Int. Explor. Mer.), Vol. 169, pp. 524–526.
Richardson, W. J., Greene, C. R., Malme, C. I., and Thomson, D. H. (**1995**). *Marine Mammals and Noise* (Academic, San Diego).
Ridgway, S., and Joyce, P. L. (**1975**). ''Studies on seal brain by radiotelemetry,'' Int. Council for the Exploration of the Sea **169**, 81–91.
Robinson, D. E., and Watkins, C. S. (**1973**). ''Psychophysical methods in modern Psychoacoustics,'' in ''*Foundations of Modern Auditory Theory*,'' edited by J. V. Tobias (Academic, New York), Vol. 2, pp. 99–131.
Ryan, B. F., and Joiner, B. L. (**1994**). *Minitab handbook*, 3rd ed. (Duxbury, Belmont, CA).
Schevill, W. E., and Watkins, W. A. (**1966**). ''Sound structure and Directionality in *Orcinus* (Killer Whale),'' Zoologica **51**, 70–76.
Schevill, W. E., Watkins, W. A., and Ray, C. (**1966**). ''Analysis of underwater *Odobenus* calls with remarks on the development and function of the pharyngeal pouches,'' Zoologica **51**, 103–106.
Schusterman, R. J., Balliet, R. F., and Nixon, J. (**1972**). ''Underwater audiogram of the California sea lion by the conditioned vocalization technique,'' J. Exp. Anal Behav. **17**, 339–350.
Schusterman, R. J., and Moore, P. (**1978a**). ''The upper limit of underwater auditory frequency discrimination in the California sea lion,'' J. Acoust. Soc. Am. **63**, 1591–1595.
Schusterman, R. J., and Moore, P. (**1978b**). ''Underwater audiogram of the northern fur seal (*Callorhinus ursinus*),'' J. Acoust. Soc. Am. Suppl. 1 **64**, S87.
Severinsen, T. (**1990**). ''Effects of disturbance on marine mammals,'' in *Environmental Atlas Gipsdalen, Svalbard*, edited by T. Severinsen and R. Hansson, Norwegian Polar Research Institute Report no. 66, Vol. 3, pp. 41–63.
Stirling, I., Calvert, W., and Cleator, H. (**1983**). ''Underwater vocalizations as a tool for studying the distribution and relative abundance of wintering pinnipeds in the High Arctic,'' Arctic **36**, 262–274.
Stirling, I., Calvert, W., and Spencer, C. (**1987**). ''Evidence of stereotyped underwater vocalizations of male Atlantic walruses (*Odobenus rosmarus rosmarus*),'' Can. J. Zool. **65**, 2311–2321.
Terhune, J. M. (**1988**). ''Detection thresholds of a harbor seal to repeated underwater high-frequency, short duration sinusoidal pulses,'' Can. J. Zool. **66**, 1578–1582.
Terhune, J. M. (**1989**). ''Underwater click hearing thresholds of a harbor seal,'' Aquatic Mammal **15**, 22–26.
Terhune, J. M., and Ronald, K. (**1972**). ''The harp seal, *Pagophilus groenlandicus* (Erxleben, 1777). III. The underwater audiogram,'' Can. J. Zool. **50**, 565–569.
Terhune, J. M., and Ronald, K. (**1975**). ''Underwater hearing sensitivity of two ringed seals (*Pusa hispida*),'' Can. J. Zool. **53**, 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. E. Nachtigall (De Spil, Woerden), pp. 81–93.
Thomas, J., Moore, P., Withrow, R., and Stoermer, M. (**1990**). ''Underwater audiogram of a Hawaiian monk seal (*Monachus schauinslandi*),'' J. Acoust. Soc. Am. **87**, 417–420.
Turnbull, S. D., and Terhune, J. M. (**1990**). ''White noise and pure tone masking of pure tone thresholds of a harbour seal listening in air and underwater,'' Can. J. Zool. **68**, 2090–2097.
Turnbull, S. D., and Terhune, J. M. (**1994**). ''Descending frequency swept tones have lower thresholds than ascending frequency swept tones for a harbor seal and human listeners,'' J. Acoust. Soc. Am. **96**, 2631–2636.
Verboom, W. C., and Kastelein, R. A. (**1995**). ''Rutting whistles of a male Pacific walrus (*Odobenus rosmarus divergens*),'' in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil, Woerden), pp. 287–298.
Watson, C. S., and Gengel, R. W. (**1969**). ''Signal duration and signal frequency in relation to auditory sensitivity,'' J. Acoust. Soc. Am. **46**, 989–997.
Zar, J. H. (**1984**). *Biostatistical Analysis*, 2nd ed. (Prentice-Hall International, Englewood Cliffs, NJ).

The magnitude of radiation force on ultrasound contrast agents

Paul A. Dayton,[a] John S. Allen, and Katherine W. Ferrara
Department of Biomedical Engineering, University of California, 1 Shields Avenue, Davis, California 95616

(Received 22 January 2002; revised 29 July 2002; accepted 30 July 2002)

High-speed photography of insonified bubbles with a time resolution of 10 ns allows observations of translation due to radiation force, in addition to the visualization of radial oscillations. A modified version of the Rayleigh–Plesset equation is used to estimate the radius–time behavior of insonified microbubbles, and the accuracy of this model is verified experimentally. The translation of insonified microbubbles is calculated using a differential equation relating the acceleration of the bubble to the forces due to acoustic radiation and the drag imposed by the fluid. Simulations and experiments indicate that microbubbles translate significant distances with clinically relevant parameters. A 1.5 micron radius contrast agent can translate over 5 microns during a single 20-cycle, 2.25 MHz, 380 kPa acoustic pulse, achieving velocities over 0.5 m/s. Therefore, radiation force should be considered during an ultrasonic examination because of the possibility of influencing the position and flow velocity of the contrast agents with the interrogating acoustic beam. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1509428]

PACS numbers: 43.80.Qf, 43.80.Jz, 43.25.Qp [FD]

I. INTRODUCTION

Ultrasonic contrast agents are encapsulated bubbles on the order of 1–10 μm in diameter. These microbubbles are injected into the blood stream in order to increase blood/tissue contrast during an ultrasonic exam. Contrast agent microbubbles are filled with air, or a gas with a lower solubility in blood than air, such as a perfluorocarbon. A shell consisting of albumin, phospholipid, or other material, encapsulates the gas core. Due to this construction, ultrasound contrast agents are highly compressible, and therefore highly echogenic. The use of these microbubble contrast agents has been shown to have clinical importance, particularly in cardiology, where contrast echocardiography has already proven to be a powerful diagnostic tool.[1–3] In addition, there is increasing interest in using contrast agents for radiology applications.[4]

One of the assumptions when using ultrasound contrast agents is that the hydrodynamic behavior of these microbubbles is similar to that of red blood cells. Using intravital microscopy, previous researchers have demonstrated that microbubbles circulate at velocities similar to erythrocytes *in vivo*.[5] Although this may be true for many conditions, Dayton *et al.* observed that it is possible to alter the flow patterns of microbubbles with ultrasound if the imaging parameters create significant acoustic radiation force.[6–8] Under these circumstances, blood velocity estimates made with an ultrasound scanner may be biased. Tortoli *et al.* has demonstrated that insonation of microbubbles with increasing pressure can cause unexpected shifts in the Doppler spectrum.[9] Tortoli *et al.* hypothesized that these changes in the Doppler spectrum were due to the radiation force. With the potential to affect microbubble velocity, it is important to estimate the magnitude of radiation force, and its effect for specific imaging parameters, in order to eliminate errors in blood velocity estimation in a clinical setting.

In addition, the ability to manipulate microbubbles with acoustic energy has the potential to be useful for biological targeting or industrial applications. One of methods of targeting microbubbles involves attaching monoclonal antibodies, biotin, or other adhesion molecules to the shell.[10–14] Many of the binding molecules that are useful for such targeted agents are active on a nanometer scale, so it becomes important to localize bubbles near the site for attachment, typically the vessel wall. Since radiation force has the ability to localize microbubbles along a vessel wall, this may be a potential means to assist the targeting of microbubbles. Also, radiation force may be useful in industrial applications, such as those in microtechnology, where it is desirable to move and manipulate small particles, cells, or bubbles. Previous researchers have discussed the possibilities of using radiation force for the sedimentation of DNA, concentrating erythrocytes for removal from suspension, and the enhancement of immunoprecipitation reactions.[15–17] Moreover, given the widespread use and benefit of acoustic contrast agents in clinical ultrasound, and the potential uses for radiation force in industry, it is important to better understand the effects of the radiation force on microbubbles. To date, the effect of radiation force produced on micron-sized encapsulated gas bubbles with clinical ultrasound has received little attention. Early experimental studies of the influence of an acoustic radiation force on bubbles involved unencapsulated bubbles, kiloHertz frequencies, and bubbles significantly larger than ultrasound contrast agents.[18,19] Clinical ultrasound operates in the low megaHertz range, and ultrasound contrast agents are encapsulated bubbles typically 1–10 microns in diameter. In addition, previous researchers have not incorporated experimentally measured radius–time oscillations of insonified contrast agents into a theoretical treatment of radiation force.

The high-speed optical imaging system discussed in this paper provides high-resolution measurements of the radius–time oscillations of insonified microbubbles, in addition to

[a]Author to whom correspondence should be addressed; electronic mail: padayton@ucdavis.edu

the effects of radiation force on single microbubbles over an acoustic pulse. In this paper, theory and experimental measurements are compared to provide a comprehensive analysis of the effects of the radiation force on microbubbles. Our paper is organized as follows. First, we introduce a modified version of the Rayleigh–Plesset equation used to simulate the oscillations of insonified microbubbles. Second, the forces upon a highly compressible bubble much smaller than the acoustic wavelength that is oscillating and translating in response to an acoustic pulse will be discussed, along with the corresponding equation. We then describe experimental methods and materials, followed by results. Simulations, which model the translation of an insonified microbubble, will be introduced, and then compared with optical observations of contrast agents translating due to an acoustic radiation force. A discussion of results and conclusion follows.

II. THEORY

In this section, we outline the equations used for calculating the displacement of individual microbubbles produced by the radiation force. Two ordinary differential equations are used: one for predicting the radius–time behavior of an insonified microbubble, and another for calculating the translational displacement of the oscillating bubble due to the radiation force from the incident acoustic field. For simplicity, the oscillation and translation of the microbubble are assumed to be independent. The coupling between these effects is assumed to be small, and is further detailed by Brennen.[20]

A. Radial oscillations of the insonified microbubble

A model based on the Rayleigh–Plesset equation is used to predict radius–time oscillations of insonified microbubbles. This model, similar to the modified Herring equation,[21] has been extended for ultrasound contrast agents by Chomas *et al.* and Morgan *et al.*[22,23] Equation (1) describes the model. The pressure at the liquid interface, $p(R,t)$, is given by Eq. (2). The notation for equations described in this manuscript is given in Table I:

$$\rho_l\left(R\ddot{R}+\frac{3}{2}\dot{R}^2\right)=\left(P_0+\frac{2\sigma}{R_0}+\frac{2\chi}{R_0}\right)\left(\frac{R_0^3-R_0^3\left(\frac{b}{V_m}\right)}{R^3(t)-R_0^3\left(\frac{b}{V_m}\right)}\right)^{\gamma}+\frac{R}{c}\frac{d}{dt}p(R,t)-\frac{4\mu\dot{R}}{R}-\frac{2\sigma}{R}-\frac{2\chi}{R}\left(\frac{R_0}{R}\right)^2-12\mu_{\mathrm{sh}}\varepsilon\frac{\dot{R}}{R(R-\varepsilon)}-(P_0+P_{\mathrm{driv}}(t)), \quad (1)$$

$$p(R,t)=\left(P_0+\frac{2\sigma}{R_0}+\frac{2\chi}{R_0}\right)\left(\frac{R_0^3-R_0^3\left(\frac{b}{V_m}\right)}{R^3(t)-R_0^3\left(\frac{b}{V_m}\right)}\right)^{\gamma}-\frac{2\sigma}{R}-\frac{2\chi}{R}\left(\frac{R_0}{R}\right)^2. \quad (2)$$

TABLE I. Notation.

A	Cross-sectional bubble area
b	van der Waals constant
c	Speed of sound in liquid
ε	Thickness of lipid shell
F_{B-G}	Force due to buoyancy
$\mathbf{g}$	Gravitational constant
γ	Polytropic gas exponent
λ	Lift coefficient
p	Pressure at gas/liquid interface
P_{driv}	Acoustic driving pressure
P_l	Pressure in liquid
P_o	Hydrostatic pressure
ρ_l	Liquid density
ρ_b	Bubble density
R	Instantaneous bubble radius
$\dot{R}$	Wall velocity of bubble
$\ddot{R}$	Acceleration of bubble wall
R_o	Bubble resting radius
$\langle R-R_o\rangle$	Time-averaged expansion
s	Data (simulated)
σ	Interfacial tension
t	Time
u_r	Relative velocity of bubble to liquid
u_l	Velocity of liquid
u_b	Velocity of bubble
μ	Medium viscosity
μ_{sh}	Viscosity of lipid shell
μ_f	Coefficient of friction
ν	Kinematic viscosity
V_b	Bubble volume
V_m	Universal molar volume
ω	Angular frequency
x	Direction
χ	Elastic modulus of lipid shell
y	Data (experimental)
z	Error

This model differs from the Rayleigh–Plesset equation in that it incorporates additional terms to account for the shell properties and acoustic radiation damping. The first and second terms on the right of Eq. (1) describe the gas pressure in the bubble and acoustic radiation damping. The viscosity of the surrounding liquid is accounted for with the third term. The fourth term approximates the surface tension of the bubble, and the fifth term approximates shell elasticity. Viscous losses due to the shell are described with the sixth term, and finally, the seventh term describes the driving pressure. In addition, for the interior gas pressure, described with term 1, a van der Waals gas correction $(R_0^3(b/V_m))$ is included to accurately simulate the radial velocities of an encapsulated contrast agent bubble filled with a high molecular weight gas, as given by Chomas *et al.*[21] The variables $R^*=R/R_0$ and $t^*=\omega t$ were used to put the differential equation, Eq. (1), into a nondimensional form, after which it was solved using the ode45 function in MATLAB. This solver uses an explicit Runge–Kutta method. The initial conditions used are $R=R_0$ and $\dot{R}=0$ at time $t=0$. This model provides the bubble radius as a function of time in response to the specified driving pressure also as a function of time.

B. Translation of the insonified microbubble

The translational motion of a microbubble in a fluid during insonation is calculated by solving a particle trajectory

equation, and may be derived from first principles for an unsteady Stokes force on a rigid sphere.[20] The particle is assumed to have a nondeformable boundary and to be isolated from any boundaries or neighboring particles. In this work, we assume that the encapsulated bubble behaves as a rigid sphere in the flow field, and hence remains spherical for drag considerations. However, the bubble is allowed to change size in response to the acoustic field, and the equations used in this work are modified from those that describe a fixed-radius bubble, in order to account for large radial wall oscillations. The reader is referred to the text by Brennen for a more comprehensive review of bubble and particle motion in fluids.[20]

Equation (3) describes the relation used for the drag and driving forces on an insonified contrast agent:

$$\rho_b V_b \frac{du_b}{dt} = -V_b \frac{dP_l}{dx} + \frac{1}{2}\rho_l |u_r| u_r A \frac{24}{\frac{2R|u_l-u_b|}{\nu}} \times \left(1 + 0.197\left(\frac{2R|u_l-u_b|}{\nu}\right)^{0.63} + 2.6 \times 10^{-4}\left(\frac{2R|u_l-u_b|}{\nu}\right)^{1.38}\right) + \left[\frac{1}{2}\rho_1 \frac{d}{dt}(V_b(u_l-u_b))\right] + \left[\frac{3}{2}\frac{\rho_l V_b (u_l-u_b)}{R}\frac{dR}{dt}\right] + (V_b(\rho_l-\rho_b)\mathbf{g} - \mu R^2 (u_l-u_b)^2 \rho_l)\mu_F. \quad (3)$$

The term on the left in Eq. (3) is the product of the mass of the bubble and its acceleration. The first term on the right is the radiation force on a highly compressible bubble due to the acoustic pressure wave.[24–26] The second term on the right is the quasistatic drag force. This term represents translational drag on the bubble from the viscosity of the fluid. Empirical approximations of this term have been determined by researchers for air bubbles of different sizes under various conditions,[27–29] and we use values determined by Meyer *et al.*[29]

The value of u_r is defined as the difference between the bubble velocity and the liquid velocity, $u_r = u_b - u_l$. The liquid velocity, u_l, is in this case determined from the acoustic pressure field, as defined in Eq. (4):[30]

$$\frac{\partial u_l}{\partial t} = -\frac{1}{\rho_l}\frac{\partial P_l}{\partial x}. \quad (4)$$

The third term on the right is the added mass. The term describes the force that must be exerted in order to accelerate a rigid sphere in the surrounding fluid, and is commonly derived in the literature.[20,29,31–33]

The fourth term on the right is an additional added mass term used by Meyer *et al.*, which takes into account the acceleration due to the oscillating bubble wall.[29] The fifth term was incorporated to consider the frictional force prominent in this experimental system. This contribution arises from the drag caused by the bubble sliding against the cellulose tube. In the experimental system, the microbubbles are contained within a cellulose tube. The tube can be approximated as flat for the small translation distances achieved, however, the floating bubble experiences frictional drag as it translates along the bubble wall. We attempted to estimate the magnitude of this effect with the term $(F_{B-G} - \lambda)\mu_F$, where μ_F is the coefficient of friction between the bubble and the cellulose wall. The lift (λ) can be expressed as $\mu R^2 u_r^2 \rho_l$, as described by Krishnan,[34] and the force due to buoyancy is $V_b(\rho_l - \rho_b)\mathbf{g}$. Although the exact value of μ_F for an encapsulated bubble sliding against a cellulose membrane is unknown, the magnitude of this term was calculated for several values within the range from 0.01–0.9. The resulting magnitudes of this drag effect were an order of magnitude smaller than other driving and drag forces acting on the bubble. We hypothesized that due to the fact that the buoyancy and the lift on the bubble are small, the frictional effects between the bubble and the tube wall are also small. Simulations presented were calculated with a friction coefficient of 0.5.

The Basset Force, or history term, which is due to the temporal diffusion of vorticity created at the surface of a particle moving in an unsteady flow field, was disregarded. Several researchers have shown that the history term provides only a small contribution to the total drag on a translating bubble. Park *et al.* estimate that without the history term, the displacement of a bubble of approximately 32 microns in diameter rising through water is overestimated by less than five percent over ten millimeters of travel.[35] The specific form for an oscillating bubble wall is not well understood so this term is neglected for simplicity.

In order to be solved numerically in MATLAB, Eq. (3) is written in a nondimensional form. This resulted in introducing the following nondimensional variables, $R^* = R/R_o$, $t^* = \omega t$, and $u_b^* = u_b/R_o\omega$. An ordinary differential equation solver (ode 113), which is based on a variable order Adams–Bashforth–Moulton PECE solver, was then used to solve for bubble translational velocity (u_b) and acceleration (du_b/dt). Integration of the velocity yields the net translational displacement. The corresponding initial conditions are $u_b = 0$ and $du_b/dt = 0$ at $t = 0$.

III. METHODS

In this section, we detail the materials and methods used for studies described in this paper. In the first section, a high-speed imaging system used to optically record the oscillation and translation of insonified bubbles is introduced. In the second section, the experimental validation of the modified Rayleigh–Plesset model described in Sec. II is discussed.

A. Experimental system for observation of radius–time curves

An experimental system, which interfaces a high-speed camera to a microscope, is used for the purpose of observing radius–time curves of insonified microbubbles.[36,37] The camera (Imacon 468 DRS Hadland) captures seven still images and one streak image of the subject. The streak image is an image of a single line of sight over time, and can be thought of as an optical M-mode image. For an oscillating

bubble, this translates to a radius–time curve. This camera provides a time resolution of 10 ns. An Olympus IX70 inverted microscope was interfaced to the high-speed camera in order to provide sufficient magnification of the micron-sized bubbles.

The experimental contrast agent MP1950 was used for all of the studies described in this work. MP1950 is a phospholipid-shelled microbubble with a decafluorobutane core. The contrast agent was diluted in distilled water until an optical observation confirmed approximately one bubble per optical field of view (approximately 1:50 000). The contrast agent solution was pumped through a 200 μm cellulose tube (Spectrum Labs, Inc.) with a manual microinjector (IM-5B, Narishige, Inc.). The cellulose tube was positioned in a water bath chamber, where it was placed at the mutual focus of both the objective and ultrasonic transducer. The cellulose tube was nearly optically and acoustically transparent. A spherically focused 2.25 MHz transducer (V305, Panametrics, Inc.) was used as the ultrasound source. This transducer had a 6 dB bandwidth of 1.6–3.3 MHz, an aperture diameter of 0.8 in., and a focal length of 2 in. An arbitrary waveform generator (AWG 2021, Tektronix, Inc.) generated the rf transmission waveform with variable parameters. This allowed the use of insonation pulses of different length, amplitude, and center frequency. The transmission signal was amplified approximately 55 dB with an rf amplifier (3100LA, ENI). The transducer was then calibrated while it was focused on the optical field of view with a needle hydrophone (PZT-Z44-0200, Specialty Engineering Associates) and a preamplifier (A17dB, Specialty Engineering Associates) connected to an oscilloscope. The tip of the needle hydrophone was positioned in the optical field of view so that the pressure measured was the same as that experienced by bubbles in the optical field of view.

A computer system (XPS 550, Dell) with timing software (IMACON 468, DRS Hadland) allowed precise timing of the acoustic pulse, the camera shutter, and a high-intensity flash for illumination. Timing of the arbitrary waveform trigger, strobe flash, and camera frames were controlled independently with nanosecond accuracy. The flash unit was a custom designed xenon flash, which back-illuminated the microsphere through a fiberoptic cable, positioned approximately 3 mm above the microscope objective, on the opposite side of the cellulose tube.

The recorded optical radius–time curves were transferred to a computer (Macintosh G4, Apple, Inc.), where the data were digitized by tracing in Photoshop (Adobe).

B. Validation of modified Rayleigh–Plesset model

In order to confirm the validity of the theoretical model used to predict bubble oscillations, experimentally determined radius–time curves were compared with simulations. The oscillations of over 70 bubbles were recorded using the high-speed camera system previously described. Bubbles from 0.8 to 3 microns in resting radius were insonified with 10- and 20-cycle pulses, at 180 and 380 kPa. Insonation with higher acoustic pressures resulted in destruction for a large percentage of contrast agents, and therefore acoustic pressures were kept below 400 kPa.

The mean square error (MSE) of the simulated and experimental radius–time curves was calculated using Eq. (5), where z is the error, s is the simulation, and y is the experimental data:

$$z^2=\frac{1/2[\overline{((s-\overline{s})-(y-\overline{y}))^2}]}{\overline{(s-\overline{s})^2+(y-\overline{y})^2}}. \tag{5}$$

The average mean square error between the theoretical and experimental radius–time curves for the ensemble of bubbles analyzed was 0.34 for 180 kPa insonation, and 0.32 for 380 kPa insonation. These values are in agreement with those determined during validation of the model by previous researchers.[23] Therefore, we conclude that the model accurately simulates radius–time behavior of insonified microbubbles at low acoustic pressures.

IV. RESULTS

In the following section, we detail simulations and observations of the displacement of microbubbles due to acoustic radiation force. In the first section we provide simulations of displacement as predicted with the models described in Sec. II. In the second section we present optical observations of the displacement of microbubbles due to radiation force. Experimentally measured displacement is compared with the simulated displacement. In the third section, simulations are used to further examine the magnitude of drag forces acting on the bubble.

A. Simulations of displacement for the insonified microbubble

In this section, the displacement of individual microbubbles due to a single 20-cycle acoustic pulse was estimated. Using Eq. (1) and Eq. (3), the displacement was calculated for individual microbubbles moving through water for various acoustic pressures, center frequencies, bubble resting radii, and pulse lengths. Figures 1(a)–(c) illustrate the effects of acoustic pressure and bubble resting radius on microbubble displacement for 1, 2.25, and 5 MHz, respectively. In each case, the simulated waveform was a single, 20-cycle acoustic pulse. The resting radius was varied from 0.25 to 3.5 microns, and the acoustic pressure was varied from 100 to 400 kPa. The size range for the simulations was chosen based on the range of bubble sizes most commonly observed experimentally, and the acoustic pressure range was chosen to minimize bubble destruction.[21,36] The effect of the pulse length on microbubble displacement was evaluated by insonifying microbubbles with identical acoustic parameters except for pulse length.

B. Experimental observations of displacement for the insonified microbubble

Streak images for more than 175 bubbles were examined with the high-speed camera in order to record displacement during insonation. These images were analyzed for translational displacement over the duration of one acoustic pulse due to radiation force.

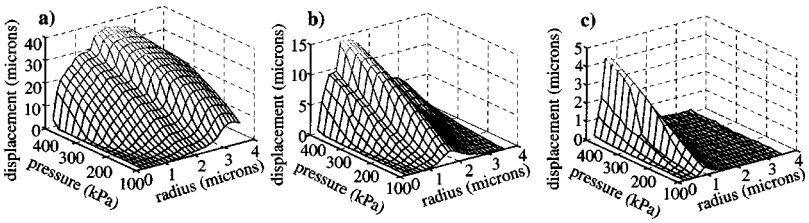


FIG. 1. Simulated radiation force displacement for single bubbles insonified with a 20-cycle pulse, as a function of acoustic pressure from 100 to 400 kPa, and resting radius from 0.25 to 3.5 microns. (a) 1 MHz center frequency. (b) 2.25 MHz center frequency. (c) 5 MHz center frequency. Maximum displacement decreases with increased center frequency, increases with increased acoustic pressure, and is maximized at bubble resonance.

Microbubbles insonified with a 20-cycle acoustic pulse at 180 and 380 kPa were observed to translate several microns in the direction of the acoustic wave propagation. This translation was recorded in the streak image along with the radius–time oscillation. Several examples of the translation and radius–time curves of insonified microbubbles are shown in Fig. 2 and Fig. 3. In these figures, the streak image shows a line of sight through the bubble, and thus illustrates the bubble diameter and axial position in the transducer beam over time. The bubble oscillates due to pressure rarefaction and compression, and simultaneously translates away from the transducer due to acoustic radiation force. The images are recorded through an inverted microscope, such that translation toward the bottom of the image corresponds with displacement away from the transducer. Figure 2 shows streak images of microbubbles of 2.4, 1.7, and 1.5 micron resting radius oscillating and translating in response to 20-cycle 180 kPa pulses at 2.25 MHz. Bubbles translate by approximately 0.7, 3.4, and 3.3 microns, respectively, over 20 cycles. Figure 3 shows streak images for similar parameters, except at a higher acoustic pressure of 380 kPa. In this example, 2.1 and 2.0 micron bubbles are observed to translate 7.2 and 8.7 microns, respectively. Theoretical radius–time curves have been overlaid on the streak images in order to illustrate the similarity between experimental and predicted radius–time curves and reillustrate the accuracy of the model. At 380 kPa, the streak images show that subharmonic oscillation is induced for bubbles of this size range at 2.25 MHz. The resulting translation, coincident with the subharmonics in the bubble oscillation at these parameters, is also shown in Fig. 3.

Figure 4 summarizes the net displacement as a function of resting radius in response to a 20-cycle 180 kPa acoustic pulse. Bubbles are displaced approximately 0.5–4.0 microns, with those of approximately 1.5 microns in resting radius translating furthest. Overlaid on the experimental data is the displacement predicted for the same parameters.

The translational displacement for bubbles insonified with a 20-cycle 380 kPa acoustic pulse at center frequencies of 1.5, 2.25, and 3 MHz is shown in Figs. 5(a)–(c). As shown in Fig. 5(a), single 20-cycle acoustic pulses at 380 kPa and 1.5 MHz displaced microbubbles from 3 to 17 microns, with the maximum displacement occurring for bubbles of approximately 2 microns in resting radius. Figure 5(b) demonstrates that a center frequency of 2.25 MHz produced microbubble displacement from 2 to 8 microns, less than demonstrated at 1.5 MHz, although the maximum displacement appeared to occur for bubbles of approximately the same size. Figure 5(c) indicates that insonation with a center frequency of 3.0 MHz produces microbubble displacement in the range of 0.25 to 4 microns, which is less than that observed for the 2.25 MHz case. The maximum displacement at 3.0 MHz occurred for bubbles of approximate radii of 1.3 microns.

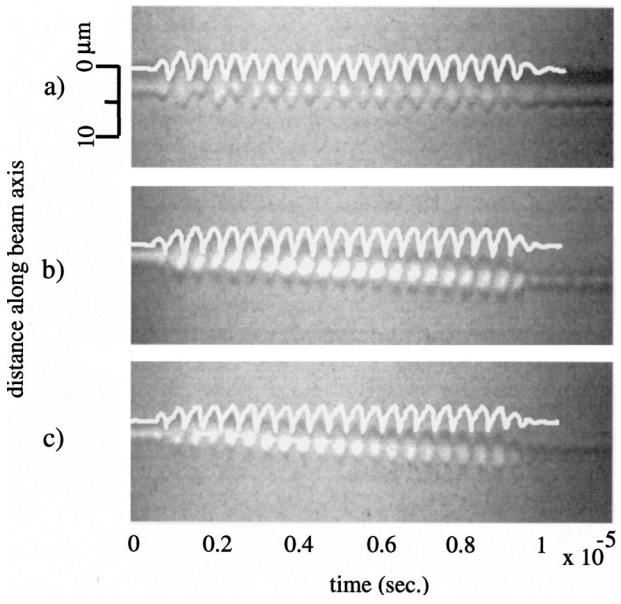


FIG. 2. Streak images show displacement of 2.4 (a), 1.7 (b), and 1.5 (c) micron resting radius bubbles over a 20-cycle, 2.25 MHz, 180 kPa pulse. Simulated *r-t* curves are overlaid to demonstrate the accuracy of the model. Measurement along the *y* axis, which is the displacement along the beam axis, allows a calculation of translational displacement as well as diameter. (Note: simulated *r-t* curves do not incorporate translation.)

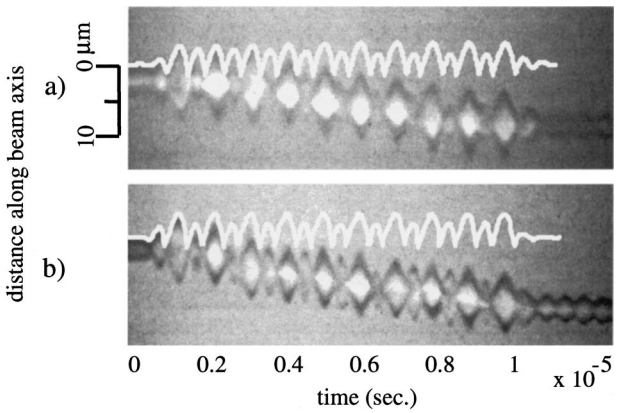


FIG. 3. Streak images show displacement of 2.1 (a) and 2.0 (b) micron resting radius bubbles over a 20-cycle, 2.25 MHz, 380 kPa pulse. Simulated *r-t* curves are overlaid to demonstrate the accuracy of the model. Measurements along the *y* axis, which is the displacement along the beam axis, allows a calculation of the translational displacement as well as diameter. (Note: simulated *r-t* curves do not incorporate translation.)

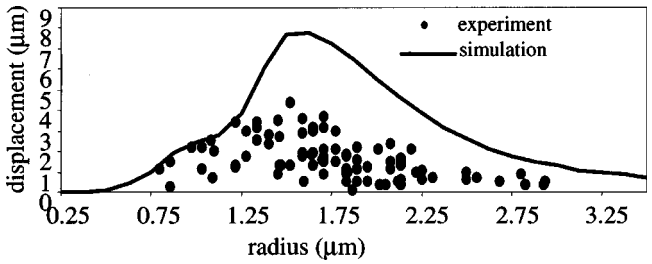


FIG. 4. Experimental and simulated translational displacement for individual bubbles insonified with a single acoustic pulse. Circles indicate measurements from optical recordings of bubble movement. Acoustic parameters include a 20-cycle pulse at 2.25 MHz and 180 kPa. The solid line indicates simulated translation for the same parameters.

C. Simulations illustrating magnitude of drag terms

Figures 6(a)–(c) illustrate simulated radius–time oscillation, translational velocity, and translation distance for 0.5 micron and 2.0 micron resting radius microbubbles, respectively. Results are shown for the first 3 cycles of a 20-cycle, 2.25 MHz, 100 kPa acoustic pulse. In Fig. 6(a), the radius–time curves (normalized by the initial radius) of the two bubbles are plotted as a function of time. The pressure gradient is shown on the same time axis so the reader can observe the relative phase between the pressure and radius oscillation. In Fig. 6(b), the translational velocity of two bubbles is plotted as a function of time. In Fig. 6(c), the corresponding translational displacement is shown on the same time axis.

Table II lists the magnitude of the radiation force produced with a simulated 20-cycle acoustic pulse with a center frequency of 2.25 MHz and an acoustic pressure of 100 kPa. The force is listed in conjunction with the maximum radius, the phase difference between the pressure gradient and radius–time curve, and the time-averaged expansion ($\langle R - R_o \rangle$). For MP1950 insonified at 2.25 MHz, the linear resonance occurs at a resting radius of approximately 1.6 microns. The combination of these parameters produces the trends in minimum and maximum radiation force illustrated in the previous sections.

The phase difference between $R(t)$ and $dP(t)/dx$ is evaluated in Fig. 7 for several bubbles of different initial radius. Parameters include 2.25 MHz center frequency, resting radii from 0.25 to 3.5 microns, and acoustic pressures from 100 to 400 kPa. In this plot, the minimum phase is shown to occur for bubbles of approximately 1.6 microns in resting radius at a low acoustic pressure. A black line in the image denotes the minimum phase lag, where the phase differences on the left of this line indicate that the maximum pressure gradient leads the maximum expansion, and phase differences on the right indicate that the maximum expansion leads the maximum pressure gradient.

The magnitude of the added mass drag terms for bubbles above and near the resonance frequency is shown in Figs. 8(a)–(d). These values are calculated in response to a 20-cycle, 2.25 MHz acoustic pulse at 180 kPa. The magnitude of the translational and oscillatory added mass terms

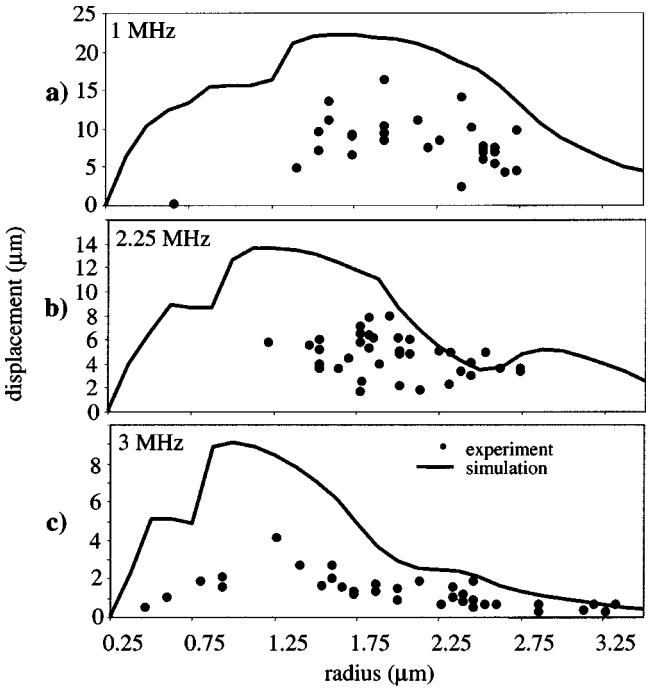


FIG. 5. Experimental and simulated translational displacement for individual bubbles insonified with a single acoustic pulse. Circles indicate measurements from optical recordings of bubble movement. Acoustic parameters include a 20-cycle pulse at 380 kPa and (a) 1 MHz, (b) 2.25 MHz, and (c) 3 MHz. The solid line indicates a simulated translation for the same parameters.

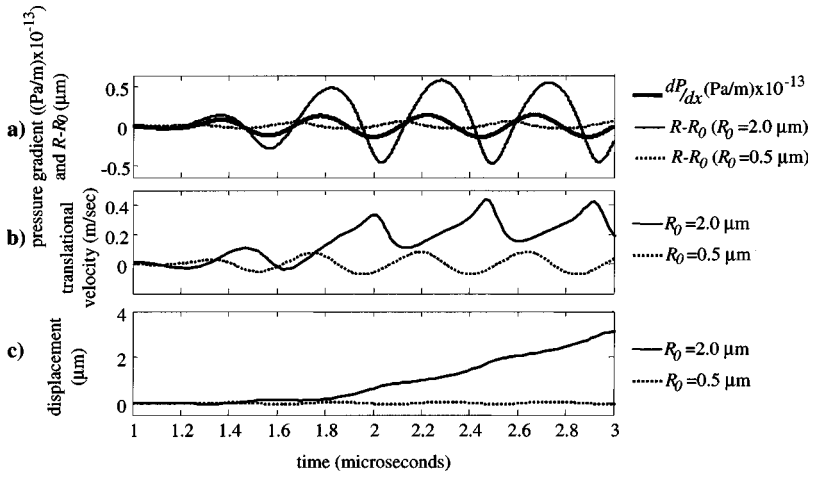


FIG. 6. Simulations of (a) time-varying pressure gradient overlaid on radius–time curves, (b) translational velocity, and (c) translational displacement, for 2.0 and 0.5 micron bubbles insonified at 2.25 MHz and 100 kPa. Plots illustrate approximately the first three cycles of a 20-cycle acoustic pulse.

TABLE II. Table of radiation force as a function of resting radius. Data indicate maximum radius, time-averaged expansion, and phase difference between $R(t)$ and $dP(t)/dx$ at 100 kPa.

Initial radius (microns)	Radiation force (Newtons)	Maximum radius (microns)	$\langle R-R_o \rangle$ (microns)	Phase difference between $R(t)$ and $dP(t)/dx$ (degrees)
0.5	4.8×10^{-9}	0.55	0.00	82.2
1.38	4.0×10^{-6}	1.8	0.09	7.3
1.63	1.07×10^{-5}	2.3	0.15	−11.4
2.88	3.60×10^{-6}	3.2	0.01	−68.4
3.50	2.43×10^{-6}	3.7	0.00	−78.2

is compared for bubbles of different resting radius in Figs. 8(a), (b), and Figs. 8(c), (d), respectively. Bubbles in these figures have resting radii of 0.5 and 2.0 microns, and have resonance frequencies above and near the 2.25 MHz driving frequency, respectively. Not illustrated in this figure is the driving force, which is approximately twice the translational added mass.

V. DISCUSSION

In this section, simulations and experimental results for microbubble displacement due to the radiation force are discussed and related to radius–time oscillation and acoustic parameters. Simulations of microbubble translation (illustrated in Fig. 1) demonstrate the effects of center frequency, microbubble resting radius, and acoustic pressure. In general, increasing acoustic pressure increases the translational displacement of a microbubble. At low acoustic pressures, translational displacement appears to increase linearly, however, with acoustic pressures above 200 kPa, the effect of acoustic pressure becomes nonlinear. This is likely due to the prominent nonlinear oscillations of the bubble at higher acoustic pressures.

For insonation at different center frequencies, it is observed that maximum displacement occurs for microbubbles

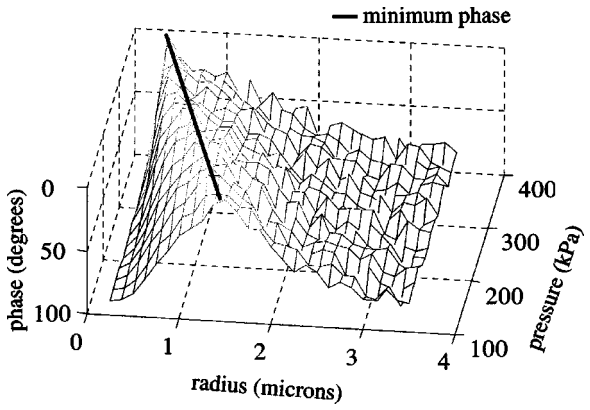


FIG. 7. The calculated phase difference between $R(t)$ and $dP(t)/dx$ for a single bubble insonified with a 20-cycle, 2.25 MHz pulse, as a function of acoustic pressure from 100 to 400 kPa, and resting radius from 0.25 to 3.5 microns. The black line on the plot denotes the minimum phase lag, where the phase differences on the left of this line indicate that the maximum pressure gradient leads the maximum expansion, and phase differences on the right indicate that the maximum expansion leads the maximum pressure gradient.

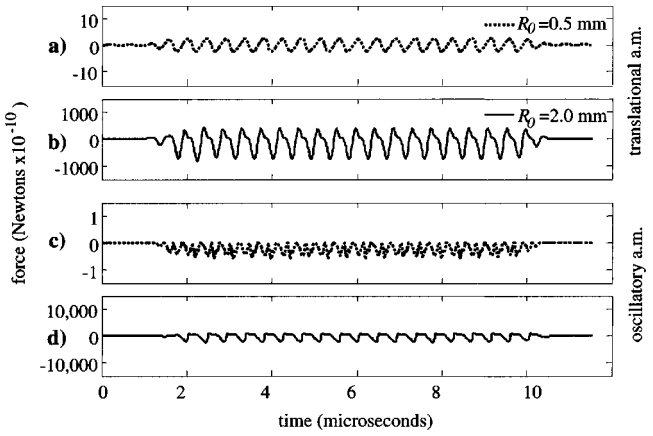


FIG. 8. The simulation values of the translational added mass (a), (b), and oscillatory added mass (c), (d) for 2.0 and 0.5 micron bubbles insonified with a 20-cycle acoustic pulse. Acoustic pressure and center frequency were 100 kPa and 2.25 MHz, respectively.

with a resonance corresponding with the driving frequency (''resonance size'') at the lower-pressure amplitudes. The radius that corresponds with linear resonance for MP1950 is calculated to be approximately 0.8, 1.6, and 3.25 microns at 5, 2.25, and 1 MHz, respectively.[38]

Maximum displacement for 1 MHz insonation is predicted for bubbles of resting radius near 2.2 microns at 400 kPa, and near 3.25 microns at 100 kPa, as shown in Fig. 1(a). Figure 1(b) illustrates displacement with a center frequency of 2.25 MHz, where the maximum displacement occurs for a resting radius of 1.3 microns at 400 kPa, and 1.6 microns at 100 kPa. For 5 MHz insonation, which is illustrated in Fig. 1(c), the maximum displacement occurs near 0.6 microns at 400 kPa, and 0.8 microns at 100 kPa. As acoustic pressure increases, the effective resonance frequency for a fixed resting radius increases due to nonlinear effects.[39]

Similar results were observed experimentally, and are shown in Fig. 4 and Figs. 5(a)–(c). Bubbles insonified with the lowest center frequency tested, 1.5 MHz, demonstrated the greatest translation on average, whereas bubbles insonified with the highest frequency tested, 3 MHz, demonstrated the smallest translation on average.

Additionally, simulations indicate that microbubble translation is maximized when the center frequency matches the fundamental or harmonic resonance frequencies of the insonified microbubble. Bubbles with a resting radius near 1.5 microns experienced a maximum displacement at 2.25 MHz, whereas displacement tapered to half the maximum near 1.0 and 2.25 microns resting radius. For the 1.5 and 3.0 MHz case, the relationship between resonance and displacement was not as clear since the range of bubble diameters was not adequately sampled because of the size distribution of the contrast agents. Another compounding factor was that at 380 kPa, many of the smaller bubbles (less than 1.5 micron resting radius) are destroyed by the acoustic pulse at 2.25 or 1.5 MHz, and hence displacement data for the lower frequencies shows only larger bubbles.

Simulations indicate that the magnitude of microbubble translation is highly dependent on resting radius. For a rest

ing radius far from resonance size, the bubble oscillates minimally, and out of phase with the pressure gradient, shown in Fig. 6(a). The resulting translation is oscillatory toward and away from the source, and is nearly symmetrical, resulting in a zero net translation over an acoustic pulse. This is demonstrated by the 0.5 micron resting radius bubble, which is smaller than the resonance size. The resulting translation is symmetric, as illustrated in Fig. 6(b), and thus in Fig. 6(c), the bubble shows no net displacement over the 20-cycle acoustic pulse. For bubbles near resonance size, translation is preferentially away from the wave source. The 2.0 micron bubble, which is closer to the resonance size, exhibits large oscillations in phase with the pressure gradient, demonstrated in Fig. 6(a). This results in translation preferentially in the direction of the wave propagation, shown in Fig. 6(b), with a resulting displacement of approximately 3.5 microns, illustrated in Fig. 6(c).

Simulations also demonstrated that the translation produced by radiation force increases linearly with pulse length. We hypothesize that this is due to bubbles achieving maximum drag-limited velocity between each acoustic cycle. Optical observations supported this hypothesis.

Data in Table II illustrate that the maximum radiation force is produced on bubbles of approximately 1.6 microns in resting radius, which is approximately the size for linear resonance at 2.25 MHz, as mentioned previously. For bubbles driven at the resonance frequency, the maximum radius is large, the phase difference between $R(t)$ and $dP(t)/dx$ is small, and the time-averaged expansion is large. The combination of effects produces the maximum product of the volume–time curve and the pressure gradient, and maximizes radiation force. For small bubbles (such as 0.5 micron in resting radius), driven below their resonance frequency, the maximum radius is small, time-averaged expansion is small, and the bubble oscillations and the pressure gradient are almost 90° out of phase. For large bubbles (such as 3.5 microns in resting radius), driven above their resonance frequency, the maximum radius is greater than that for bubbles driven at resonance frequency, but the time-averaged expansion is small, and the phase difference between $R(t)$ and $dP(t)/dx$ is nearly 80°. These effects produce a radiation force magnitude that is less than that for bubbles driven at resonance. For the case of the 3.5 micron resting radius bubble (shown in Table II), the magnitude of radiation force is larger than for the 0.5 micron resting radius bubble, since the maximum radius is slightly larger and the phase difference is smaller. The minimum phase lag occurs near the resonance frequency of the bubble, although the phase minimum shifts to smaller radii as the acoustic pressure increases (illustrated in Fig. 7).

For bubbles that expand and contract significantly during insonation, such as those near resonance, the effect of the oscillatory mass is significant. Near resonance, the magnitude of this term is greater than the driving force $(-V_b(dP_l/dx))$. Similarly, the translational added mass terms are also a significant fraction of the magnitude of the driving force. The relative magnitude of these terms for bubbles of different radii is illustrated in Fig. 8.

From simulations of displacement calculated by Eq. (3), the authors estimate that an imaging system operating at 2.25 MHz, 400 kPa, 5 kHz PRF, and with a five-cycle pulse length, will cause a translation of interrogated microbubbles of resonance size by approximately 3.5 mm over 200 μs. Correspondingly, the average translational velocity of these agents would be approximately 17.5 mm/s away from the transducer. This translation may be significant during the measurement of slow flow velocities.

Simulations overestimated experimentally measured displacements by approximately 20%–30%. These results suggest that there were significant uncontrolled variables in each experiment. One likely source of error may be interaction between the bubble and the cellulose tube. Although theoretical analysis indicated that sliding friction between the bubble and the tube wall should be small in comparison to radiation force, bubbles were often observed to stick to the cellulose wall. During the experiment, efforts were made to interrogate bubbles that appeared to be free floating, however, it is probable that some adhesive interaction may have occurred, causing less displacement in some cases than in others. Frictional drag due to adhesion forces between the bubble and the cellulose was not accounted for in the model. A second likely source of error was uncertainties with the measurement of acoustic pressure. The needle hydrophone used for pressure measurements at the time of the experiment was calibrated with an accuracy of $\pm 10\%$. However, experimental comparisons of simulated radius–time curves to those determined experimentally demonstrated that the closest agreement was often achieved with an acoustic pressure up to 15% different from that measured with the hydrophone. Additionally, the nonlinear propagation of the simulated acoustic pulse at higher acoustic pressures was not taken into account in the model. The experimental system had an optical resolution limit near 500 nanometers which also contributes to errors in the radius–time curves, which propagate into the empirical fit of the radial model and hence the translation calculation.

VI. CONCLUSIONS

A model accounting for drag terms on an oscillating bubble overestimates displacement distance but accurately reproduces trends observed experimentally in microbubble displacement as a function of acoustic pressure and initial bubble radius. The authors hypothesize that overestimates may be due to friction between the bubble shell and containment tube, and errors in the acoustic pressure measurements. Using a high-speed imaging system, the authors have shown that a single microbubble can be displaced over 5 microns at a velocity of over 0.5 m/s with a 20-cycle, 380 kPa, 2.25 MHz acoustic pulse. From simulations, the authors estimate that an imaging system operating at clinically relevant parameters may cause an error in estimation of the translational velocity of the interrogated contrast agents, which may be significant during the measurement of slow flow velocities. This mean velocity increases substantially with higher acoustic pressures, with longer pulse lengths, and with higher PRFs.

Translational displacement is determined by the combination of the driving force, the translational added mass, the oscillatory added mass, and the quasistatic drag force. The added mass terms, in particular, are dependent upon the fluctuating bubble volume, and must be included for an accurate analysis of an oscillating bubble. The driving (radiation) force is primarily determined by the maximum bubble expansion, the time-averaged expansion, and the phase between bubble oscillations and the pressure gradient.

Increasing the acoustic pressure increases the displacement due to the radiation force, although in a nonlinear fashion depending on initial bubble size relative to the driving frequency. Varying the resting radius significantly affects bubble translation, with maximum displacement occurring when the bubble is of resonance size. Lower center frequencies result in greater displacement, and higher center frequencies result in smaller displacement.

In summary, it has been shown that the radiation force produced by ultrasound with clinical parameters can translate contrast agent microbubbles significant distances and at significant velocities. This result may have important implications for contrast-assisted ultrasound, such as aiding the manipulation of targeted microbubbles.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Karen Morgan, who contributed to the development of the modified Rayleigh–Plesset model described in this paper. The assistance of Dr. James Chomas is appreciated for the modification of the model to consider a van der Waals gas. Also, we appreciate Jana Gut, for her assistance in data analysis. This work was supported by Grant No. NIH CA76062.

[1] K. Wei and S. Kaul, ''Recent advances in myocardial contrast echocardiography,'' Curr. Opin. Cardiol. **12**, 539–546 (1997).
[2] S. L. Mulvagh, A. N. DeMaria, S. B. Feinstein, P. N. Burns, S. Kaul, J. Miller, M. Monaghan, T. R. Porter, L. J. Shaw, and F. S. Villanueva, ''Contrast echocardiography: Current and future applications,'' J. Am. Soc. Echocardiogr. **13**, 331–342 (2000).
[3] J. R. Lindner, F. S. Villanueva, J. M. Dent, K. Wei, J. Sklenar, and S. Kaul, ''Assessment of resting perfusion with myocardial contrast echocardiography: Theoretical and practical considerations,'' Am. Heart J. **139**(2 Pt.1), 231–240 (2000).
[4] ''Radiological applications of contrast,'' in *Ultrasound Contrast Agents*, edited by B. B. Goldberg, J. S. Raichlen, and F. Forsberg (Martin Dunitz, London, 2001), Part III, pp. 239–3360.
[5] A. R. Jayaweera, N. Edwards, W. P. Glasheen, F. S. Villanueva, R. D. Abbott, and S. Kaul, ''*In vivo* myocardial kinetics of air-filled albumin microbubbles during myocardial contrast echocardiography. Comparison with radiolabeled blood cells,'' Circ. Res. **74**, 1157–1165 (1994).
[6] P. Dayton, A. Goode, S. Klibanov, G. Brandenburger, and K. Ferrara, ''Action of microbubbles when insonified: experimental evidence,'' Proc. IEEE, Vol. 2 1131–1134 (1996).
[7] P. A. Dayton, K. E. Morgan, S. A. Klibanov, G. H. Brandenburger, K. Nightingale, and K. W. Ferrara, ''A preliminary evaluation of the effects of primary and secondary radiation forces on acoustic contrast agents,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**, 1264–1277 (1997).
[8] P. Dayton, A. Klibanov, G. Brandenburger, and K. Ferrara, ''Acoustic radiation force *in-vivo*: a mechanism to assist the targeting of microbubbles,'' Ultrasound Med. Biol. **25**, 1195–1201 (1999).
[9] P. Tortoli, V. Michelassi, F. Guidi, and M. Pratesi, ''Unexpected Doppler effects from microbubbles moving through an ultrasonic beam,'' Proc. IEEE, Vol. 2 1729–1732 (1999).
[10] G. M. Lanza, K. D. Wallace, S. E. Fischer, D. H. Christy, M. J. Scott, R. L. Trousil, W. P. Caheris, J. M. Miller, P. J. Gaffney, and S. A. Wickline, ''High-frequency ultrasonic detection of thrombi with a targeted contrast system,'' Ultrasound Med. Biol. **23**, 863–870 (1997).
[11] G. M. Lanza, K. D. Wallace, M. J. Scott, W. P. Cacheris, D. R. Abendschein, D. H. Christy, A. M. Sharkey, J. G. Miller, P. J. Gaffney, and S. A. Wickline, ''A novel site-targeted ultrasonic contrast agent with broad biomedical application,'' Circulation **95**, 3334–3340 (1997).
[12] A. L. Klibanov, M. S. Hughes, F. S. Villanueva, R. J. Jankowski, W. R. Wagner, J. K. Wojdyla, J. H. Wible, and G. H. Brandenburger, ''Targeting and ultrasound imaging of microbubble-based contrast agents,'' Magn. Reson. Mater. Phys., Biol., Med. **8**, 177–184 (1999).
[13] R. Langer, ''Drug delivery and targeting,'' Nature (London) **392**, 5–10 (1998).
[14] F. S. Villanueva, R. J. Jankowski, S. Klibanov, M. L. Pina, S. M. Alber, S. C. Watkins, and G. H. Brandenburger, ''Microbubbles targeted to intracellular adhesion molecule—1 bind to activated coronary artery endothelial cells,'' Circulation **98**, 1–5 (1998).
[15] K. Yasuda, S. S. Haupt, S. Umemura, T. Yagi, M. Nishida, and Y. Shibatam, ''Using acoustic radiation force as a concentration method for erythrocytes,'' J. Acoust. Soc. Am. **102**, 642–645 (1997).
[16] W. T. Coakley, G. Whitworth, M. Grundy, R. Gould, and R. Allman, ''Ultrasonic manipulation of particles and cells,'' Bioseparation **4**, 73–83 (1994).
[17] M. Groschl, ''Ultrasonic separation of suspended particles—part I: fundamentals,'' Acustica **84**, 432–447 (1998).
[18] I. C. Macedo and W. Yang, ''Acoustic effects of gas bubbles in the flows of viscous fluids and whole blood,'' J. Acoust. Soc. Am. **53**, 1327–1335 (1973).
[19] L. A. Crum, ''Bjerknes forces on bubbles in a stationary sound field,'' J. Acoust. Soc. Am. **57**, 1363–1370 (1975).
[20] C. E. Brennen, *Cavitation and Bubble Dynamics* (Oxford University Press, Oxford, 1995).
[21] K. Vorkurka, ''Comparison of Rayleigh's, Herring's, and Gilmore models of gas bubbles,'' Acustica **59**, 214 (1986).
[22] J. E. Chomas, P. A. Dayton, D. May, J. Allen, A. Klibanov, and K. Ferrara, ''Optical observation of contrast agent destruction,'' Appl. Phys. Lett. **77**, 1056–1058 (2000).
[23] K. E. Morgan, ''Experimental and theoretical evaluation of ultrasonic contrast agent behavior,'' Ph.D. dissertation, University of Virginia, 2000.
[24] L. A. Crum and A. I. Eller, ''The motion of bubbles in a stationary sound field,'' Office of Naval Research Technical Memorandum, 1969, Vol. 61.
[25] V. N. Alekseev, ''Force produced by the acoustic radiation pressure on a sphere,'' Sov. Phys. Acoust. **29**, 77–81 (1983).
[26] A. Prosperetti, ''Bubble phenomena in sound fields: part two,'' Ultrasonics **22**, 97–144 (1984).
[27] V. E. Johnson and T. Hsieh, ''The influence of the trajectories of gas nuclei on cavitation inception,'' *6th Naval Hydrodynamics Symposium*, 1966, pp. 163–182.
[28] T. Watanabe and Y. Kukita, ''Translational and radial motions of a bubble in an acoustic standing wave field,'' Phys. Fluids A **5**, 2682–2688 (1993).
[29] R. S. Meyer, M. L. Billet, and J. W. Holl, ''Freestream nuclei and traveling bubble cavitation,'' Trans. ASME **114**, 672–679 (1992).
[30] H. J. Rath, ''The translational motion of a nonlinear oscillating gas bubble in a compressible liquid in the presence of an inhomogeneous sound field,'' Acustica **44**, 148–155 (1980).
[31] R. Panton, *Incompressible Flow* (Wiley, New York, 1984).
[32] T. R. Auton, J. C. R. Hunt, and M. Prud'homme, ''The force exerted on a body in an inviscid unsteady non-uniform rotational flow,'' J. Fluid Mech. **197**, 241–257 (1988).
[33] D. A. Drew and R. T. Lahey, Jr., ''Some supplemental analysis concerning the virtual mass and lift force on a sphere in a rotating and straining flow,'' Int. J. Multiphase Flow **16**, 1127–1130 (1990).
[34] G. P. Krishnan and D. T. Leighton, Jr., ''Inertial lift on a moving sphere in contact with a plane wall in a shear flow,'' Phys. Fluids **7**, 2538–2545 (1995).
[35] W. C. Park, J. F. Klausner, and R. Mei, ''Unsteady forces on spherical bubbles,'' Exp. Fluids **19**, 167–172 (1995).
[36] J. E. Chomas, P. A. Dayton, D. May, and K. Ferrara, ''Threshold of fragmentation for ultrasonic contrast agents,'' J. Biomed. Opt. **6**, 141–150 (2000).

[37] P. A. Dayton, J. E. Chomas, A. F. H. Lum, J. S. Allen, J. R. Lindner, S. I. Simon, and K. W. Ferrara, ''Optical and acoustical dynamics of microbubble contrast agents inside neutrophils,'' Biophys. J. **80**, 1547–1556 (2001).
[38] K. E. Morgan, J. S. Allen, P. A. Dayton, J. E. Chomas, A. L. Klibanov, and K. W. Ferrara, ''Experimental and theoretical evaluation of microbubble behavior: effect of transmitted phase and bubble size,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 1494–1508 (2000).
[39] M. S. Plesset and A. Prosperetti, ''Bubble dynamics and cavitation,'' Annu. Rev. Fluid Mech. **9**, 145–185 (1977).

Erratum: "Cooperative interaction as the physical basis of the negative stiffness in hair cell stereocilia" [J. Acoust. Soc. Am. 111, 2208–2212 (2002)]

K. H. Iwasa
Section on Biophysics, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, Maryland 20892

G. Ehrenstein
Biophysics Section, National Institute of Neurological Disorders, and Stroke, National Institutes of Health, Bethesda, Maryland 20892

(Received 9 May 2002; accepted for publication 9 May 2002)

[DOI: 10.1121/1.1489690]

PACS numbers: 43.80.Lb, 43.80.Ev, 43.80.Jz, 43.10.Vx

The last two paragraphs in Discussion (p. 2212) should read:

"Negative stiffness appears in physics literature in connection with buckling of elastic objects (Bazant and Cedolin, 1991) and ferroelasticity (Knowles and Steinberg, 1978; Lakes, 2001). The former is deterministic and is related to deformations of macroscopic objects. Buckling increases compliance with increased strain, resulting in negative stiffness. In contrast, the ferroelastic form of negative stiffness is statistical and involves cooperative interactions between microscopic objects analogous to ferroelectricity and ferromagnetism (Salje, 1990).

Negative stiffness resulting from Sachs-Lecar model (Sachs and Lecar, 1991) is analogous to ferroelasticity (Salje, 1990) in that it is the consequence of cooperative interactions."

Bazant, Z., and Cedolin, L. (**1991**). *Stability of Structures* (Oxford U. P., Oxford, UK).

Knowles, J. K., and Steinberg, E. (**1978**). "On the failure of ellipticity and the emergence of discontinuous deformation gradients in plane finite elastostatics," J. Elast. **8**, 329–379.

Lakes, R. S. (**2001**). "Extreme damping in composite materials with a negative stiffness phase," Phys. Rev. Lett. **86**, 2897–2900.

Sachs, F., and Lecar, H. (**1991**). "Stochastic models for mechanical transduction," Biophys. J. **59**, 1143–1145.

Salje, E. (**1990**). *Phase Transitions in Ferroelastic and Co-elastic Crystals* (Cambridge U. P., Cambridge, England).

PROGRAM OF

The First Pan-American/Iberian Meeting on Acoustics • Cancun, Mexico
2–6 December 2002

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, 2 DECEMBER 2002 GRAND CORAL 1 AND 2, 9:30 TO 11:30 A.M.

Opening Ceremonies

Introduction of Choir

Welcoming Remarks

Dr. Juan Ignacio Garcia Salvidea, Mayor of Cancun, Mexico

Richard Stern, President, Acoustical Society of America

James E. West, Meeting Cochair, Acoustical Society of America

Samir N. Y. Gerges, Meeting Cochair and President, Iberoamerican Federation of Acoustics

Sergio Beristain, Meeting Cochair and President, Mexican Institute of Acoustics

MONDAY AFTERNOON, 2 DECEMBER 2002 GRAND CORAL 2, 1:30 TO 3:05 P.M.

Session 1pAAa

Architectural Acoustics and Noise: Design of Offices to Enhance Speech Privacy and Work Performance

Kenneth P. Roy, Chair
Innovation Center, Armstrong World Industries, P.O. Box 3511, Lancaster, Pennsylvania 17604

Chair's Introduction—1:30

Invited Papers

1:35

1pAAa1. Designing for speech privacy in offices. Kenneth P. Roy (Armstrong World Industries, Innovation Ctr., 2500 Columbia Ave., Lancaster, PA 17603)

Speech privacy is an important component of the building interior environment, as it affects the satisfaction and performance of the occupants and thus their productivity. Privacy, both in terms of the confidentiality of conversations and relative to distractions and annoyance, needs to be considered. Field measurements of interzone attenuation, S/N ratio, and privacy calculations will be presented for various open and closed office designs from evaluations in several buildings. Field sound measurement techniques, and electronic masking system design and tuning will be discussed.

2:00

1pAAa2. Design variables for speech privacy in offices. Joel A. Lewitz (Lewitz and Assoc., Inc., 1505 Bridgeway, Ste. 128, Sausalito, CA 94965)

Open plan office environments have a lower potential for speech privacy and freedom from distraction than conventional enclosed offices. The three areas of concern for the acoustical designer are the source, path and listener environments. Given a source spectrum, level and directivity, design elements, which influence attenuation paths, will be addressed with evaluations of their relative importance. The spectrum, level, uniformity and incoherence of background sound at the listener are critical to establishing adequate signal-to-noise ratio for high privacy criteria. Case studies will be addressed comparing challenges, solutions, and results for various open and closed plan spaces. These will be compared to a computer model, which evaluates the relative contribution to privacy of the most important open plan office design variables. Sound masking was shown to be the most powerful and cost effective design tool for establishing acoustical environments with varying privacy criteria.

2:25

1pAAa3. The influence of spatialized background noise on ease of hearing in rooms. Yasushi Shimizu, Abigail Stefaniw, Dana Smith, and Rendell Torres (Prog. in Architectural Acoust., School of Architecture, Rensselaer Polytechnic Inst., Troy, NY 12180)

The influence of irrelevant noise on worker's performance has not yet been studied from the standpoint of spatialized background noise in rooms. This paper presents the detriment to Ease of Hearing due to spatial auditory events produced by intruding sound in open plan offices. Auralization is used to represent a generic open plan office noise environment that includes HVAC noise, an intruding voice, and a focus voice. Two different speech communication environments are presented to the listener. In the first presentation, the intruding voice is located far from the listener's workstation (in the reverberant field) and is not localizable. In the second presentation, the intruding voice is localized near the listener's workstation. The focus voice is localized in front of the listener for both presentations. These sounds are played over a five-channel auralization system. Listeners are tested for the degree of Ease of Hearing (EOH) of the focus voice. Psycho-acoustical test results are presented to show the difference between the Ease of Hearing in each case. Improvement strategies to keep high quality communication are also presented based on the results.

Contributed Paper

2:50

1pAAa4. ClassTalk system for predicting and visualizing speech in noise in classrooms. Murray Hodgson (School of Occupational and Environ. Hygiene, Univ. of British Columbia, 3rd Fl., 2206 East Mall, Vancouver, BC V6T 1Z3, Canada)

This paper discusses the ClassTalk system for modeling, predicting and visualizing speech in noise in classrooms. Modeling involves defining the classroom geometry, sources, sound-absorbing features, and receiver positions. Empirical models, used to predict speech and noise levels, and reverberation times, are described. Male or female speech sources, and overhead-, slide-, or LCD-projector, or ventilation-outlet noise sources, can have four output levels; values are assigned based on ranges of values found from published data and measurements. ClassTalk visualizes the floor plan, speech- and noise-source positions, and the receiver position. The user can walk through the room at will. In real time, six quantities—background-noise level, speech level, signal-to-noise level difference, useful-to-detrimental energy fraction (U50), Speech Transmission Index, and speech intelligibility—are displayed, along with occupied and unoccupied reverberation times. An example of a large classroom before and after treatment is presented. The future development of improved prediction models and of the sound module, which will auralize speech in noise with reverberation, is discussed.

MONDAY AFTERNOON, 2 DECEMBER 2002 GRAND CORAL 2, 3:20 TO 4:50 P.M.

Session 1pAAb

Architectural Acoustics and Noise: Acoustical Design at Hotels

David E. Marsh, Chair
Pelton Marsh Kinsella, 1420 West Mockingbird Lane, Suite 400, Dallas, Texas 75247-4932

Invited Papers

3:20

1pAAb1. Design challenges and coordination issues in hotel projects. Melvin L. Saunders IV and David Marsh (Pelton Marsh Kinsella, 1420 W. Mockingbird Ln., Dallas, TX 75247, saundersml@c-b.com)

The design of a Five Star hotel facility encompasses a myriad of design dilemmas. On the same note, the design of a One Star or Two Star hotel has many dilemmas of its own. The ability of an acoustical consultant, as an integral part of the design team, to recognize the differences between these types of projects can be the difference between a successful hotel project and miserable failure. Different quality hotels require different levels of design criteria. Proper coordination and timing between trades and installations, such as loudspeakers, ceiling coffers, chandeliers, sprinkler heads, and ductwork, is also very important for the success of the overall project. This paper will discuss techniques and methods to produce successful hotel projects, as well as various noise sources throughout these spaces. It will also highlight a number of tips learned through many hotel design experiences.

3:45

1pAAb2. Structure-borne noise at hotels. George Paul Wilson and Deborah A. Jue (Wilson, Ihrig & Assoc., Inc., 5776 Broadway, Oakland, CA 94618-1531, gwilson@wiai.com)

Hotels present a challenging environment for building designers to provide suitable noise and vibration isolation between very incompatible uses. While many are familiar with ways to reduce traditional sources of airborne noise and vibration, structure-borne noise and vibration are often overlooked, often with costly repercussions. Structure-borne noise can be very difficult to pinpoint, and troubleshooting the sources of the vibration can be a tedious process. Therefore, the best approach is to avoid the problem altogether during design, with attention to the building construction, potential vibration sources, building uses and equipment locations. In this paper, the relationship between structure-borne vibration and noise are reviewed, typical vibration sources discussed (e.g., aerobic rooms, laundry rooms, mechanical equipment/building services, and subway rail transit), and key details and design guidance to minimize structure-borne noise provided.

4:10

1pAAb3. The isolation of low frequency impact sounds in hotel construction. John J. LoVerde and David W. Dong (Veneklasen Assoc., 1711 Sixteenth St., Santa Monica, CA 90404)

One of the design challenges in the acoustical design of hotels is reducing low frequency sounds from footfalls occurring on both carpeted and hard-surfaced floors. Research on low frequency impact noise [W. Blazier and R. DuPree, J. Acoust. Soc. Am. **96**, 1521–1532 (1994)] resulted in a conclusion that in wood construction low frequency impact sounds were clearly audible and that feasible control methods were not available. The results of numerous FIIC (Field Impact Insulation Class) measurements performed in accordance with ASTM E1007 indicate the lack of correlation between FIIC ratings and the reaction of occupants in the room below. The measurements presented include FIIC ratings and sound pressure level measurements below the ASTM E1007 low frequency limit of 100 Hertz, and reveal that excessive sound levels in the frequency range of 63 to 100 Hertz correlate with occupant complaints. Based upon this history, a tentative criterion for maximum impact sound level in the low frequency range is presented. The results presented of modifying existing constructions to reduce the transmission of impact sounds at low frequencies indicate that there may be practical solutions to this longstanding problem.

Contributed Paper

4:35

1pAAb4. Meeting space acoustics. Paul L. Tan (Shen Milsom & Wilke, Inc. (Las Vegas), 5729 Bracana Court, Las Vegas, NV 89141)

A general synopsis of acoustic requirements in large dedicated, or multiuse spaces in entertainment, such as casinos, hotels, and themed restaurants is presented. This paper will explore various acoustical performance criteria such as building system noise and vibration, space volume, architectural geometry, sound isolation, and factors influencing speech and music intelligibility, such as reverberation, echoes, etc., Also included in this exploration will be the integration of modern 21st century multimedia and telecommunications technology.

MONDAY AFTERNOON, 2 DECEMBER 2002 CORAL KINGDOM 2 AND 3, 1:00 TO 5:40 P.M.

Session 1pAB

Animal Bioacoustics: Coral Reef Bioacoustics

David Mann, Chair

College of Marine Science, University of South Florida, 140 Seventh Avenue South, Saint Petersburg, Florida 33701

Chair's Introduction—1:00

Invited Papers

1:05

1pAB1. Diversity of fish courtship and spawning sounds and application for monitoring reproduction. Phillip Lobel (Boston Univ. Marine Prog., Marine Biological Lab., Woods Hole, MA 02543)

It has been known for a long time that many fishes produce sounds especially during reproduction. Recent studies using a novel system for synchronous audio and video recordings underwater reveal that several species produce species specific and behavior specific sounds exclusively associated with courtship and the mating act. Passive acoustic technology can utilize these sounds to remotely monitor the breeding behavior of reef fishes. This application is being tested as an alternative or supplement to the traditional methods involving sampling fish gonads, ichthyoplankton net collection, and direct observation by scuba diving. This presentation will show a video with sounds of fishes courting and spawning and will describe their acoustic characteristics.

1:25

1pAB2. Acoustic competition in the gulf toadfish *Opsanus beta*. Michael L. Fine and Robert F. Thorson (Dept. of Biol., Virginia Commonwealth Univ., Richmond, VA 23284-2012, mfine@atlas.vcu.edu)

Nesting male *Opsanus beta* produce boatwhistle advertisement calls and agonistic grunt calls. The grunt is a short-duration pulsatile call, and the boatwhistle is a complex call typically consisting of zero to three introductory grunts, a long tonal boop note, and zero to three shorter boops. The beginning of the boop note is also gruntlike. Anomalous boatwhistles contain a short duration grunt embedded in the tonal portion of the boop or between an introductory grunt and the boop. Embedded grunts have sound pressure levels and frequency spectra that correspond with those of recognized neighbors, suggesting that one fish is grunting during another's call, a phenomenon here termed acoustic tagging. Snaps of nearby pistol shrimp may also be tagged, and chains of tags involving more than two fish occur. The stimulus to tag is a relatively intense sound with a rapid rise time, and tags are generally produced within 100 ms of a trigger stimulus. Time between the trigger and the tag decreases with increased trigger amplitude. Tagging is distinct from increased calling in response to natural calls or stimulatory playbacks since calls rarely overlap other calls or playbacks. Tagging is not generally reciprocal between fish suggesting parallels to dominance displays.

1:45

1pAB3. Nocturnal sound production by longspine squirrelfish (*Holocentrus rufus*) with notes on sound production by fishes on the Turneffe Atoll coral reef in Belize. Joseph Luczkovich, Christopher Stewart (Dept. of Biol., East Carolina Univ., Greenville, NC 27858), and Mark Sprague (East Carolina Univ., Greenville, NC 27858)

Reef fishes produce sounds in response to disturbance, during agonistic encounters in defense of territories, and in spawning events. We recorded in captivity the disturbance sounds of reef fishes in the grunt family (*Haemulidae*), triggerfish family (*Balistidae*), and squirrelfish family (*Holocentridae*) in Belize. Using these recordings and published sonograms for reef fishes, we have identified longspine squirrelfish (*Holocentrus rufus*) in field recordings made on Calabash Caye, Turneffe Atoll, in Belize. We monitored the production of sound by longspine squirrelfish over the course of a transition from day to night. The calling rate of the longspine squirrelfish was greatest after sunset, suggesting that the increased activity of these fish was due to their nocturnal agonistic behavior. A sonobuoy that recorded on a timer 90 s every hour for 24 h failed to detect any longspine squirrelfish calls, perhaps due to their intermittent nature. An unknown sound of biological origin was detected on the sonobuoy recording, but it did not match any of the captive species recorded. A reef fish sound catalogue from captive individuals is being collected for use in the identification of unknown sounds and to monitor the condition of the reefs remotely.

2:05

1pAB4. Snapping shrimps: The bane of bioacoustics monitoring of coral reefs. Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734)

Snapping shrimps are the major source of biological noise in shallow bays, harbors, inlets, and coral reefs in temperate and tropical waters. The sounds are impulsive with a very broad spectrum spanning from several hundred Hz to over 200 kHz. Peak frequencies are typically between 1–10 kHz with a very slow drop-off in energy with increasing frequency. At 200 kHz, the intensity of the sound may be only 20 dB below the intensity at the peak frequency. The peak-to-peak source level can be higher than approximately 189 dB *re* 1 mPa. When a shrimp rapidly closes its snapper claw, a high-velocity water jet that has speed exceeding cavitation conditions is produced. The sound is the result of a cavitating bubble collapsing. In a body of water, the overall level of snapping shrimp sounds can be as high as 20 dB above the noise level typical of sea state 6. The noise is present continuously, with little diurnal and seasonal variations and its broad frequency extent make it extremely difficult to filter this noise. Therefore, snapping shrimp sounds can severely limit the use of underwater bioacoustic monitoring in a coral reef environment.

Contributed Papers

2:25

1pAB5. Choruses from fishes in the Great Barrier Reef. Robert D. McCauley (Ctr. for Marine Sci. and Tech., Curtin Univ. of Technol., Bentley, WA 6102, Australia) and Douglas H. Cato (Defence Sci. and Technol. Organisation, Pyrmont 2009, Australia)

Sounds from fish and invertebrates are a major component of the ambient sea noise around Australia, particularly within the Great Barrier Reef. Choruses from large numbers of calling animals are common and cause large variations in noise level. Long term statistics from the central Great Barrier Reef showed fish calling was a major contributor, both as continuous choruses and as individual calls repeated frequently. Four calling types predominated, with each displaying unique call characteristics and calling patterns through time and space. Durations varied from less than 10 ms to several seconds long, and comprised from one to nearly 40 pulses. These sounds were most common during the Australian summer with one call type also displaying lunar trends. They all showed diurnal variation with the highest activity generally at night, but times of maximum calling rate for each source type differed, consistent with avoiding competition for sound space. The acoustical characteristics of the call types provided information on the fishes calling physiology and behavior. The structure of each pulse could be related to swimbladder mechanics, with swimbladders normally lightly damped. On some occasions, a call type attributed to nocturnal planktivorous fishes may have ensonified much of the Great Barrier Reef.

2:40

1pAB6. Sound production by the toadfish *Sanopus astrifer*. David A. Mann (Univ. of South Florida, College of Marine Sci., 140 7th Ave. S., St. Petersburg, FL 33701, dmann@marine.usf.edu), Wei-Li Diana Ma (Boston Univ., Boston, MA 02215), and Phillip S. Lobel (Boston Univ., Woods Hole, MA 02543)

Sounds produced by the toadfish *Sanopus astrifer* were recorded at Glovers Atoll, Belize. These are the first recorded sounds by a toadfish in the *Sanopus* genus. The boatwhistle call is similar to the boatwhistle produced by *Opsanus* species, in that it consists of a harmonic call about 250

ms in duration with a fundamental frequency of about 180 Hz. Boatwhistles are made in bouts of 1–4 boatwhistles with an inter-boatwhistle interval of 140–170 ms. The first boatwhistle in a bout tends to be longer than subsequent boatwhistles. *S. astrifer* also shows the acoustic tagging phenomenon described by Fine and Thorson for *Opsanus beta*, in which the male produces a short grunt during calling by another male. However, instead of tagging during the boatwhistle like *Opsanus*, *Sanopus* tags between boatwhistles in a bout.

2:55–3:05 Break

Invited Papers

3:05

1pAB7. Sound detection by coral reef fishes. Arthur N. Popper (Dept. of Biol. and Neurosci. & Cognit. Sci. Prog., Univ. of Maryland, College Park, MD 20742, apopper@umd.edu)

Many coral reef fishes use sounds for communication. However, they, as all fishes, learn a good deal about their environment by listening to ambient noises produced by swimming fishes, water moving across the reef, rain, and many other sources. Such information provides these fishes with a sensory world that is far more extensive than from visual, chemical, or tactile cues. There is also a growing body of evidence that suggests that larval coral reef fishes that have not yet settled are able to detect, and swim towards, a reef through the use of the ambient sound of the reef. Studies of hearing by coral reef fishes show considerable variability in sound detection capabilities, sometimes even within the same taxonomic family. Studies have shown that different damsel fishes (*Pomacentridae*) can actually discriminate between sounds produced by their own species and other species within the group. Within the holocentrids, there are similarities in the sounds produced by different species, but considerable differences in the structure of the ear and in hearing capabilities of the same species. This paper will discuss sound detection by coral reef fishes, and the ability of fishes to detect and use the ''auditory scene.''

3:25

1pAB8. The laterophysic connection: A novel specialization thought to enhance sound pressure sensitivity in butterflyfishes (*Chaetodontidae, Chaetodon*). Jacqueline F. Webb (Dept. of Biol., Villanova Univ., Villanova, PA 19087, jacqueline.webb@villanova.edu), W. Leo Smith (American Museum of Natural History, New York, NY), Timothy C. Tricas (Univ. of Hawaii at Manoa, Honolulu, HI 96822), and Darlene R. Ketten (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Morphological specializations that enhance sound reception have been described among fishes, but sources of behaviorally-significant sound stimuli and sound transduction mechanisms are known for very few bony fishes. Butterflyfishes in the genus Chaetodon have a unique sensory specialization, the laterophysic connection (LC), defined anterior swim bladder extensions (horns) in association with medial openings in the lateral line canals of the cranial supracleithral bones. The LC is hypothesized to enhance sensitivity of both the lateral line and ear to sound pressure. Variation in LC morphology among species is defined by variation in soft tissues and is thought to be functionally significant. Preliminary field studies demonstrate that C. multicinctus produces sounds during agonistic interactions, thus providing a source of potentially relevant sound stimuli. Presence of an inflated swim bladder in close proximity to well-developed supracleithral openings in pelagic tholichthys stage individuals (15 mm SL) and horn development at the end of the tholichthys stage (25–29 mm SL) in C. ocellatus suggests that sound may also play an important role in the early life history of Chaetodon. [Work supported by the NSF IBN-9603896 to J.F.W.]

3:45

1pAB9. Larval orientation to reefs and reef sound. John Montgomery, Nicholas Tolimieri (Leigh Marine Lab., Univ. of Auckland, Auckland, New Zealand), and Andrew Jeffs (Natl. Inst. for Water and Atmospheric Res., New Zealand)

Reef systems form a discontinuous spatial mosaic leading many reef organisms to have a dispersal phase in their early life history. At the end of this phase, the larvae must settle onto a reef. Recent evidence supports the view that both larval fish and invertebrates are more active participants in finding suitable settlement habitat than had previously been thought. Not only are they more active swimmers, but field evidence shows that they orient toward reefs from some distance away. Of the possible distance clues, sound is potentially one of the most attractive. Both abiotic and biotic sounds provide habitat-specific information that propagates well underwater and can convey directional clues. Using a combination of reef sound replay, light traps, and choice chambers, it has been shown that larval fish and invertebrates can be attracted to reef sound. Hearing in most fish is based on the otoliths of the inner ear that detect displacement in the sound field. Current theory holds that otolithic hearing is capable of detecting the axis of sound propagation, but not the direction to the source. These experiments provide evidence that fish larvae (even those without a swimbladder) can resolve this 180 degree ambiguity.

4:05

1pAB10. Sound detection of reef sounds by larvae of coral-reef fishes. Jeffrey M. Leis, Brooke M. Carson-Ewart (Australian Museum, 6 College St., Sydney, NSW 2010, Australia, jeffl@austmus.gov.au), and Douglas H. Cato (Defence Sci. and Technol. Organisation, Pyrmont 2009, Australia)

Nearly all teleost reef fishes are pelagic during the larval stage, returning at the end of this stage to settle in a reefal habitat. Little is known about their sensory capabilities or how larvae detect and locate a reef. Sound is the most likely cue, particularly since reef shrimp and fish choruses are audible at considerable distances. Playback experiments off Lizard Island, Great Barrier Reef, broadcast nocturnal reef sounds and artificial sounds (pure tones) in the presence of settlement stage larvae of the coral-reef damselfish *Chromis atripectoralis*, released during the day within 100 m of the underwater speaker 500–1000 m from the nearest reef. Larvae consistently swam to the south in the absence of playback and during playback of artificial sounds. During playback of reef sounds, they showed

no overall swimming direction, indicating that they can hear and can distinguish between a sound with potential biological significance and one without. Light traps with and without playback of reef sounds on 33 paired (playback: nonplayback) trap nights resulted in higher catches for 14 of 17 families for which >10 individuals were trapped. This suggests that not only can larvae hear, but they also have some aural localization capability.

4:25

1pAB11. Ontogenetic development of mechanosensory ability of a pomacentrid coral fish, the red saddleback clownfish (*Amphiprion ephippium*). Hong Yan (School of Biological Sci., Univ. of Kentucky, Lexington, KY 40506-0225, hyyan00@pop.uky.edu), Steve Simpson (Dept. of Biol., Univ. of York, P.O. Box 373, York YO10 5YW, UK), and Matthew Wittenrich (East Aurora, NY 14052)

Recent work on the settlement of coral reef fish suggests that ambient sound could be used as a cue for navigation by the pelagic larvas of reef fishes. Coral reef fish larvas made directional movements toward sound stimulus indicating they are capable of detecting and localizing sound. These findings suggest that some degrees of hearing abilities must have developed prior to settlement. We used embryos and newly hatched larvas of a clown fish (*Amphiprion ephippium*) to investigate the ontogenetic development of the auditory function of this fish. The heart beat rates of embryos and newly hatched larvas (up to 7 days after hatching) in response to various frequencies of sound were measured to indicate their reaction to mechanosensory stimuli. The auditory brainstem response (ABR) recording protocol was used to measure acoustically evoked potentials of Day-8 to Day-36 larvas. Preliminary findings indicate that both the embryos' and larvas reaction to sound stimuli improved (both frequency range and sound pressure level) throughout the development. The electrophysiological observations on embryos and larvas of Amphiprion ephippium provide evidence to support the hypothesis that acoustic stimuli could be used as a cue for coral reef larval fish migration use. [Work supported by NIH, NOHR, IMLS, NERC, AIMS.]

4:45

1pAB12. The application of passive acoustics to assess, monitor, and protect the coral reef ecosystems of the U.S. Pacific Islands. Russell E. Brainard, Kevin B. Wong (NOAA Fisheries, Honolulu Lab., 1125-B Ala Moana Blvd., Honolulu, HI 96825), Scott Ferguson, and Molly Timmers (Joint Inst. for Marine and Atmospheric Res., Univ. of Hawaii)

Activities of the NOAA Fisheries Honolulu Laboratory Coral Reef Ecosystem Investigation (CREI) are presented and areas where the application of passive acoustic methods may significantly contribute to scientific, management, and operational objectives are highlighted. To understand coral reef ecosystems and to reduce adverse human impacts are two fundamental themes of The National Action Plan to Conserve Coral Reefs. The CREI was established to assess, monitor, map, restore, and protect the coral reef ecosystems of the U.S. Pacific Islands. Activities include rapid ecological assessments of fish, corals, algae, and invertebrates, digital video surveys of habitat and fish, acoustic seabed classification surveys, an *in-situ* collection of oceanographic conditions, and satellite remote sensing observations. Resource constraints and the large distances between the U.S. Pacific Islands often limit *in-situ* work to 1–2 years between site visits and prevent adequate temporal monitoring of the biological responses. We propose to develop passive acoustic techniques to monitor some aspects of the health of these remote ecosystems and to develop warning systems to elert scientists and resource managers of large changes or potential threats.

5:05

1pAB13. Analysis of coral reef sound recordings in the Phoenix and Line Islands. Mark W. Sprague, Joseph J. Luczkovich (East Carolina Univ., Greenville, NC 27858, spraguem@mail.ecu.edu), Russell E. Brainard, and Kevin B. Wong (NOAA Fisheries, Honolulu Lab., Honolulu, HI 96814)

Automated acoustic recording stations are being considered to allow managers to monitor remote coral reefs. Sounds of interest include those of biological, environmental, and anthropogenic origin. Preliminary data about remote acoustic monitoring of reefs were obtained in February–March of 2001 and 2002 during National Oceanic and Atmospheric Administration (NOAA) cruises to the Phoenix and Line Islands in the western Pacific Ocean. In 2001, sonobuoys manufactured by East Carolina University (ECU) were deployed to make a 90-s acoustic recording every hour for a 24-h period at each location. In 2002, continuous 24-h acoustic recordings were made using an experimental Remote Underwater Digital Acoustic Recording (RUDAR) system. The ECU sonobuoys recorded diurnal and locational variations in the sounds of snapping shrimp as well as variations in ambient noise. At most locations, the snapping shrimp sound level increased near sunset and decreased near sunrise. The experimental RUDAR system recorded limited data due to some technical problems arising from field use. Analysis of the recordings as well as recommendations for future studies will be presented.

Contributed Paper

5:25

1pAB14. The importance of sound for navigation at settlement in coral reef fishes, with reference to the development of the auditory response. Stephen D. Simpson (Dept. of Biol., Univ. of York, P.O. Box 373, York YO10 5YW, UK), Matthew L. Wittenrich (1256 Luther Rd., East Aurora, NY 14052), Mark G. Meekan (Australian Inst. of Marine Sci., Darwin, NT 0909, Australia), and Hong Y. Yan (Univ. of Kentucky, Lexington, KY 40506)

The behavioral mechanisms that larval coral reef fishes employ to locate suitable settlement habitats on their return from the plankton are poorly understood. The response of settlement-stage reef fishes to reef noise was studied at Lizard Island, Great Barrier Reef, Australia. Catches of fishes in light traps attached to underwater speakers playing reef sounds were twice those of silent traps. To demonstrate that this attraction is important during settlement, sound was used to favorably attract settlers to noisy (over silent) artificial reefs. Many coral reef fish families lay eggs demersally on the reef. By monitoring the heart rates of clownfish embryos in response to artificial noise signals, a broadening of the spectral range of response and a lowering of the threshold of response (an increase in sensitivity) was identified during their development. The thresholds suggest that these embryonic fish will have experienced coral reef noises

prior to leaving the reef. These studies clearly demonstrate the use of sound as a navigation cue for settlement-stage coral reef fishes, and suggest the potential for imprinting. Our findings not only have major implications for how we model recruitment, but also offer some potential management tools for reef systems. [Work supported by Natural Environment Research Council (UK), British Association, Australian Institute of Marine Science, National Organization for Hearing Research (USA), Institute of Museum and Library Services (USA), Department of Education (USA).]

Session 1pAO

Acoustical Oceanography and Underwater Acoustics: David Weston Memorial Session I

D. Vance Holliday, Cochair
BAE Systems, 4669 Murphy Canyon Road, Suite 102, San Diego, California 92123-4333

Ian Roebuck, Cochair
DSTL Naval Systems, Winfrith Technology Center, Dorchester, Dorset DT2 8WX, United Kingdom

Chair's Introduction—1:00

Invited Papers

1:05

1pAO1. Formulas for signal and reverberation with mode-stripping and Lambert's law. Chris H. Harrison (SACLANT Undersea Res. Ctr., Viale San Bartolomeo, 400, 19138 La Spezia, Italy, harrison@saclantc.nato.int)

Extending David Weston's notion of ray invariants and flux it is possible to derive closed-form solutions for two-way propagation and reverberation in range-dependent ducts. In particular, there is interesting and sometimes surprising behavior when the propagation obeys ''mode-stripping'' (the high angles are preferentially attenuated by bottom losses) while the scattering obeys Lambert's Law (high angles are preferentially backscattered). There is a regime where reverberation and target follow the same range law, postponing ''reverberation-limiting'' indefinitely! From the experimental and databasing point of view it is therefore extremely important to know whether Lambert's Law is actually obeyed in practice. These issues and their implications for signal-to-background are addressed. The formulas allow for range-dependent environments and the transition from mode-stripping to single mode propagation at long range. A modified version for range-independent environments includes the effects of refraction on reverberation and signal-to-background. While these formulas cannot compete with numerical calculations in terms of detail, they show at a glance the dependence on such parameters as bathymetry, critical angle, bottom loss (within the critical angle). Here, the behavior is illustrated graphically and compared with a numerical model, SUPREMO.

1:25

1pAO2. David Weston—Ocean science of invariant principles, total accuracy, and appropriate precision. Ian Roebuck (DSTL, Winfrith Technol. Ctr., Dorchester, Dorset DT2 8WX, UK, iroebuck@dstl.gov.uk)

David Weston's entire professional career was as a member of the Royal Navy Scientific Service, working in the field of ocean acoustics and its applications to maritime operations. The breadth of his interests has often been remarked upon, but because of the sensitive nature of his work at the time, it was indeed much more diverse than his published papers showed. This presentation, from the successors to the laboratories he illuminated for many years, is an attempt to fill in at least some of the gaps. The presentation also focuses on the underlying scientific philosophy of David's work, rooted in the British tradition of applicable mathematics and physics. A deep appreciation of the role of invariants and dimensional methods, and awareness of the sensitivity of any models to changes to the input assumptions, was at the heart of his approach. The needs of the Navy kept him rigorous in requiring accuracy, and clear about the distinction between it and precision. Examples of these principles are included, still as relevant today as they were when he insisted on applying them 30 years ago.

1:45

1pAO3. Wavefront modeling in shallow water acoustics. Chris T. Tindle (Phys. Dept., Univ. of Auckland, Auckland, New Zealand)

Early work in collaboration with David Weston on the relationship between ray and normal mode propagation led to ray theory with beam displacement which showed that ray based solutions were not restricted to high frequencies. In turn this has led to acoustic wavefront modeling which allows the efficient calculation of transient acoustic fields in conditions of rapid range dependence. In wavefront modeling, conventional ray tracing is used to find the wavefronts but the amplitude, phase, and travel time of pulses are found from a depth–time diagram at the receiver range. The method allows for rapidly changing water depth due to surface waves. Reflection of sound beneath surface wave crests leads to focusing and caustics which are handled through approximations to a phase function. There is good agreement with experimental results. [Work supported by ONR.]

2:05

1pAO4. Sound absorption due to fish: From David Weston's discoveries to recent developments. Orest Diachok (Naval Res. Lab., Washington, DC 20375, orest@wave.nrl.navy.mil)

David Weston's experiments in the early 1960's on sound propagation between bottom mounted sources and receivers led to an unexpected result: transmission loss in limited frequency bands changed abruptly during twilight. The magnitude of the change: as large as 40 dB at 20 km. Temporal changes in absorption losses coincided with the times when the character of backscattered signals changed from discrete echoes during daytime, to diffuse reverberation at night. Weston's hypotheses: both effects were due to fish with swim bladders, and diurnal changes in the attenuation and reverberation were due to changes in the depth and separation between fish in schools. Weston's discoveries and inferences provided the stimulus for the design of multi-disciplinary ''bioacoustic absorption spectroscopy'' experiments in 1995 in the Gulf of Lion in concert with Ifremer, France (Diachok, 1999), and in 2001 in the Santa Barbara Channel in concert with the Southwest Fisheries Science Center, USA (Diachok *et al.*, 2002). These experiments provided compelling evidence of (1) the biological causes of absorption lines, (2) frequency changes associated with changes in depth and separation at twilight, and (3) good agreement between number densities derived from concurrent absorptivity and fisheries echo sounder measurements. [Research supported by ONR.]

2:25

1pAO5. The contributions of David Weston to the acoustics of gas-filled bladders in fishes. Mardi C. Hastings (Biomed. Eng. Ctr. and Dept. of Mech. Eng., Ohio State Univ., 1080 Carmack Rd., Columbus, OH 43210, hastings.6@osu.edu)

Thirty-five years ago David Weston published ''Sound Propagation in the Presence of Bladder Fish'' [in *Underwater Acoustics*, Vol. 2, edited by V. M. Albers, 1967, pp. 55–88], a seminal contribution that provided the foundation for future research pertaining not only to scattering by fish, but also swimbladder mechanics. In this paper he outlined the theory for the acoustic response of a single swimbladder and then extrapolated to estimate scattering of sound by large schools of fish in the ocean. He addressed the effects of swimbladder geometry, tissue surrounding the swimbladder, and ocean depth on its resonance and scattering characteristics. At that time he pointed out that his work was primarily theoretical because very few experiments had been reported. In the last two decades advances in instrumentation and continued interest in this subject have led to many experiments that have provided data that now confirm much of his original theory. A review of experimental studies inspired by his work and correlation of data with his creative theoretical predictions will be presented.

Contributed Paper

2:45

1pAO6. Three applications of David Weston's work to studies of scattering from fish. C. Feuillade, R. W. Nero, C. H. Thompson (Naval Res. Lab., Stennis Space Center, MS 39529-5004), and R. H. Love (BayouAcoust., Pass Christian, MS 39571-2111)

The Naval Research Laboratory has for many years conducted a program, comprising both experimental and theoretical components, to determine the low- to mid-frequency acoustic scattering characteristics of individual fish and fish schools. This paper discusses three procedures, developed during the course of our work, whose intellectual genesis can be traced directly from David Weston. First, we discuss the ''Weston correction'' to the monopole resonance frequency of a prolate spheroidal air bubble. This is used to model variations in the resonance frequency of a fish swimbladder as its aspect ratio changes under compression, and to facilitate fish species identification and abundance estimation through the water column. Second, we discuss a scattering model, inspired by Weston's work on arrays of air bubbles, which incorporates coherence and multiple scattering effects between swimbladders of comparable size to investigate the levels and fluctuations of scattering from fish schools. Third, we discuss an acoustic reference target, constructed from sheets of ''Bubble-Wrap®'' packaging material and successfully deployed and tested, which is a practical implementation of Weston's theory for the scattering response of planar arrays of identical bubbles. [Work supported by ONR.]

Invited Papers

3:00

1pAO7. Source levels of impulsive sound sources in underwater acoustics. N. Ross Chapman (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada)

Impulsive sound sources have been used extensively in underwater acoustics for many different research applications. Since the initial work by Weston in developing a simple analytical model for an underwater explosion, there have been several theoretical and experimental programs designed to determine source levels. More recently, other types of sources such as air guns and water guns have been introduced from marine seismic research, and there is renewed interest in knowing accurate source levels for assessing the impact on marine environments. In this paper the results of a series of experiments carried out to measure the source levels of several different types of impulsive sources are summarized. These included traditional 0.82-kg SUS charges, small and medium sized air guns from 5–185 cu. in., and a 160 cu. in. water gun. The SUS charges were exploded at shot depths from 18–200 m, and the air guns and water guns were fired at shallow depths from 1–5 m, corresponding to the conventional operating depths. The experiments provided high-quality shot waveforms that were processed to determine calibrated source levels in 1/3 octave frequency bands from 10–600 Hz. The measured values are compared to predictions from Weston's simple model.

3:20–3:35 Break

3:35

1pAO8. Frequency-dependent attenuation in unconsolidated marine sediments. Michael D. Richardson (Marine Geosciences Div., Naval Res. Lab, Stennis Space Center, MS 39529-5004, mike.richardson@nrlssc.navy.mil)

Wood and David Weston (1964) reported that *in situ* attenuation varied linearly with frequency (constant Q) over the frequency range 4 to 70 kHz in mud sediments from Emsworth Harbour, Hampshire UK. This linear frequency dependence seemed at odds with propagation models according to Wood and Weston. Since this pioneering work, actually conducted in 1951–52, numerous authors have measured attenuation in a variety of unconsolidated sediments and over a wide range of frequencies. Hamilton (1972, 1985), Kibblewhite (1989), and Bowles (1997) provide excellent compilations of frequency-dependent attenuation in unconsolidated sediments. Given the measurement difficulties, the wide variety of techniques used to measure attenuation, and changes in the nature of sediment at different spatial scales, the exact form of frequency dependence is still a widely debated issue. Recent measurements over the frequency range of 20–100 kHz in sandy sediments tend to support a linear dependency of attenuation with frequency, at least over this frequency range. [Work supported by ONR.]

Contributed Papers

3:55

1pAO9. Ten years of hydroacoustical observations of the behavior and distribution of small pelagic fish in the west coast of Baja California, Mexico. Carlos Robinson (Instituto de Ciencias del Mar y Limnologia, UNAM, Ciudad Universitaria, Mexico 04500, robmen@servidor.unam.mx)

The main goal is to present results on the behavior and distribution of small pelagic fish on the west coast of Baja California, Mexico using hydroacoustics and related to oceanographic parameters. From December 1993 to September 2002, 20 oceanographic surveys have been done off the west coast of Baja California, Mexico aboard the R/V ''EL PUMA.'' Two areas are covered, a Northern area from Punta Colnet to Punta Baja (30 54′, 116 40′ W to 29 26′ N 115 29′ W) and a Southern area from Punta Eugenia to Bahia Tortugas (27 29′ N 115 22′ W to 26 47′ N 113 55′ W). Results show that since 1993 echo-counting has been reduced significantly in both areas. Results are discussed in terms of oceanographic parameters and the presence of the 1997 El Niño. Before the 1997 El Niño, echo-counting was high in both areas. The behavior is reversed during El Niño. However, in the central area results from March 2000 may suggest a possible change in this tendency.

4:10

1pAO10. Measuring acoustic backscattering of deepwater fish *in situ* using a manned submersible. Kelly Benoit-Bird, Whitlow Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), Christopher Kelley, and Christopher Taylor (Univ. of Hawaii at Manoa, Honolulu, HI 96822)

An outstanding problem in fisheries acoustics is the depth dependence of scattering characteristics, required for sonar surveys, of swimbladder-bearing fish. The swimbladder, which is the major source of acoustic reflectivity, may become compressed due to increasing pressure with water depth, changing backscatter characteristics. To address this, echoes from deepwater snappers were obtained from a manned submersible using broadband sonar. A low-light camera mounted next to the sonar transducer permitted simultaneous identification of species, fish size, and orientation. Data were compared to echoes from captured fish measured at the surface. The relationship between fish length and *in situ* target strength showed no difference from the relationship measured at the surface. No differences in the species-specific temporal echo characteristics were observed between surface and *in situ* measures. This indicates that the size and shape of the fish's swimbladder are maintained both at the surface and at depths of up to 250 m.

4:25

1pAO11. Broadband temporal and spectral structure of acoustic backscatter from Hawaiian Lutjanid snappers. Whitlow W. L. Au and Kelly J. Benoit-Bird (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734)

The characteristics of acoustic echoes from six species of deep dwelling (up to 350 m) Hawaiian Lutjanid snappers were determined by backscatter measurements at the surface. A broadband linear frequency modulated signal and a short dolphin-like sonar signal were used as the incident signal. The fishes were anesthetized and attached to a monofilament net which was in turn attached to a rotor so echoes could be collected along the roll, tilt, and lateral axes. The echo waveforms were complex with many highlights and varied with the orientation of the fish. The highlight structure was determined by calculating the envelope of the cross-correlation function between the incident signal and the echoes. In the tilt plane, the strongest echo occurred when the incident angle was perpendicular to the long axis of the swim bladder. The number of highlights was the fewest at this perpendicular orientation and increased as the fish is tilted from this orientation. The echo structures were easily distinguishable between species and were generally consistent within species. The highlight structure of the echoes resulted in the transfer function being rippled, with local maxima and minima at different frequencies.

4:40

1pAO12. *In situ* acoustic estimates of the swimbladder volume of Atlantic herring, *Clupea harengus*. Redwood W. Nero, Charles H. Thompson (Naval Res. Lab., Stennis Space Center, MS 39529-5004), and J. Michael Jech (Northeast Fisheries Sci. Ctr., Woods Hole, MA 02543)

Most marine fish maintain swimbladder volumes equivalent to 4%–5% of their body weight in order to maintain neutral buoyancy. In many fish the addition or removal of gas from the swimbladder is accomplished with the gas gland, a blood invested portion of the swimbladder wall. However, several families, including the herring family, Clupeidae, lack a gas gland. Instead, these fish possess a pneumatic duct between the esophagus and the swimbladder by which they are believed to inflate their swimbladders by ''gulping'' atmospheric air at the sea surface. Acoustic measurements at 1.5–5 kHz on fish in the Gulf of Maine showed a swimbladder resonance peak near 2.3 kHz at 180 m depth. Midwater trawls confirmed that the fish were Atlantic herring (*Clupea harengus*) of 19–28 cm length. Calculations using a model of swimbladder resonance gives swimbladder volumes of 1.3% at 180 m. Extrapolation using Boyle's law suggests that at the sea surface, these herring would need to inflate their swimbladders by up to four times the volume required for neutral buoyancy. In general, swimbladders of the Clupeidae may resonate at lower frequencies than previously expected. [Work supported by ONR and NMFS.]

4:55

1pAO13. Fine-scale diel migration dynamics of an island-associated sound-scattering layer. Kelly Benoit-Bird and Whitlow Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734)

The Hawaiian mesopelagic boundary community, an island-associated, midwater sound-scattering layer, undergoes diel vertical and horizontal migrations. To understand the dynamics of the community's migration at small temporal scales and large spatial scales, five bottom-mounted, 200-kHz active-acoustic mooring that transmitted ten signals every 15 min, from dusk until dawn for 5 days. Two layers within the boundary community were observed to undergo simultaneous diel vertical and horizontal migration. Vertical migration rates were measured at 0–1.7 m/min, while the horizontal rate averaged 1.67 km/h, swamping the vertical movement. The vertex of the migration pattern was observed 45 min before the midpoint between sunset and sunrise. Until the vertex, animal density increased relatively constantly as the animals migrated towards shore, with the highest animal densities found in the shallowest areas at midnight. Animal abundance estimates at the leading and trailing edge of the layer support the hypothesis that increased animal densities near shore are related to packing, as mesopelagic animals avoid the surface and the bottom. We observed high levels of biomass moving rapidly, over a great distance, into shallow waters very close to shore, providing insight into the significant link the mesopelagic boundary community provides between nearshore and oceanic systems.

5:10

1pAO14. Tidal matched field processing inversion for water depth and source range in the Intimate96 test. A. Tolstoy (ATolstoy Sci., 8610 Battailles Court, Annandale, VA 22003, atolstoy@ieee.org), S. Jesus, and O. Rodríguez (Univ. of Algarve, Faro, Portugal)

Examining Intimate96 hydrophone data (300 to 800 Hz) we see clearly the effects of tidal changes, i.e., of changing water depths. In this work we will examine Matched Field Processing (MFP) sensitivity at that range of frequencies to expected tidal changes (the depth varies ± 1.0 m from the nominal of 135 m). Is it possible to invert such data to accurately and uniquely estimate water depth D as a function of time (tides)? What about accurate, unique, *simultaneous* estimates of source range r_{sou}? What happens when we use multiple frequencies and when D is known to shift in a predictable fashion? Can the r_{sou} vs D ambiguity ever be resolved for a successful, unique MFP inversion for those parameters?

5:25

1pAO15. Influence of fetch limited surface roughness on mid-to-high frequency acoustic propagation in shallow water. Robert Heitsenrether, Mohsen Badiey (Ocean Acoust. Lab., College of Marine Studies, Univ. of Delaware, Newark, DE 19716, rheits@udel.edu), James Kirby (Ctr. for Appl. Coastal Res., Univ. of Delaware, Newark, DE 19716), and Steve Forsythe (Naval Undersea Warfare Ctr., Newport, RI 0234)

Surface waves are among several environmental parameters that can influence broadband mid-to-high frequency (1–18 kHz) acoustic wave propagation. Understanding the interaction of sound waves at a rough surface requires a detailed description of the ocean wave spectrum. In shallow water regions, due to proximity to land, surface waves are usually fetch and duration limited with reduced spectral level and higher frequency components. A model that approximates the wave spectra for a fetch limited sea has been combined with acoustic ray-based method for analysis of forward scattered acoustic signals in such a coastal environment. Numerical modeling is employed to investigate time-frequency-angle characteristics of ray paths reflected from a rough sea surface. Temporal variability of acoustic signal fluctuations has been examined as a function of varying sea surface. To validate this model, results are compared against a unique set of experimental data collected in a fetch limited region. The experimental design allowed an examination of time evolution of a single surface bounced ray path. Simultaneous wind speed and acoustic propagation measurements allowed correlation between the individual ray paths and the sea surface at varying sea state conditions. For low wind speed conditions, model results predict the temporal fluctuations of the measured acoustic signal propagation.

MONDAY AFTERNOON, 2 DECEMBER 2002 CORAL GARDEN 1, 1:00 TO 3:55 P.M.

Session 1pBB

Biomedical Ultrasound/Bioresponse to Vibration and Signal Processing in Acoustics: Acoustic Microscopy

Joie P. Jones, Chair
Radiological Sciences, University of California, Irvine, California 92697-5000

Invited Papers

1:00

1pBB1. A short history of acoustical microscopy. Joie Jones (Dept. of Radiol., Univ. of California, Irvine, Irvine, CA 92697-5000, jpjones@uci.edu)

Optical microscopy has a long and interesting history, going back thousands of years to discoveries made in both Assyrian and Mayan cultures. Acoustical microscopy, on the other hand, has had a much shorter but equally interesting history, going back only to the mid-20th century. This presentation traces the development of acoustical microscopy from its very beginnings to the present. Comparisons with other microscopic techniques will point out the unique features offered by acoustical microscopy. A wide range of application areas will be reviewed and future prospects and potentials discussed.

1:30

1pBB2. Scanning tomographic acoustic microscopy. Hua Lee (Dept. of Elec. & Computer Eng., Univ. of California, Santa Barbara, Santa Barbara, CA 93106, hualee@ece.ucsb.edu)

This paper provides an overview of the design and development of the scanning tomographic acoustic microscopy (STAM). This research effort spans over a period of more than 12 years, which successfully elevated the acoustic microscopy from the traditional intensity-mapping mode to the level of holographic and tomographic imaging. The tomographic imaging capability of STAM was developed on the platform of the scanning laser acoustic microscope (SLAM), which operates in a coherent transmission mode with plane-wave illumination and scanning laser wavefield detection. The image formation techniques were based on the backward propagation method implemented in the plane-to-plane format. In this paper, the key elements of the design and development, including the modification of the data-acquisition hardware, implementation of image reconstruction algorithms for multiple-frequency and multiple-angle tomography, and the high-precision phase-correction and image registration techniques for the superposition of coherent sub-images, will be discussed. Results of full-scale experiments will also be included to demonstrate the capability of holographic and tomographic image formation in microscopic scale.

2:00

1pBB3. Intravascular ultrasound: From the acoustic microscope into the catheterization laboratory. Antonius F. W. van der Steen, Chris L. de Korte, and Cornelis J. Slager (Thorax Ctr., Erasmus Medical Ctr., Ee 23.02, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, vandersteen@tch.fgg.eur.nl)

One of the explicit goals of acoustic microscopy (AM) of biological tissues is to gain information and insights *in vitro* that will improve clinically used diagnostic ultrasound. This lecture will focus on the transfer of knowledge on vascular ultrasound obtained from acoustic microcopy to interventional cardiology. Spectral AM has shown that calcified atherosclerosis can be discriminated from fibrous and fatty plaques. The discrimination between fibrous and fatty plaques is less pronounced, but AM has shown these have distinct different anisotropy properties. Furthermore, it has been shown *in vitro* that fibrous tissues have much higher Youngs moduli than fatty tissues. AM has also shown that endothelial cells are sensitive to shear stress. In the catheterization laboratory this knowledge plays an important role in vulnerable plaque detection. A vulnerable plaque is a nonobstructing plaque consisting of a fatty pool covered by a thin fibrous cap. Breaking of the cap annually kills 225 000 people in the USA. A combination of biplane angiography, 3D IVUS and computational fluid dynamics allows one to determine shear stress at the endothelium. This has proven to be a predictor for the long term outcome of interventions such as balloon angioplasty and (drug eluting) stent placement.

2:30–2:45 Break

2:45

1pBB4. Application of acoustic microscopy to assessment of cardiovascular biomechanics. Yoshifumi Saijo, Hidehiko Sasaki, Shin-ichi Nitta, Motonao Tanaka (Dept. of Medical Eng. & Cardiology, Inst. of Development, Aging and Cancer, Tohoku Univ., 4-1 Seiryomachi, Aoba-ku, Sendai 980-8575, Japan), Claus S. Joergensen, and Erling Falk (Aarhus Univ., 8000 Aarhus, Denmark)

Acoustic microscopy provides information on physical and mechanical properties of biological tissues, while optical microscopy with various staining techniques provides chemical properties. The biomechanics of tissues is especially important in cardiovascular system because its pathophysiology is closely related with mechanical stresses such as blood pressure or blood flow. A scanning acoustic microscope (SAM) system with tone-burst ultrasound in the frequency range of 100–200 MHz has been developed, and attenuation and sound speed of tissues have been measured. In human coronary arteries, attenuation and sound speed were high in calcification and collagen, while both values were low in smooth muscle and lipid. Another SAM system with 800-MHz–1.3-GHz ultrasound was applied for aortas of Apo-E deficient mouse, which is known to develop atherosclerosis. Attenuation of ultrasound was significantly higher in type 1 collagen compared to type 3 collagen. Recently, a new type FFT-SAM using a single-pulse, broadband frequency range ultrasound (20–150 MHz) has been developed. Cardiac allograft was observed by FFT-SAM and the acoustic properties were able to grade allograft rejection. SAM provides very useful information for assessing cardiovascular biomechanics and for understanding normal and abnormal images of clinical ultrasound.

3:15

1pBB5. Comparative study of optical detection for scanning laser acoustic microscopy. Moises Cywiak (Centro de Investigaciones en Optica, A.C., Apdo, Postal 1-948, 37000 Leon, GTO, Mexico) and Glen Wade (Univ. of California, Santa Barbara, CA 93106, wade@ece.ucsb.edu)

Scanning laser acoustic microscopy (SLAM) provides a way of visualizing and measuring surface features of nanometric order by a combination of acoustical and optical means. The acousto-optical interface in SLAM is an acoustically transparent and optically reflecting surface embedded in the acoustic medium. Acoustical excitation causes a dynamic surface deformation which is picked up by a focused scanning laser beam which, after reflection, is modulated in its phase and angle of reflection. There are three basically different demodulation schemes reported in the literature: the knife-edge detector, the time-delay interferometric detector, and the reference-beam detector. (In addition, a new detector derived from the knife-edge detector, namely the pyramidal-mirror detector, is scheduled for separate treatment in this session.) This paper presents comparative studies of the above three basic methods of detection. The performance characteristics of each are analyzed in detail by a careful examination of the transfer function for each of the systems. Computer simulations for their operation provide a comprehensive picture of the strengths and weaknesses of each.

3:35

1pBB6. Increasing image resolution using the pyramidal detector. Steve Isakson (Isakson Engineering, 8355 San Rafael Rd., Atascadero, CA 93422), Luis Roberto Sahagun (Centro de Investigaciones en Optica, A.C., Apdo. Postal I-948, 37000 Leon, GTO, Mexico), and Glen Wade (Univ. of California, Santa Barbara, CA 93106)

The knife-edge detector (KED) has been used in transverse wave, scanning laser acoustic microscopes (SLAM) successfully for many years. It possesses a good balance between spatial bandwidth, detector noise insensitivity, and detector grain stability. Many detectors have tried to increase the bandwidth to improve the image resolution, only to be difficult to use and susceptible to vibration and mechanical variations. The authors have invented and designed a pyramidal detector (PD) that combines the stability and noise insensitivity of the KED with a significantly improved bandwidth in both axes of the object plane. The detector uses the entire signal provided at the detector to increase the signal to noise ratio and significantly decrease the cover slip spatial frequencies that are difficult to detect. The increased bandwidth allows image resolution approaching what an ideal detector could achieve. Additionally, unlike the KED, the PD spatial bandwidth is approximately isotropic to waves traveling in the cover slip. Increased spatial resolution and uniformity result. This property is particularly useful for back-propagation in tomographic applications.

MONDAY AFTERNOON, 2 DECEMBER 2002 CORAL SEA 1 AND 2, 1:00 TO 3:15 P.M.

Session 1pMU

Musical Acoustics: Musical Instrument Acoustics

Uwe J. Hansen, Chair
Department of Physics, Indiana State University, Terre Haute, Indiana 47809

Contributed Papers

1:00

1pMU1. Shape optimization techniques for musical instrument design. Luis Henrique (Instituto Politecnico do Porto, Escola Superior de Musica e Artes do Espectaculo, R. da Alegria, 503, 4000-046 Porto, Portugal), Jose Antunes (Instituto Tecnologico e Nuclear, 2686 Sacavem Codex, Portugal), and Joao S. Carvalho (Universidade Nova de Lisboa, 1050-069 Lisboa, Portugal)

The design of musical instruments is still mostly based on empirical knowledge and costly experimentation. One interesting improvement is the shape optimization of resonating components, given a number of constraints (allowed parameter ranges, shape smoothness, etc.), so that vibrations occur at specified modal frequencies. Each admissible geometrical configuration generates an error between computed eigenfrequencies and the target set. Typically, error surfaces present many local minima, corresponding to suboptimal designs. This difficulty can be overcome using global optimization techniques, such as simulated annealing. However these methods are greedy, concerning the number of function evaluations required. Thus, the computational effort can be unacceptable if complex problems, such as bell optimization, are tackled. Those issues are addressed in this paper, and a method for improving optimization procedures is proposed. Instead of using the local geometric parameters as searched variables, the system geometry is modeled in terms of truncated series of orthogonal space-funcitons, and optimization is performed on their amplitude coefficients. Fourier series and orthogonal polynomials are typical such functions. This technique reduces considerably the number of searched variables, and has a potential for significant computational savings in complex problems. It is illustrated by optimizing the shapes of both current and uncommon marimba bars.

1:15

1pMU2. The acoustical Klein–Gordon equation: A time-independent perturbation analysis. Barbara J. Forbes (Dept. of Environ. and Mech. Eng., The Open Univ., Walton Hall, Milton Keynes MK7 6AA, UK)

The acoustical Klein–Gordon equation, describing one-dimensional wave propagation in a duct of varying cross-section, is discussed. Dispersive solutions, not elucidated by the Webster equation, are examined and perturbative geometric ''potentials'' are defined. A time-independent, first-order expression for the perturbed eigenfunctions—defined as the potential energy per unit length of duct—is derived. Eigenvalue shifts are defined in terms of a sinusoidal perturbative term, and Ehrenfest's theorem is used to obtain the first-order shifts in the duct resonances. It is found that the perturbation may be in or out of phase with the radiation pressure, thus strengthening or weakening the resonances, respectively; and that the perturbation may raise, lower, or have no effect on the resonance frequencies, depending on the interaction of the phases of the terms. The results are compared with the standard theory of pipe resonances, which assumes perturbations to the cross-section of the pipe rather than directly to the eigenvalues.

1:30

1pMU3. Holographic interferometry and finite element modeling of handbells. John R. Buschert (Goshen College, Goshen, IN 46526, johnrb@goshen.edu)

Holographic interferometry was used to map out the vibration modes of a C4 handbell. A finite element model (FEM) of the handbell was also created. The FEM program can be made to present its results in a graphic format which mimics the stripes seen in interferometry. Thus, the FEM model can easily be compared in detail to the holograms. Excellent agreement between the two has been obtained, which is evidence for the validity of the finite element model. The FEM has also been used to explore the effects of changing the shape of the bell. The effects on the overtone structure of changing the bell in some unusual ways will be presented.

1:45

1pMU4. Transverse and torsional modes of vibration of American organ reeds. Blake M. Dirksen (Loras College, Dubuque, IA 52004, dirkbm@loras.edu) and James P. Cottingham (Coe College, Cedar Rapids, IA 52402)

A reed from an American reed organ consists of a brass tongue riveted to a frame with an opening just large enough for the reed to pass through during the course of its oscillation. To a first approximation, the reed tongue can be modeled as a cantilever beam of uniform cross section, and the mode frequencies of transverse and torsional modes can be calculated. The reeds used in this study differ from this simple model in several ways, including a nonuniform thickness along the reed length. In addition, the

reed includes a curved spoon-like section at the tip of the reed tongue as well as an asymmetric twist at the tip. Transverse and torsional modes of vibration of a mechanically excited reed tongue have been studied using a proximity sensor (variable impedance transducer), a laser vibrometer, and a TV holography system. The motion of the air-driven reed has also been investigated. Although the motion of air-driven reed is dominated by the first transverse mode, some evidence of higher-order transverse modes and torsional modes has been observed. [Work supported by the NSF from REU Grant No. 0139096.]

2:00

1pMU5. The influence of reed curvature on the tone quality of lingual organ pipes. G. R. Plitnik (Dept. of Phys., Frostburg State Univ., Frostburg, MD 21532, gplitnik@frostburg.edu) and J. Angster (Fraunhofer Institut fur Bauphysik, Stuttgart, Germany)

Given certain design constraints, such as the type of stop being voiced and the desired tone quality, reed voicers must use consummate skill to curve each tongue so as to produce the best and most stable tone, as well as maintaining a consistent tone quality across an entire rank of pipes. The curvature given to a reed tongue influences not only the harmonic structure of the steady-state sound, but also the attack. Two fundamentally different types of curvature are typically employed, the chorus reed (trompette) curve (which yields a bright sound) and the smooth-toned curve employed for solo reeds such as the clarinet. This study investigated the effect of reed curvature on the vibration and tone of reed tongues of both types. Two *F*2 pipes (a trompette and a clarinet) were constructed and voiced with 6 different tongues each to produce a variety of tones. The reed's vibration was measured under typical conditions by laser vibrometer; the pressure waves in the boot and in the shallot were measured by means of one-eighth inch microphones and the emitted sound was recorded at the egress. By performing various measurements simultaneously, phase differences were also determined.

2:15

1pMU6. The acoustics of the echo cornet. Robert W. Pyle, Jr. (11 Holworthy Pl., Cambridge, MA 02138, rpyle@post.harvard.edu) and Sabine K. Klaus (America's Shrine to Music Museum, Landrum, SC 29356)

The echo cornet was an instrument produced by a number of makers in several countries from about the middle of the nineteenth to the early twentieth centuries. It consists of an ordinary three-valve cornet to which a fourth valve has been added, downstream of the three normal valves. The extra valve diverts the airstream from the normal bell to an ''echo'' bell that gives a muted tone quality. Although the air column through the echo bell is typically 15 cm longer than the path through the normal bell, there is no appreciable change of playing pitch when the echo bell is in use. Acoustic input impedance and impulse response measurements and consideration of the standing-wave pattern within the echo bell show how this can be so. Acoustically, the echo bell is more closely related to hand-stopping on the French horn than to the mutes commonly used on the trumpet and cornet.

2:30

1pMU7. Significance of torsion modes in bowed-string dynamics. Octavio Inacio (Instituto Politecnico do Porto, Escola Superior de Musica e Artes do Espectaculo, R. Alegria, 503, 4000-046 Porto, Portugal), Jose Antunes (Instituto Tecnologico e Nuclear, 2686 Sacavem Codex, Portugal), and Luis Henrique (Instituto Politecnico do Porto, 4000-046 Porto, Portugal)

Several aspects of bowed-string dynamics are still inadequately clarified. The importance of torsion modes on the motion regimes is one such issue. Experiments involving torsion are difficult and most of the results available pertain to numerical simulations. The authors' approach differs from previous efforts in two main aspects: (1) the development of a computational method distinct from the wave-propagation approach pioneered by McIntyre, Schumacher, and Woodhouse and (2) an extensive and systematic analysis of the coupling between torsion and transverse motions is performed. The numerical simulations are based on a modal representation of the unconstrained string and a computational approach for friction that enables accurate representations of the stick-slip forces and of the string dynamics, in both time and space. Many relevant aspects of the bowed-string can be readily implemented, including string inharmonic behavior, finite bow-width, and torsion effects. Concerning the later aspect, a realistic range of the torsional to transverse wave-speed ratio is investigated, for several values of the bow velocity and normal force. Results suggest that torsion modes can effect both transient durations and steady state regimes, in particular when the above-mentioned ratio is <4. Gut strings should then be particularly prone to torsion effects.

2:45

1pMU8. The acoustics of the bagana. Stephanie Weisser (Aspirant FNRS-Laboratoire de Phonolgie, Univ. Libre de Bruxelles, Ave. F. Roosevelt, 50 CP 175, B 1050, Brussels, Belgium, stephanieweisser@swing) and Didier Demolin (Univ. Libre de Bruxelles, Brussels, Belgium)

The bagana is a big Ethiopian lyre with ten strings. The instrument is found in the area of the Amhara, the culturally dominant tribe of Ethiopia. It is an intimate instrument, played only with the voice, for prayer and meditation. It can be tuned in two pentatonic scales, and can be plucked with the fingers, string by string or with a plectrum, all strings together. The box of the bagana is made of wood covered with leather, and the strings are made of gut. They are very thick and between them and the bridge, there are ten buzzers. Therefore, the bagana produces a very deep and buzzing sound. The paper will analyze the acoustical function of several parts of this instrument, e.g., the sound box and its crosslike hole at the back. The influence of the buzzers on the fundamental frequency, the spectrum, the duration, and the intensity of the signal will be closely examined. The role of the five unused strings (the ''rest'' strings) will also be examined.

3:00

1pMU9. Acoustic characteristics of Ekonda scrapers. Didier Demolin and Stephanie Weisser (Free Univ. of Brussels, 50 av. F. D. Roosevelt, 1050 Brussels, Belgium)

This paper describes some acoustic characteristics of a musical instrument played with the complex vocal polyphonies of the Ekonda of the Congo. The complexity of these songs and the remarkable pitch accuracy of their intonation are striking given that there is no tuning fork or fixed pitch instrument which can account for the pitch stability. Ekonda scrapers are required to accompany the polyphonies. An acoustic study of these idiophones was made to understand their contributions to the songs. The scrapers are made in a hollow piece of a palm tree branch. The instrument is played by rubbing a piece of hard wood upon splits made on the piece of palm tree. In addition to the friction noise due to the rubbing of the piece of hard wood, there are some very distinct resonances which can be identified. An important observation is that there is always a very intense resonance at the frequency of the musical scales tonic. The relation between the pitch of the tonic and the frequency of the scrapers resonances which act as a tuning fork are detailed.

Session 1pNS

Noise: Environmental Noise Planning and Legislation

Lawrence S. Finegold, Cochair
Finegold & So, 1167 Bournemouth Court, Centerville, Ohio 45459-2647

Eugenio Collados, Cochair
Depto. de Fisica, Universidad de Santiago de Chile, Santiago, Chile

Chair's Introduction—1:00

Invited Paper

1:05

1pNS1. Development of a model local noise ordinance standard. Lawrence S. Finegold (Finegold & So, Community & Environ. Noise Consultants, 1167 Bournemouth Ct., Centerville, OH 45459, lsfinegold@earthlink.net) and Bennett M. Brooks (Brooks Acoust. Corp., Vernon, CT 06066)

A comprehensive, easy to understand and enforceable local noise ordinance can be a powerful tool in the effort to manage the noise environment at the community level. At the present time, however, many communities do not have a local noise ordinance, and those that do exist can vary considerably. In order to provide guidance to government officials involved in the development of noise ordinances, the American National Standards Institute (ANSI) Standards Committee on Noise (S12), Working Group 41 is developing a Model Local Noise Ordinance. Although this effort is still in the early conceptual stages, this paper will describe some of the issues and approaches that are being considered in the development of this document. It is envisioned that Part 1 of this possible new ANSI Standard will address general community noise control issues that should be considered in the development of a noise ordinance and Part 2 will provide a sample noise ordinance with a menu of standardized language sections that can be selected to fit individual community situations. Ranges of recommended noise exposure criteria will be provided, along with a brief discussion of issues that local officials should consider in choosing these criteria.

Contributed Papers

1:30

1pNS2. Railroads, noise and the city: A manual for authorities. Fernando J. Elizondo-Garza, Cesar A. Leal-Chapa, Miguel Cupich-Rodriguez, and Jorge Cupich-Guerrero (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)

Trains crossing cities produce different problems and annoyances; one of them is noise, which directly affects the neighborhoods near the railroad tracks. This situation must be solved jointly by the city authorities and the railroad companies in two ways: planning, which looks to avoid future problems, and short-term solutions, which have the objective to control the present problems. In this paper a manual for authorities is presented that gives an overview of the available options that help in the decision-making process necessary to achieve a problem-free interaction between railroads and cities in relation to noise.

1:45

1pNS3. Conflicting noise ordinances in a metropolitan city. Allan K. Kennedy (Environ. Health and Safety Services, Black & Veatch Corp., 11401 Lamar Ave., Overland Park, KS 66211)

Controlling community noise has become an important aspect in most communities. The desire to provide an environment that is acceptable and peaceful to residents and conducive to businesses sometimes creates conflicts in the regulatory requirements. This conflict is compounded in metropolitan cities where many communities are adjacent. Each community has a set of priorities for its citizens directing how ordinances are written and what sound limits are permitted on certain land uses. Noise ordinances in a United States metropolitan area with a population of 1.5 million were reviewed for their similarity and dissimilarity with respect to zoning. It was discovered that even though the communities share the same borders, they have vastly different noise limits in terms of amplitude, time of day, required parameter (Leq, Ldn, etc.), and land use. The conflicts arise for noise sources that are located at the borders of these cities.

2:00

1pNS4. The potential application of auditory neuroscience to improving the quality of environmental noise standards. George A. Luz (U.S. Army Ctr. for Health Promotion and Preventive Medicine, 5158 Blackhawk Rd., Aberdeen Proving Ground, MD 21010) and John G. Neuhoff (The College of Wooster, Wooster, OH 44691)

In the Public Policy for Noise session at ASA's last meeting, Fidell stated that U.S. environmental noise regulations are hampered by (1) treating people as if they had a sound level meter in their brain and (2) over-reliance on the Schultz curve. This paper looks at four stages in replacing the SLM in the brain: (1) Pre-1980 when knowledge of psychoacoustics exceeded the capability of instruments (Case study: U.S. Environmental Protection Agency's 1974 decision to use A-weighting instead of a loudness-based measure). (2) The 1980s when instrumentation could imitate psychoacoustics (Case study: Development of the loudness meter). (3) The 1990s when psychoacoustic research had relatively little influence on American noise standards (Case study: The decision of ANSI S12.9 to reject a Japanese proposal for assessing the annoyance of impulsive sounds [Ogura *et al.*, J. Noise Control Eng. **40**, 231–240 (1993)] in favor of a subjective ''impulsive'' and ''highly impulsive'' dichotomy). (4) Present opportunities for creative teamwork between equipment developers, standards writers, psychoacousticians, and neuroscientists (Case study: Explaining the U.S. Air Force's correction to the Sound Exposure Level for the added annoyance of low level, fast moving jet aircraft by ''auditory looming'').

2:15

1pNS5. Comparison between procedures of previsions in sound propagation in a quarry. Emanuele Sarri and Guido Alfaro Degan (Dept. of Mech. Eng., Roma Tre Univ., 79 Via Della Vasca Navale, Roma 00146, Italy)

The aim of this work is to compare the accuracy of numerical previsions of sound propagation in a quarry obtained with different algorithms. A comparison is made between the ISO 9613-1/2:1993 (E) standard for outdoor sound propagation, a new simplified algorithm based on the ISO code, a pyramid beam tracing method, and experimental measurements. The ISO 9613 standard proposes a detailed and complex method taking into account air absorption. It also takes into account the effects caused by the propagation over soil with varying properties, shielding both from thin and thick obstacles as well as the effects of vegetation layers and excess attenuation. The new algorithm simplifies ISO codes proposing an easy method, specific for a quarry. Pyramid beam tracing, implemented in Ramsete sound propagation software, is typically used to simulate sound fields in large rooms and also outdoors. Experimental and simulated data were used to built graphic plots, enabling a direct comparison of the results.

2:30

1pNS6. Assessment of adequacy of ray acoustics approach for prediction of barrier insertion loss in the presence of a reflecting ground. Daniel F. P. Pazos, Ricardo E. Musafir (Dept. of Mech. Eng./ COPPE, Universidade Federal do Rio de Janeiro C.P. 68503, 21945-970 Rio de Janeiro, Brazil), and Eldad J. Avital (Univ. of London, London E1 4NS, UK)

The precision of insertion loss prediction obtained by applying Kurze–Anderson formula to each of the four possible ray paths connecting source and observer in the presence of a reflecting ground, and computing interference appropriately, is discussed. To this end, the field behind the barrier, given by the method described, is compared to that obtained by solving numerically the continuity and momentum equations simultaneously, using the appropriate boundary conditions to account for the perfectly reflecting ground and the barrier. Analysis of the agreement of the interference patterns obtained in both cases, for pure tones and for octave bands, permits determining the range in which the much simpler ray method provides sufficient results. Calculations with the Kurze–Anderson formula are performed with Mathematica 4.0, while the numerical propagation code is written in Fortran.

2:45

1pNS7. Investigation of the acoustic parameters that influence traffic noise. Alejandro A. Armas and Federico M. Iasi (Laboratorio de Acustica y Luminotecnia, Cmno. Centenario y 506-Gonnet (CP1897), Buenos Aires, Argentina)

Here, we analyze the behavior of the acoustic parameters L_{eq}, L1, L10, L50, L90, L99, L_{min}, and L_{max} of vehicular noise, which were measured in controlled laboratory conditions, for various traffic noise simulations. For that, sound recordings were created of equal duration but different composition, based on real recordings of the passage of different types of vehicles (cars, motorbikes, trucks, and buses) and using pink noise as background noise. This study is based on the necessity of finding the most appropriate parameters for the characterization of the traffic noise in the cities. The study began with the test of a measurement methodology that uses the equivalent continuous sound level, the main acoustical parameter applied in the city of La Plata, Argentina. However, L_{eq} was shown insufficient and inadequate in certain situations, especially those that didn't exhibit intense traffic noise, as found in certain residential areas. (To be presented in Spanish.)

3:00–3:15 Break

3:15

1pNS8. Measurement protocol for normalized noise maps. Gabriel Viro, Francisco Ruffa, Daniel Gavinowich, Oscar Bonello (Laboratorio de Acustica y Electroacustica, FIUBA, Paseo Colon 850, Buenos Aires, Argentina, laceac_uba@yahoo.com), Pablo Ciccarella, and Erica Barkasz (FIUBA, Paseo Colon 850, Buenos Aires, Argentina)

As cities grow bigger, traffic noise gets worse. In order to help in the determination of noisy areas, the LACEAC prepared a measurement protocol to aid in this task. This protocol is the base for a recommendation to the cities' governments. To put together this protocol, the LACEAC has analyzed different legislation together with measurements taken in several pilot areas of the city of Buenos Aires and determined the best approach based on the cultural habits of the city inhabitants. This protocol thus recommends the way that measurement locations should be determined, exposure time, the way to take the measurements, to study the results and to draw the noise maps.

3:30

1pNS9. Spectral information in noise-mapping: An exploratory study. Vivian Pasch, Patricia Mosconi, Marta Yanitelli, Susana Cabanellas, Jorge Vazquez, Juan C. Rall (Grupo Ruido, Fac. de Arquitectura, Natl. Univ. of Rosario, Riobamba 220 bis, 2000 Rosario, Argentina), and Federico Miyara (Natl. Univ. of Rosario, 2000 Rosario, Argentina, fmiyara@fceia.unr.edu.ar)

International standards such as ISO 1996 and ISO 717 as well as noise regulations in several countries are increasingly relying on spectral information in order to assess the acoustical behavior of materials and structures and the effects of noise on people. Nevertheless, the new European Union Directive on the assessment and management of environmental noise reinforces the A-weighted equivalent level (with appropriate night and evening corrections) as the preferred indicator for noise mapping. Considering that noise maps are a powerful zoning and planning resource, the idea of reporting the mean spectrum of noise at each selected location at different times is proposed and thoroughly justified. Arguments in favor of its feasibility are given, showing that, in spite of the widespread opinion, costs and required time may be reduced considerably by the use of low-priced, new-technology auxiliary equipment. Then an exploratory study is reported, in which (a) the spectrum of traffic noise in Rosario (Argentina) is compared with the internationally standardized traffic noise spectrum, and (b) the noise spectrum at an open street is compared with the noise spectrum at a street with a U-profile owing to the same vehicles.

3:45

1pNS10. Evaluation of urban railways noise impact. Massimo Coppi and Stefano Grignaffini (Dept. of "Fisica Tecnica," Univ. of Rome "La Sapienza"—Via Eudossiana, 18-00184 Rome, Italy)

The study of the environmental impact of urban railways mainly regards the noise generated by trains. Italian standards provide for the characterization of territorial pertinence zones (from the outer center line and for each side of the railway tracks), inside which specific absolute limit values of noise introduction produced by the same infrastructure are permitted. Therefore, according to provisions in these areas, the noise evaluation has to be performed referring only to the infrastructure contribution, without evaluating other acoustic sources in the same areas. As a consequence, in the pertinence zones the imposed limits for the infrastructure presence and the limits imposed by the acoustic municipal zoning, taking into account the acoustic impact of the other sources, are valid simultaneously but separately when evaluating the acoustic impact. However, in these areas the general acoustic climate is due to the overlap of both acoustic classes, and the noise induced on the population is determined by the contemporary presence of both sources. It is eventually very important to evaluate the effective noise dose absorbed by the population, especially when having to design mitigation interventions.

4:00

1pNS11. Study of noise external to public hospitals in the metropolitan area of Guadalajara. Martha Orozco, Javier Garcia, Gabriela Hernandez (IMACH, Departamento de Ciencias Ambientales, CUCBA, Universidad de Guadalajara, Km 15 1/2 Carr. a Nogales, Las Agujas Zapopan, Jalisco, Mexico), Arturo Figueroa, Jorge Macias (Direccion de Ecologia, H. Ayuntamiento de Guadalajara, Guadalajara, Jalisco, Mexico), Alan Delgadillo (Secretaria de Educacion, Guadalajara, Jalisco, Mexico), Esmeralda Vazquez, and Maria del Carmen Rivera (Universidad de Guadalajara, Las Agugas, Zapopan, Jalisco)

Environmental noise is at the present time a cause of concern, for the serious nuisance that it causes, and also the effects on health and behavior of individuals, on human activities, as well as for the psychological and social consequences. Hospital centers are places where numerous factors of quality and degradation can affect the recovery of patients. Knowing the acoustical conditions in some of the main hospitals of the city will make it possible to suggest measures to assist the present problem. We sampled 7 hospitals. The measurement points were selected at corners and at representative points or areas of pedestrian and vehicular entrance, a minimum of 4 points per block. We detected critical points of environmental noise in the vicinity of the hospitals, that require urgent attention. Some of the critical levels of noise were for LEQ, 93.4 and averages of this around 84.63 and maximum levels of 114.2, caused by an ambulance siren.

4:15

1pNS12. Effects of meteorology on predicted levels of aircraft noise. Kenneth Plotkin, Bruce Ikelheimer (Wyle Laboratories, 2001 Jefferson Davis Hwy., Ste. 701, Arlington, VA 22202, kplotkin@wylelabs.com), Jerome Huber, and John-Paul Clarke (MIT, Cambridge, MA 02139)

Calculation of average noise (e.g., Ldn) around airports generally ignores the effects of atmospheric gradients as these effects are assumed to average out over the course of a year. The validity of this assumption was tested through a numerical study using the aircraft noise simulation model NMSIM. Specifically, NMSIM was extended to account for three-dimensional refraction by wind and temperature gradients, plus refraction and scattering into shadow zones, as well as ground impedance effect. Five years of meteorological data at a number of airport locations was obtained, and divided into hourly conditions. A nominal airport, based on a major air carrier airport, was defined. NMSIM was run for hourly operations and weather conditions at each airport location. The resultant Ldn contours were compared to each other by season and location, and also to contours computed for the baseline assumption of no gradients. [Work supported by NASA Langley Research Center.]

4:30

1pNS13. Military aircraft noise and nonlinear acoustics. Victor W. Sparrow, Kent L. Gee (Grad. Prog. in Acoust., Penn State, 316B Leonhard Bldg., University Park, PA 16802, vws1@psu.edu), J. Micah Downing, and Kenneth J. Plotkin (Wyle Laboratories, Arlington, VA 22202)

A new joint program between Wyle Laboratories, Penn State, and others has been initiated to develop a better noise model for military fighter aircraft. As is well known, the high sound amplitudes produced by modern jets imply the importance of nonlinear effects in the sound propagation. In the literature one can find several experiments clearly showing spectral broadening and the development and coalescence of shocks as the noise propagates with distance. Current U.S. DoD noise models, nevertheless, neglect nonlinear distortion. A number of different nonlinear propagation predictions schemes are available, and currently there is no broad consensus as to which models work best for jet noise. This talk will discuss the applicability of several of these methods to the current need. Only when nonlinear propagation methods have been coded and tested against laboratory and full scale measurements will a predictive tool be available for assessing the environmental noise impacts of modern aircraft. [Work supported by SERDP.]

MONDAY AFTERNOON, 2 DECEMBER 2002 CORAL ISLAND 1 AND 2, 1:00 TO 4:30 P.M.

Session 1pPA

Physical Acoustics: Atmospheric Acoustics and Impedance Surfaces

Craig J. Hickey, Cochair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Rodolfo Martinez, Cochair
CIIDIR, IPN, Oaxaca, Mexico

Contributed Papers

1:00

1pPA1. Effects of nonlinearity on the propagation of acoustic pulses in random media. Robin Cleveland (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215, robinc@bu.edu), Laurent Dallois (Oxford Univ., Oxford OX1 3LB, UK), and Philippe Blanc-Benon (LMFA UMR CNRS 5509, 69131 Ecully Cedex, France)

We conducted a numerical investigation into the propagation of finite-amplitude pulses in media with inhomogeneous random sound speed. An *N* wave (idealized sonic boom) was used as the pulse shape. Initial simulations considered a medium with a single spherical scattering object with a slow sound speed. This object acted as a focusing lens. As the amplitude of the *N* wave was increased nonlinear effects initially led to enhancement of focusing, reduction in shock risetime, and a shift of the peak away from the object. However, for high amplitude, energy loss at the shock led to a dramatic reduction in the amplitude of the focus and a shift towards the object. Simulations were then carried out in a two-dimensional random media. The sound speed in the random media was constructed using a Fourier mode decomposition with parameters appropriate for turbulence in the atmospheric boundary layer. For low amplitude waves the *N* wave was focused and defocused by regions of low and high sound speed, respectively. However, the presence of multiple paths means that the wave form no longer resembled an *N*-wave after propagating about 10 wavelengths. As the amplitude was increased the focusing was enhanced and more localized.

1:15

1pPA2. Effect of turbulence on the risetime of sonic booms and spark generated *N* waves. Bart Lipkens (Mech. Eng. Dept., Western New England College, 1215 Wilbraham Rd., Springfield, MA 01119) and Philippe Blanc-Benon (LMFA UMR CNRS 5509, 69134 Ecully Cedex, France)

The risetime of acoustic pulses, such as sonic booms and electric spark generated *N* waves, is affected by propagation through turbulence. Risetime increases with increasing propagation distance and turbulence intensity. Data from electric spark propagation through turbulence show that Pierce's wave front folding theory (1971) correctly describes the dependence of risetime τ on propagation distance x, i.e., $\tau \propto x^{11/7}$, and turbulence kinetic energy dissipation rate ϵ, i.e., $\tau \propto \epsilon^{4/7}$. More recently, Pierce (AIAA-95-105) also derived a dispersion relation for acoustic pulse propagation through turbulence. The dispersion relation leads to an extra term in the propagation equation and takes into account the averaged contribution of all or part of the turbulence spectrum. A review of the dispersion theory is presented. Then a Van Kármán turbulent kinetic energy spectrum is used to obtain an accurate representation of both very large and very small turbulence structures. New results are compared with previous results obtained with a Kolmogorov spectrum.

1:30

1pPA3. Scattering of sound by a penetrable sphere above a plane boundary. Wai Keung Lui and Kai Ming Li (Dept. of Mech. Eng., The Hong Kong Polytechnic Univ., Hung Hom, Kowloon, Hong Kong, mekellui@polyu.edu.hk)

The problem of acoustic scattering by a penetrable sphere irradiated by a point source is investigated. A theoretical model is developed for the case of scattering by a sphere placed above an acoustically hard and an impedance boundary. The sphere is made of an extended reaction material that is assumed to be acoustically penetrable. The problems are tackled by using the technique of variables separation and appropriate wave field expansions. By adopting an image source method, the solutions can be formulated in form of multiscattering interaction between the sphere and the image sphere near a hard and an impedance boundary. The effect of boundary impedance on the reflected sound fields is incorporated in the numerical model by using the well-known Wely–van der Pol formula. Preliminary indoor measurements are conducted in an anechoic chamber for the characterization of the acoustical properties of the penetrable sphere as well as the impedance boundary. A further set of experimental measurements is carried out to demonstrate the validity of the proposed theoretical models for various receiver locations around the sphere above the impedance boundary. [Work sponsored by the Innovation & Technology Commission, MTR Corp. Ltd., and The Hong Kong Polytechnic University under Project No. ZM07.]

1:45

1pPA4. Narrow-band pulse propagation in a one-dimensional stratified random medium. Shimshon Frankenthal and Mark J. Beran (Faculty of Eng., Tel Aviv Univ., Ramat Aviv, Israel)

In the stratified random medium considered here, the spatially fluctuating wave speed and refractivity vary only in the range direction, which is normal to the planes of stratification. To treat backscattering in such a medium, it is convenient to split the field into forward- and backward-propagating components, and derive coupled equations that govern the range evolution of the statistical moments of these components. The derivation entails an ensemble-averaging operation, which is here predicated on the key assumption that the two components, which are incident from opposite directions on a differential range slab, are statistically independent of the refractivity fluctuations within that slab. The assumption is justifiable when the fluctuations exhibit a sufficiently rapid time dependence. This procedure is used to derive equations for the bichromatic coherence of the two field components, which are then solved to determine the evolution of the power flux associated with the propagation of a signal due to a planar narrow-band pulse incident on a semi-infinite scattering slab. It is shown that the pulse signal is attenuated exponentially in the range direction, on the so-called ''localization'' scale, while the scattered power forms echo pulses which are gradually attenuated until all the incident pulse energy is reflected back out of the slab.

2:00

1pPA5. Nonlinear acoustics—Coupling between hydrodynamic and acoustic pressure fields. Mauricio Pazini Brandao (IEAv-CTA Rodovia dos Tamoios km 5,5, Sao Jose dos Campos, SP, Brazil, 12228-840)

Linear or Classical Acoustics considers that sound waves propagate essentially without any medium movement. However, in Aeroacoustics we deal with phenomena where there are the simultaneous generation and propagation of sound in a moving medium. Unless the flow velocity is very low, when in aerodynamics the fluid can be modeled as being incompressible, nonlinearities must be taken into account to better model the flow physics. In this paper, by considering the exact mass and momentum equations in the form of a generalized Ffowcs–Williams and Hawkings equation, the coupling between the hydrodynamic and acoustic pressure fields generated by the motion of a solid body in still air is studied. The extent of the validity of linear hypotheses and decoupling between the two pressure fields is discussed. Strategies to consider the coupling for faster movements are indicated. This coupling appears in computations as stepwise linear iterations between the hydrodynamic and acoustic perturbation pressure fields. The goal of this research is to devise a mathematical and computational model where the formation of normal shock waves in transonic flows appears as a nonlinear interaction process between aerodynamics and aeroacoustics.

2:15

1pPA6. Velocity measurements in the acoustic oscillatory boundary layer with laser Doppler anemometry. J. R. Castrejón-Pita, A. A. Castrejón-Pita, and G. Huelsz (Centro de Investigación en Energía, UNAM, A.P. 34, 62580 Mexico, Mor. Mexico)

Velocity measurements using laser Doppler anemometry (LDA) in the oscillatory viscous boundary layer (OVBL) produced by acoustic waves are presented. The amplitude velocity profile of acoustic standing waves in air with frequencies of 68.5, 114.5, and 343.5 Hz are reported. The results agree with the predictions of linear theories and show that the small discrepancies found in similar measurements made with hot wire anemometry (HWA) [Huelsz *et al.*, Exp. Fluids **32**, 612 (2002)] are due to the interaction of the hot wire probe with the flow and the wall. The advantages and disadvantages of the LDA technique compared with HWA are discussed for OVBL measurements. [Work supported by CONACyT Project No. 32707-U. The authors wish to thank Guillermo Hernández for his valuable technical support.]

2:30–2:45 Break

2:45

1pPA7. Effective impedance of hard rough surfaces. Keith Attenborough and Patrice Boulanger (Dept. of Eng., Univ. of Hull, Hull HU6 7RX, UK)

The effective impedance of hard rough surfaces insonified from a point source has been investigated through boss models, boundary element simulations, and measurements. The complex excess attenuation measured or predicted has been fitted for effective impedance by means of a formulation for the sound field above a smooth finite impedance plane. It has been found that the effective impedance plane is higher than the nominal one and that the results are very sensitive to the location of the specular reflection point. Polynomial expressions for the real and imaginary parts of the effective impedance have been derived for various types of roughnesses.

3:00

1pPA8. A high-order on-surface radiation condition for computing scattering from concave objects and extended surfaces. David C. Calvo (U.S. Naval Res. Lab., Washington, DC 20375)

Approximate solutions to scattering problems can be obtained efficiently using on-surface radiation conditions (OSRCs). An OSRC may be viewed as an approximation of the exact nonlocal (integral) Dirichlet-to-Neumann operator, and therefore requires less effort to compute surface quantities and far-field patterns in comparison with other methods. In a previous study [Calvo *et al.*, IEEE Trans. Antennas Propag. (in press)], a higher-order OSRC was developed for two-dimensional convex scatterers that was accurate for large scattering angles relative to the surface normal and moderate surface curvatures. A remarkable result was the exceptional accuracy obtained for end-on incidence with hard scatterers—a challenging case in which creeping waves arise and past OSRCs have had difficulty treating. In this talk, improvements to the OSRC are discussed by relaxing the convexity restriction to allow for compact objects with moderately deep concavity with favorable results. The OSRC is then applied to scattering by extended pressure-release surfaces featuring moderately deep corrugations. Results are favorable in comparison with other more costly numerical techniques. A particular advantage of the high-order OSRC occurs for grazing angles of incidence where the effects of shadowing come into play. Possible applications to ocean acoustics and microwave remote sensing of the sea surface will be addressed.

3:15

1pPA9. Experimental and numerical study of air-coupled surface waves generated above strips of finite impedance. Wenhao Zhu, Gilles A. Daigle, and Michael R. Stinson (Inst. for Microstructural Sci., Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

A surface composed of a lattice of small cavities can support the propagation of air-coupled surface waves. Energy is localized near the surface and, since the sound pressure near the surface can be greater than if the surface was rigid, passive amplification is obtained. By restricting the lattice of cavities to a strip of finite width embedded in an otherwise rigid surface, the amplification depends on the direction of the incident sound. Thus directional receivers can be designed. These directional receivers can provide more amplification than would be obtained with a semi-infinite lattice of cavities. In this paper, we will discuss various ways in which strips of finite impedance can be configured to enhance the amplification (and directivity) properties of surface waves. Experiments were carried out on a model surface and the measurements are compared with numerical results obtained using a boundary element method.

3:30

1pPA10. Sound propagation in a refracting atmosphere above an impedance discontinuity. Shahram Taherzadeh and Nick Harrop (Dept. of Environ. & Mech. Eng., The Open Univ., Milton Keynes MK7 6AA, UK)

de Jongs formulation of sound propagation above a ground with a single impedance change has been extended to include effects of a refracting atmosphere and atmospheric turbulence. The theory is compared with a numerical algorithm based on a hybrid Boundary Integral Equation/Fast Field Program developed for predicting the propagation of sound in a refracting atmosphere above an uneven, discontinuous terrain. By using the analogy of sound diffraction over curved surfaces to atmospheric refraction over flat ground surfaces, the effect of temperature and wind velocity gradients in the presence of flat ground surfaces can be studied. Measurements of the excess attenuation of sound from a point source over a mixed impedance curved surface are carried out in an anechoic chamber as well as outdoor measurements over a tarmac–grass discontinuity. These measurements are compared with predictions based on the extended de Jong theory and the hybrid BIE/FFP algorithm in the nonturbulent case. Results show that where there is a single discontinuity between acoustically hard and finite impedance surfaces both models are found to give satisfactory agreement with measured data except when the discontinuity is midway between the source and the detector.

3:45

1pPA11. Method of designing layered sound absorbing materials. Youssef Atalla and Raymond Panneton (Acoust. and Vib. Group, Univ. of Sherbrooke, 2500 Boul. de l'Universit Sherbrooke, QC J1K 2R1, Canada, yatalla@gme.usherb.ca)

A widely used model for describing sound propagation in porous materials is the Johnson–Champoux–Allard model. This rigid frame model is based on five geometrical properties of the porous medium: resistivity, porosity, tortuosity, and viscous and thermal characteristic lengths. Using this model and with the knowledge of such properties for different absorbing materials, the design of a multiple layered system can be optimized efficiently and rapidly. The overall impedance of the layered systems can be calculated by the repeated application of single layer impedance equation. The knowledge of the properties of the materials involved in the layered system and their physical meaning, allows to perform by computer a systematic evaluation of potential layer combinations rather than do it experimentally which is time consuming and always not efficient. The final design of layered materials can then be confirmed by suitable measurements. A method of designing the overall acoustic absorption of multiple layered porous materials is presented. Some aspects based on the material properties, for designing a flat layered absorbing system are considered. Good agreement between measured and computed sound absorption coefficients has been obtained for the studied configurations. [Work supported by N.S.E.R.C. Canada, F.C.A.R. Quebec, and Bombardier Aerospace.]

4:00

1pPA12. The interaction of airborne sound with the poroelastic ground. Craig J. Hickey and James M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38655, chickey@olemiss.edu)

An airborne acoustic wave impinging on the surface of the ground provides a good vibrational source for investigating the near surface. Research has been carried out on the use of this type of source for landmine detection, characterization of agricultural soils, and studying the characteristics of snow and ice cover. Since the ground and snow is porous, the impinging sound wave induces motion of the fluid within the pores as well as vibrating the solid framework. Vibrating the solid component of the ground using an airborne sound source is known in the acoustic community as acoustic-to-seismic (a/s) coupling. The most complete understanding of the interaction of airborne sound with the ground is to treat the ground as a poroelastic or poroviscoelastic medium. This treatment predicts three types of waves can propagate in a bulk homogeneous porous material with a deformable framework: two dilatational waves and one rotational wave. The dilatational waves are usually referred to as the fast, or Type I, and slow, or Type II, waves and deform both the solid and fluid constituents. This presentation will discuss the energy partition and impedance at an air-porous medium interface as related to the solid and fluid deformations.

4:15

1pPA13. Acoustical measurement system to improve agricultural implements for soil protection. Rodolfo Martinez and F. Diego (CIIDIR, IPN, Oaxaca, Mexico, rodolfo_mc@yahoo.com)

This research work is based on theoretical mechanics and agricultural mechanics theory applied to the study of the behavior of animal-pulled implements for seeding or harvesting for minimizing erosion of soil. For this purpose the pertinent acoustic measurements are presented with the objective of finding acoustical parameters useful to evaluate different designs of these kinds of implements with regard to depth and shape of the earth moving device in order to protect agriculture soil from erosion.

Session 1pSAa

Structural Acoustics and Vibration: Complex Structures and Fluid Loading

Peter H. Rogers, Chair
School of Mechanical Engineering, Georgia Institute of Technology, Graduate Box 268, Atlanta, Georgia 30332

Contributed Papers

1:00

1pSAa1. Vibration of submerged prolate spheroidal shells with internal fluid loading. Jeffrey E. Boisvert (Naval Undersea Warfare Ctr., Newport, RI 02841, boisvertje@npt.nuwc.navy.mil) and Sabih I. Hayek (Penn State Univ., University Park, PA 16802)

The equations of motion for nonaxisymmetric vibration of submerged prolate spheroidal shells of constant thickness with internal fluid loading were derived using Hamiltons principle. The shell theory used in this derivation includes shear deformations and rotatory inertias. The shell displacements and rotations were expanded in infinite series of comparison functions. These include associated Legendre functions in terms of the prolate spheroidal angular coordinate and circular functions in the azimuthal angle coordinate. The external and internal fluid loading impedances were computed using expansions of prolate spheroidal wave functions in each domain. The shell was excited by axisymmetric normal surface forces, including a point load at the shell apex and ring load at other locations. Numerical results were obtained for the driving and transfer mobilities for several shell thickness-to-half-length ratios ranging from 0.005 to 0.1, and for various shape parameters, a, ranging from an elongated spheroid shell ($a=1.01$) to a spherical shell ($a=100$). Results are presented for various combinations of external and internal fluid loading, and comparisons are made to the *in vacuo* shell vibration. [Work supported by ONR and the Navy/ASEE Summer Faculty Program.]

1:15

1pSAa2. Noise radiated from a rotating submerged elastic cylindrical thin shell. Jayme J. Caspall, Minami Yoda, and Peter H. Rogers (The George W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332)

Although the aeroacoustics of high Reynolds number boundary layers is reasonably well understood, less is known about the hydroacoustics of such flows, and the effect of fluid loading. The noise generated by the turbulent boundary layer around an elastic, thin-walled and cylindrical shell rotating in quiescent water was studied in the Georgia Tech. Underwater Acoustic Tank for Reynolds numbers up to 200 000. The steel shell, which is filled with air, has a diameter D of 0.625 m, a wall thickness of $0.004D$, and an aspect ratio of unity; the tank dimensions are $19D$ by $12D$ by $11D$. Extraneous noise sources (e.g., bearing and motor vibration) were isolated from the net signal to estimate flow noise. Radiated noise power was calculated from hydrophone data under a diffuse field assumption. To our knowledge, these results are unique in both their structural acoustics and fluid mechanics scaling.

1:30

1pSAa3. Vibration of a cylindrical membrane shell subjected to a steady rotational load. Arthur I. Koral and Mauro Pierucci (San Diego State Univ., San Diego, CA 92182-3108, mpierucci@engineering.sdsu.edu)

The vibration of a cylindrical membrane shell is driven by a point source moving along its circumferential direction. The point source is modeled by a Dirac delta function. The shell vibration is governed by the membrane equations. No variation is assumed to take place along the shell axial direction. The coupled partial differential equations for the radial and the tangential displacements $w(\text{theta},t)$ and $v(\text{theta},t)$ are decoupled by decomposition and two sets of coupled ordinary differential equations for the time varying magnitudes of the shell circumferential modes are obtained. The force starts its motion at $t=0$ and continues at a constant rate omega. The spectrum of the shell vibration, as expected, shows frequencies that are multiples of the rotating forcing frequency, multiples of the propagating wave front frequency as well as other components. For large times these spectra yield vibrations resembling those of broadband noise.

1:45

1pSAa4. Flexural waves on narrow plates. Andrew Norris (Dept. of Mech. & Aerosp. Eng., Rutgers Univ., Piscataway, NJ 08854-0909)

Flexural wave speeds on beams or plates depend upon the bending stiffnesses, which differ by a well-known factor depending on the Poisson's ratio. A quantitative analysis of a plate of finite lateral width displays the plate-to-beam transition, and permits asymptotic analysis that shows the dependence on the width. Orthotropic plates are analyzed using both the Kirchhoff and Kirchhoff–Rayleigh theories, and isotropic plates are considered for Mindlin's theory with and without rotational inertia. A frequency dependent Young's modulus for beams or strips of finite width is suggested, although the form of the correction to the modulus is not unique and depends on the theory used. The sign of the correction for the Kirchhoff theory is opposite to that for the Mindlin theory. These results indicate that the different plate and beam theories can produce quite distinct behavior. This divergence in predictions is further illustrated by comparison of the speeds for antisymmetric flexural modes on narrow plates. The four classical theories predict limiting wave speeds as the plate width vanishes, but the values are different in each case. The deviations can be understood in terms of torsion and how each theory attempts, or fails, to approximate this effect.

2:00

1pSAa5. Analysis of variable-thickness, streamlined transducer array windowing concepts for high speed underwater vehicles. Robert M. Koch (Adv. Technol. Div., Naval Undersea Warfare Ctr., Code 8232, Bldg. 1302, 1176 Howell St., Newport, RI 02841, kochrm@npt.nuwc.navy.mil)

Recent research in very high speed underwater vehicles shows they require a much more streamlined vehicle nose than the present rather flat tactical scale vehicle nose. It has been found that the common forward-facing transducer array with a constant thickness array window utilized on current lower speed vehicles is inadequate. Two newer alternate front-end array concepts suitable for lower drag, higher speed vehicles being investigated are (a) a variable thickness, streamlined array window on a flat forward facing array and (b) a streamlined vehicle nose structure consisting of a built-in distributed, conformal array. While these highly streamlined concepts are attractive from the standpoint of reducing vehicle drag at higher speeds, both require reexamination of the effects of high curvature on front-end sonar array performance. In the present paper, both above streamlined array concepts are numerically examined using the SARA2D dynamic structural-acoustic Finite Element Analysis (FEA) code. First, the acoustic performance of a hemispherical array window is compared to a common equal aperture flat window in both transmit (outgoing

acoustic wave) and receive (incident acoustic wave) modes. Second, a dynamic simulation of acoustic array beam-forming and beam-steering is performed for a conical shaped conformal array to assess the feasibility of the concept.

2:15

1pSAa6. Modeling of wall-pressure fluctuation based on time-mean flow field. Yu-Tai Lee, William Blake, Theodore Farabee, Michael Tse, and Joseph Brown (Naval Surface Warfare Ctr., Carderock Div., 9500 MacArthur Blvd., West Bethesda, MD 20817)

Time-mean flow field and turbulent flow characteristics from solving the Navier–Stokes equations are used to predict the flow direction wave number spectrum of wall-pressure fluctuations. The vertical turbulent velocity is represented by the kinetic energy contained in the local turbulent flow. A redistribution of the turbulent kinetic energy to account for the anisotropic turbulence is implemented based on the equilibrium turbulent shear flow. The spectral correlation model in predicting the wall-pressure fluctuation is obtained through the Green's function formulation, and the streamwise and spanwise wave number spectrum modelings. Predictions for the equilibrium flow agree well with measurements, and reveal an overlap region of an inverse of the frequency. Predictions for a 2-D reattached flow field after flow separation and a 3-D flow field on an airfoil with focus on the trailing edge and airfoil tip are discussed and compared to measurements. [Work supported by ONR.]

MONDAY AFTERNOON, 2 DECEMBER 2002 CORAL GALLERY 1, 2:45 TO 4:30 P.M.

Session 1pSAb

Structural Acoustics and Vibration: Analytical and Numerical Methods

Rolando Menchaca, Chair
Av. Juan de Dios Batiz S/N, Col. Lindavista, D.P. C.P. 07070, Mexico

Contributed Papers

2:45

1pSAb1. A hybrid finite-element formulation for analyzing systems of beams and plates in the midfrequency range. Sang-Bum Hong and Nickolas Vlahopoulos (Dept. of Naval Architecture and Marine Eng., Univ. of Michigan, 2600 Draper Rd., Ann Arbor, MI 48109-2145)

A hybrid finite-element analysis (hybrid FEA) formulation has been developed in the past for computing the midfrequency vibration of systems that contain one type of energy. The hybrid FEA is based on characterizing as long members in a system all the members that contain a large number of wavelengths within their dimension. All the remaining members are considered as short. The Energy Finite-Element Analysis (EFEA) is employed for modeling the behavior of the long members, while the conventional FEA method is utilized for modeling the short members. In this paper the hybrid FEA formulation is extended to plate structures that are spot-welded to a frame comprised by tubular members. The new formulation is validated by comparing the hybrid FEA solution to results produced by very dense conventional FEA models in the mid-frequency range. [Research supported by the Automotive Research Center, US. Army.]

3:00

1pSAb2. Petrov–Galerkin's method hybrid with finite element into the Helmholtz equation solution. Part I. Itzala Rabadan Malda (Academia de acustica, ICE, ESIME, IPN, Mexico, D. F.), Emigdio Salazar Cordero (ESFM, IPN, Mexico, D. F.), and Jose Angel Ortega Herrera (SEPI, ESIME, IPN, Mexico, D. F.)

In this first part, this paper's purpose is to determine a numerical algorithm according to certain advantages that both Petrov–Galerkins and finite element method offer to resolve the wave equation completely, for simple geometry domains like circular and rectangular membranes, cones and cylinders, and parabolic and spherical volumes. It will permit to contemplate the Helmholtz equations reduced case. Then it can consider real domains like membranes, cones, cylinders, pipes, and parabolic and spherical volumes, polymaterial or discontinuous (with holes will be treated in a second part). This will allow determination of important design parameters for different acoustic devices.

3:15

1pSAb3. Petrov–Galerkin's method hybrid with finite element into the Helmholtz equation solution. Part II. Itzala Rabadan Malda (Academia de Acustica, ESIME, IPN, Mexico, D. F.), Emigdio Salazar Cordero (ESFM, IPN, Mexico, D. F.), and Jose Angel Ortega Herrera (IPN, SEPI, ESIME, Zacatenco, Mexico, D. F.)

This work proposes a hybridization between Petrov–Galerkins numeric method and finite element method (FEM) to resolve Helmholtz equation when dominion is an open or semiopen tube-shaped configuration and with determinate number of holes over cylindrical surface. It's pretended to solve these kind of cavities, thereby it allows us to obtain very important design parameters like: cavity length, quantity, size and distance between toneholes, form and size of mouthpiece or outlet. These parameters are design basis into acoustic and musical instrumentation: baffles outlet pipes, diffusers, silencers, flutes, oboes, saxophones, trumpets, quenas, and many more. In this way it's expected to determine advantages of this numeric method above another using actually.

3:30

1pSAb4. Simulation of the transient behavior of impact machinery using finite elements package ANSYS. Jose de Jesus Villalobos-Luna, Pedro Lopez-Cruz, Patricia Gonzalez-Guajardo, and Fernando J. Elizondo-Garza (Vib. Lab., FIME, Univ A. de Nuevo Leon, P.O. Box 28 "*F*," Cd. Universitaria, San Nicolas, 66450, N.L., Mexico, villalobosluna@hotmail.com)

In this paper the finite elements method is used to simulate the transient behavior of impact machinery and its foundation. Using the software package ANSYS, two cases are modeled: first, the impact machine directly mounted on the floor, and second, the machine mounted over an inertia block foundation. The results of the modeling permit clear visualization of transient vibrations in the machine and how they are transmitted to the foundation and the floor. How to use this modeling for didactic purposes is discussed.

3:45

1pSAb5. Nonstationary bifurcations with additive noise. Huw G. Davies (Univ. of New Brunswick, Fredericton, NB E3B 5A3, Canada, davies@unb.ca)

The response of nonlinear systems with slow variation of a parameter through a bifurcation may be critically affected by very-low-level noise. Examples to be discussed include passage through resonance of a periodically pumped laser, a transcritical bifurcation associated with a coupled pitch/roll model of ship motion, and a period-doubling cascade of a discrete map. In each case the original form of the response persists as the control parameter varies through the bifurcation, even though this original form is now unstable. Noise is required to trigger the bifurcated form; then there is a rapid transition with exponential growth from one form of response to another. Sinusoidal parameter variation through the bifurcation at sufficiently high frequency may effectively stabilize the original form by preventing an appreciable growth of the bifurcated form; noise is destabilizing. The two competing effects are analyzed using matched-asymptotic-expansions. The innermost of a nested set of three expansions describes how noise causes the growth of the bifurcating form from the original unstable form. Previous results for period-doubling [H. G. Davies and K. Rangavajhula, Proc R. Soc. London, Ser. A **457**, 2965–2982 (2001)] are modified for the other types of bifurcation, and are extended to a higher order period-doubling cascade using renormalization.

4:00

1pSAb6. Transient modeling of Lamb waves emitted by integrated transducers using a hybrid finite element–normal mode expansion technique. Laurent Duquenne, Emmanuel Moulin, Jamal Assaad, and Christophe Delebarre (Le Mont Houy, 59313 Valenciennes, Cedex 9, France, laurent.duquenne@univ-valenciennes.fr)

For a few years, the concept of integrated health monitoring of aeronautic structures has become an important issue. A promising solution is the emission of ultrasonic Lamb waves using thin piezoelectric transducers integrated to the structure. Indeed Lamb waves are known to be able to propagate over long distance and to interact with inner structural defects. One important aspect when applying Lamb waves to health monitoring is the control and optimization of their generation. Therefore, a hybrid finite element–normal mode expansion modeling technique adapted to piezoelectric transducers has been recently developed. This technique allows the determination of the amplitude of each Lamb mode as a function of the transducer parameters. So far, it had been successfully applied to the case of harmonic excitation only. In this paper, adaptation of this hybrid technique to transient excitation is presented, which enables to simulate actual experimental conditions. Contrary to other techniques presented in the literature, the present technique allows to take into account the actual electromechanical behavior of the integrated transducer. Predicted results have been successfully compared to experimental measurements performed on different types of materials.

4:15

1pSAb7. Sound radiation of a discontinuous structure calculated with various semianalytical and numerical methods. Berndt Zeitler and Michael Möser (Institut für Technische Akustik, Sekr. TA 7, Einsteinufer 25, 10587 Berlin, Germany)

The subject of investigation is an infinitely long radiator with the cross section of a pie slice. The modal radiation efficiency and directivity pattern are calculated for given normal velocity on the surface; rigid legs are assumed. Three semianalytical methods and one numerical method are used. The Wave-Approach (WA) splits the field into two regions assuming different wave types in each. Their amplitudes are chosen to fulfill the surface boundary conditions and to ensure a smooth transition between the regions. The Multi-Point–Multipole-Approach (MPMA), as the name suggests, uses several multipoles scattered within the structure, whose amplitudes are chosen such that the reconstructed normal velocity on the surface ideally equals the original. In this study the order is limited; only one mono- and two perpendicular dipoles are used. No restriction of order is made in the Single-Point–Multipole-Approach (SPMA), but of course one of quantity. For the Finite-Element-Method (FEM) Sommerfeld's radiation condition was chosen as the outer boundary condition. The results for different criteria and under certain circumstances are compared and discussed. The sound radiation is also visualized through animations displaying particle movement.

MONDAY AFTERNOON, 2 DECEMBER 2002 GRAND CORAL 3, 1:30 TO 4:00 P.M.

Session 1pSC

Speech Communication: Speech Perception and Production by Hearing Impaired Individuals

Mario A. Svirsky, Cochair
Department of Otolaryngology, Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, Indiana 46202

Jorge Gurlekian, Cochair
Laboratorio de Investigaciones Sensoriales, CONICET-UBA, Marcelo T. de Alvear 2202, 1122 Buenos Aires, Argentina

Chair's Introduction—1:30

Invited Papers

1:35

1pSC1. Language-specific, hearing-related changes in vowel spaces: A study of English- and Spanish-speaking cochlear implant users. Joseph Perkell (Speech Commun. Group, Res. Lab. of Electron., MIT, Cambridge, MA 02139), Marek Polak, Thomas Balkany (Univ. of Miami Med. School, Miami, FL), Jennell Vick, Harlan Lane, Ellen Stockmann, Mark Tiede, and Majid Zandipour (MIT, Cambridge, MA 02139)

This study investigates the role of hearing in vowel productions of postlingually deafened cochlear implant (CI) users. Based on the hypothesis that competing demands of intelligibility and economy of effort influence vowel production, we predicted that speakers of a language with a more crowded vowel space, such as American English, would show an increase in average vowel spacing (AVS

the average inter-vowel distance in the F1-F2 plane) with the provision of hearing from a CI; whereas speakers of a language with fewer vowels, such as Spanish, would not. Results of a preliminary on–off study supported the hypothesis: with hearing from a CI, 7 English-speaking CI users all increased AVS; however, among 7 Spanish speakers, some increased AVS and some reduced it. The current study is exploring these findings further with new groups of CI users and normal-hearing speakers of English and Spanish and an elaborated paradigm that includes different speaking rates and comparisons between pre-implant and post-implant recordings. Preliminary results will be reported. [Work supported by NIH.]

2:00

1pSC2. Computational models of vowel identification by English- and Spanish-speaking cochlear implant users. Mario A. Svirsky, Daniel C. Hadley, Heidi S. Neuburger, Su-Wooi Teoh (Dept. of Otolaryngol.-HNS, Indiana Univ. School of Medicine, Indianapolis, IN 46202), and Eloy Villasuso (Univ. of Miami, Miami, FL)

Cochlear implant (CI) users achieve significant levels of speech recognition on average, but cross-listener variability is very high. Unfortunately, our understanding of the speech perception mechanisms employed by CI users is still incomplete. To address this issue we have developed the multidimensional phoneme identification (MPI) model, which aims to predict phoneme identification for individual cochlear implant users based on their discrimination along specified acoustic dimensions. The MPI model was used to fit vowel confusion matrices from English- and Spanish-speaking CI users. Good agreement between predicted and observed matrices was obtained for both English and Spanish. Some of the acoustic dimensions required to obtain these fits were the same for both languages (e.g., $F1$ and $F2$), but others were not (e.g., $F3$ is required to obtain a good fit in English, but not in Spanish). These results are consistent with differences in the acoustic phonetics of the two languages: a low value of $F3$ is used in English to encode the retroflex vowel /r/, and this sound does not exist in Spanish. These results raise the possibility that the optimal stimulation strategies may differ across languages. [Work supported by NIDCD (R01-DC03937), NOHR, DRF, and BID/CONICYT (Uruguay).]

2:25

1pSC3. Use of the Spanish adaptation of the phonetic and phonologic Ling tests to evaluate progress in children with cochlear implants and hearing aids. Gonzalo Corvera, Maria Cristina Manrique, Juan Pablo Manrique, and Valeria Lapilover (Clinica Lomas Altas, Paseo de la Reforma 2608-9, Mexico D.F. 11950, Mexico)

The decision of when to implant a prelingually deafened child is based on his or her progress with hearing aids. It comes down to a subjective decision regarding whether the child could do better with an implant or not, is heavily influenced by the therapists' experience with cochlear implants, and is difficult to transmit objectively to new teams. The speech discrimination tests usually used tend to use words or phrases in and out of context scenario. Dr. Daniel Ling's phonetic and phonologic tests (1990) include aspects on auditory abilities, language comprehension, and articulation, and are evaluated in the course of therapy and not in a special evaluatory session. The Spanish adaptation of these tests (adapted by Dr. Ling in collaboration with Santiago Manrique and Cristina de Manrique, 1996) were used for this study, analyzing the results in children with cochlear implants and with hearing aids, to explore their use (or lack of it) as an adjunct for the prescription of a cochlear implant.

2:50–3:10 Break

3:10

1pSC4. Development of an Argentine Spanish hearing impaired speakers database. Jorge A. Gurlekian (Lab. de Investigaciones Sensoriales, CONICET-UBA, Marcelo T. de Alvear 2202, (1122) Buenos Aires, Argentina, jag@fmed.uba.ar) and Ana Sanchez Navarro (Centro de Medicina Física y Rehabilitación, (1058) Buenos Aires, Argentina)

Argentine Spanish hearing impaired speech has not been systematically registered or studied. A database with standardized methods for data collection and labeling is presented here. Speech of 600 speakers, adults, and children, with different types of hearing losses is recorded at 44100 Hz and 16 bits. The corpus consists of word lists, grouped by tonality and frequency of use, and five short sentences, containing all the phonemes and allophones. Semi-spontaneous speech, consisting of a description of two sets of simple pictures, is also recorded. Data are acoustically analyzed and labeled at the segmental—graphemic, phonetic and acoustic tiers—and suprasegmental levels, using a labeling method that incorporates psychoacoustic information into a combination of categorical and continuous descriptions of fundamental frequency. Syntactic labels are added to semi-spontaneous speech data. Labeled files are loaded into a relational SQL database to allow different types of queries. A complementary database provides information about personal and medical data, and results of hearing tests. These databases contribute to the description of acoustic-phonetic features of deaf speech, a comparison between speech perception and production, and new approaches for the adjustment of auditory prostheses. [Work supported by FONCYT, CONICET and Mutualidad Argentina de HipoacB4usicos.]

3:35

1pSC5. Cortical reponses to electrical stimulation from cochlear implants. Emily A. Tobey, Michael D. Devous, Sr., and Peter S. Roland (Univ. of Texas at Dallas, 1966 Inwood Rd., Dallas, TX 75235)

In order to examine factors contributing to speech perception performance variations in cochlear implant users, single photon emission computed tomography (SPECT) was used to examine cortical activity elicited by the electrical stimulation of multichannel cochlear implants. Subjects watched a 15-min videotaped story under several conditions: (a) audio presented monaurally in the aided right or left ears; (b) audio presented binaurally with aids; (c) a visual-only presentation; and (d) signals processed through the implant. Five minutes after the start of each story, 99mTc HMPAO was injected over 30 s to insure subjects were unaware of the tracer administration. Right and left monaural stimulation in normal-hearing subjects resulted in significant bilateral activation of Brodmann areas 41, 42, 22, 21, and 38. Despite similar hearing losses to each ear, significant differences in per-operative auditory cortex

activation was observed between ears across the subjects. Post-implant, individuals with relatively high levels of open-set speech perception demonstrated bilateral activation of cortex; however, the extent of activation was significantly less than that observed for normal-hearing individuals. Individuals with minimal open-set speech perception scores demonstrated unilateral activation of the cortex on the hemisphere contralateral to the ear of implantation. [Work supported by NIH.]

MONDAY AFTERNOON, 2 DECEMBER 2002 GRAND CORAL 1, 1:00 TO 5:30 P.M.

Session 1pUW

Underwater Acoustics: Localization and Inversion Issues

Monica Montanari, Cochair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Te-Chih Liu, Cochair
Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Contributed Papers

1:00

1pUW1. Rapid inversion of waveguide environmental parameters using an analytic adjoint method. Aaron Thode (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093)

A method is presented for inverting sound-speed perturbations in a stratified waveguide environment over several kilometers range. An adjoint Green's function formalism is used to backpropagate the difference between a measured and modeled received field through the modeled waveguide environment, in a manner reminiscent of acoustic time-reversal methods. Using a set of linearized equations similar in form to the Born approximation, the backpropagated error field is converted into an estimate of an error surface gradient, or Frechet derivative [S. J. Norton, J. Acoust. Soc. Am. **106**, 2653–2660]. The gradient can then be used to locate the error surface minimum via a series of iterations. The use of a normal-mode waveguide Green's function enables a considerable analytic simplification of the equations. The similarities and differences between the adjoint approach, matched-field processing, and diffraction tomography are discussed. [Work sponsored by ONR.]

1:15

1pUW2. Beam-based matched field steering for large volumetric arrays. Henry Cox and Kevin D. Heaney (ORINCON Industries, 4350 N. Fairfax Dr., Ste. 470, Arlington, VA 22203)

A new approach is introduced to significantly improve the efficiency of generation of matched field steering vectors for large volumetric arrays. For a volumetric array it is shown that there is a virtual Equivalent Vertical Line Array (EVLA) with many fewer elements that captures all of the information contained in the acoustic propagation from a far field source to the receiving array. A relationship between the size of the virtual EVLA and the geometry of the volumetric array is derived. Element-based matched field steering vectors for the virtual EVLA can be computed in the usual way and then used to generate beam-based matched field steering vectors for shaded plane wave beams. By matching the mainlobes of the shaded plane wave beams of the virtual EVLA with the vertical beam pattern of the beams of the volumetric array, the beam-based steering vectors for the large real volumetric array are obtained without needing to compute its element-based steering vectors. For large volumetric arrays the savings can exceed an order of magnitude.

1:30

1pUW3. Characteristics of small boat acoustic signatures. Martin L. Barlett and Gary R. Wilson (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, barlett@arlut.utexas.edu)

Small boats are often a dominant noise source in harbors, coastal regions, and lakes. However, detailed information about acoustic characteristics is not generally available. To remedy this deficiency, measurements of underwater acoustic signatures from various small boats have been conducted under controlled conditions. Boats used in the measurements were powered by a variety of gasoline and diesel motors in outboard, inboard–outboard, and inboard drive configurations. Measurements were made using a bottom mounted hydrophone in about 100 feet of water. In many instances, the boats were instrumented with a system that recorded the GPS position and engine RPM permitting accurate determinations of platform operating parameters. Measured small boat signatures contain both narrowband and broadband acoustic components. Narrowband components are generally associated with sources from the engine or drive. Narrowband levels in the observed spectra were found to exhibit significant variations and are believed to originate from fluctuations in the load on the engine and drive. Broadband energy was observed to be a dominant noise source at frequencies above several hundred Hertz and is a characteristic feature in all small boat signatures. Examples of small boat signatures will be presented and representative acoustic characteristics of this type of watercraft will be discussed.

1:45

1pUW4. A 3D forward-look sonar simulation model. Pierre-Philippe J. Beaujean, Julia Gazagnaire, and Joseph M. Cuschieri (Dept. of Ocean Eng., Florida Atlantic Univ., 101 N. Beach Rd., Dania Beach, FL 33004, pbeaujea@seatech.fau.edu)

A modeling tool for 3-D Forward-Look Sonar (3-D-FLS) or equivalent has been developed. The model replicates the entire process of a 3-D Forward-Look sonar generating a 3-D image of a target. The 3-D-FLS operates at a central frequency of 225 kHz and has a range of 100 m. The transmitted signal is a frequency-modulated chirp with 30 kHz of bandwidth. The source is a line array and the projector is a 64-element line array. Ray theory is used to calculate the trajectory of the rays. The model takes sound velocity profiles into account. The reverberation levels due to the sea bottom and sea surface are calculated using Kuos model. Targets of various shapes and sizes are modeled as a collection of reflecting data points. The target strengths of each point on a specific target are assumed equal. The final product is a modeling tool which can be used to define the

sonar hardware and processing software necessary to achieve various operational needs. Simulated results are validated using 3-D-FLS data collected during several at-sea experiments.

2:00

1pUW5. Autonomous underwater vehicle adaptive path planning for target classification. Joseph R. Edwards and Henrik Schmidt (Dept. of Ocean Eng., MIT, 77 Massachusetts Ave., Rm. 5-204, Cambridge, MA 02139)

Autonomous underwater vehicles (AUVs) are being rapidly developed to carry sensors into the sea in ways that have previously not been possible. The full use of the vehicles, however, is still not near realization due to lack of the true vehicle autonomy that is promised in the label (AUV). AUVs today primarily attempt to follow as closely as possible a preplanned trajectory. The key to increasing the autonomy of the AUV is to provide the vehicle with a means to make decisions based on its sensor receptions. The current work examines the use of active sonar returns from mine-like objects (MLOs) as a basis for sensor-based adaptive path planning, where the path planning objective is to discriminate between real mines and rocks. Once a target is detected in the mine hunting phase, the mine classification phase is initialized with a derivative cost function to emphasize signal differences and enhance classification capability. The AUV moves adaptively to minimize the cost function. The algorithm is verified using at-sea data derived from the joint MIT/SACLANTCEN GOATS experiments and advanced acoustic simulation using SEALAB. The mission oriented operating system (MOOS) real-time simulator is then used to test the onboard implementation of the algorithm.

2:15

1pUW6. Analysis of shallow water experimental acoustic bistatic data. Richard Brothers, Gary Heald, and Samantha Dugelay (QinetiQ, Newton's Rd., Weymouth DT4 8UR, UK, rbrothers@qinetiq.com)

During February 2002, data were collected in controlled experimental conditions in Weymouth Bay, UK. The experimental setup consists in a broadband source with variable depth and variable azimuthal and tilt angles and a 16 element receiver array located on a tower. The tower is placed at 27.4 m or 33 m from the source and the receiver is pointed in various directions for each pulse. Transmitted pulses consist mainly in CW and linear chirps with central frequencies at 150 kHz, bandwidths varying up to 100 kHz, and pulse lengths varying from 67 ms to 15 ms. The primary objectives of the data reported here are: (1) compare and validate the bistatic scattering strength model developed by the Applied Physics Laboratory of University of Washington; and (2) to assess the complexity and practical implementation of bistatic scattering for mine hunting purposes. Data are processed and the bistatic scattering strength is compared to the predicted bistatic scattering strength. Comparisons are discussed for various geometrical configurations and for different pulse types. Results show good concordance of the model and the experimental data despite ambient noise limitations due to difficult sea conditions.

2:30

1pUW7. Buried mine classification by means of higher-order spectral analysis. Monica Montanari and Henrik Schmidt (Dept. of Ocean Eng., MIT, 77 Massachusetts Ave., Rm. 5-204, Cambridge, MA 02139)

In the joint MIT/SACLANTCEN Generic Oceanographic Array Technology Sonar (GOATS) project, future missions are envisioned that employ one or more autonomous underwater vehicles (AUVs) for automated target detection and classification. The classification of buried mines after detection is important for reducing the false alarm rate of the mine hunting sonar system. Buried mine classification is pursued in this paper by employing higher-order spectral analysis. Due to the nature of the target, it is expected that there will be at least one elastic response following the specular return from the object. The elastic response is typically of very low signal to noise ratio, and the noise is correlated reverberation, so advanced spectral analysis must be employed to detect the elastic behavior of the target. Higher-order spectral analysis methods are applied to a time signal in a short temporal window following the specular return. The delay of the elastic return corresponds to the time necessary for the wave to travel around the perimeter of the object before it is reradiated toward the receiver. The effectiveness of this approach to mine/rock classification is shown by an application to at-sea data from the GOATS experiment series (1998, 2000, and 2002) as well as advanced acoustic simulations.

2:45

1pUW8. Deriving the spatial distribution and properties of radiating acoustic sources in a ship from a measurement of its signature in shallow water. Marshall V. Hall (DSTO Maritime Operations Div., 17 Pirrama Rd., Pyrmont, NSW 2009, Australia, Marshall.hall@dsto.defence.gov.au)

When a wideband point source passes a hydrophone in shallow water, the spectrogram of the signature shows a characteristic interference pattern. This pattern consists of a number of bands of strong signal, which decrease in frequency as the source approaches, and pass through minimum frequencies at the closest point of approach (CPA). If the environmental waveguide is range-independent the spectrogram is symmetric around the time of CPA. The geoacoustic model of the seabed is determined by using an optimization routine to match the computed with the measured spectrogram. For a hydrophone in the nearfield of a passing ship, the bands in the spectrogram are less clear, and are asymmetric even if the waveguide is range-independent. The objective is to represent a ship by a small number of point sources distributed in space, each of which has its own spectrum and effective directionality. The main source, the propeller, is directional in effect because the hull is an obstacle to transmission in the forward direction. For each source, the spectrum and directionality are expressed as functions of frequency and angle. The function coefficients are determined by using an optimization routine to match the computed with the measured signature.

3:00–3:15 Break

3:15

1pUW9. Concurrent detection and localization of buried targets using track-before-detect processing. Te-Chih Liu and Henrik Schmidt (MIT, Cambridge, MA 02139)

The GOATS (Generic Oceanographic Array Technology Systems) Joint Research Program explores the development of environmentally adaptive autonomous underwater vehicle (AUV) technology for Rapid Environment Assessment and Mine Counter Measurement (MCM) in coastal environments. MIT is developing the GOATS multistatic sonar concept which uses a low-frequency source on one AUV to subcritically insonify the seabed over a wide area, while a formation of multiple AUVs are used for mapping the associated 3-D scattered acoustic field in the water column. Exploring the different characteristics of the scattering from various buried targets, an algorithm has been developed for autonomous, concurrent detection and localization of buried targets. At the subcritical isonification angles necessary to sustain high coverage rate, the buried target detection depends on the measurement of very weak signals, in general undetectable by traditional detect-before-track algorithms. The method developed here applies a Track-Before-Detect (TBD) approach to solve this problem. This technique tracks the AUV trajectory first using the slowly changing environment information, and then the weak signal detection is declared after the AUV track is estimated at sufficient level of confidence. This new algorithm has been applied on the GOATS-2000 bistatic data. The result shows the successful navigation results by confident detection of three buried targets, two of which are particularly weak scatterers. A real-time, autonomous implementation of the algorithm was recently demonstrated in the GOATS-2002 experiment, and the result for both experiments will be discussed. [Work supported by ONR.]

3:30

1pUW10. Source motion effects on matched field processing. Katherine H. Kim and William S. Hodgkiss (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704, khkim@mpl.ucsd.edu)

Past matched field processing (MFP) analyses have typically overlooked the effects of source motion. However, in real-world applications of MFP, especially those involving the use of large arrays or in the highly dynamic shipping environments common to littoral waters, source motion can exacerbate the conflicting requirements of sufficient samples for adequate data covariance estimation and short-term stationarity assumptions. In the context of MFP, short-term stationarity can be assumed provided that the source does not traverse more than one beamwidth of the matched field processor. Thus, theoretically derived estimates of MFP resolution for a vertical line array are compared to those determined empirically from a shallow water acoustics experiment off the coast of San Diego. Predictions of the eigenvalue spreading that arises from source motion, based upon the work of Baggeroer and Cox [Baggeroer and Cox, Proc. of 1999 Asilomar Conference **1**, 103–108 (1999)], are then compared to experimental eigenspectra. Finally, the impact of interferer motion on adaptive MFP processor performance is presented.

3:45

1pUW11. Broadband matched-field tracking with horizontal line arrays in shallow water. Paul A. Baxley, Ahmad T. Abawi, and Randall Brannan (SPAWAR Systems Ctr., San Diego, Code 2857, 49575 Gate Rd., Rm. 170, San Diego, CA 92152-6435, baxley@spawar.navy.mil)

The use of three-dimensional matched-field processing to obtain time-dependent range, depth, and bearing estimates with a horizontal line array (HLA) can be computationally restrictive. The problem becomes much more tractable, however, when beam forming is first used to obtain bearing estimates versus time, which are then used to obtain range-depth ambiguity surfaces versus time. For the case of a moving source, matched-field tracking (MFT) compensates for source motion by integrating matched-field correlations over candidate source tracks through the bearing-range-depth space. The true track is determined by the highest integrated correlation, which also results in a gain in detectibility for the true source track. The effectiveness of this approach for the tracking of broadband sources in shallow water is studied via the analysis of HLA data collected during the Shallow Water Evaluation Cell Experiment 1996 (SWellEX-96), which occurred in 200 m water, 6 km southwest of San Diego. Broadband signals in this experiment included a narrowband comb and the broadband signature emanating from the source-tow ship. The MFT methoc can be used as a stand-alone tracker or as a tool to obtain initial track estimates for the initialization of a more efficient conventional tracker. [Work supported by ONR 321US.]

4:00

1pUW12. Identifying individual sources in matched-field processing by modifying the covariance matrix. Eric M. Giddens, Philippe Roux, and W. A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California, San Diego, 8820 Shellback Way, La Jolla, CA 92093-0238)

A method for identifying multiple sources in matched-field processing (MFP) is proposed. The ambiguity surface produced by MFP may contain sidelobe structures that confound individual source identification. Data corresponding to the location of the loudest source is added to the signal and the dominant eigenvector of the covariance matrix is removed. The addition of data prevents the degenerate case where a single eigenvector describes multiple sources. The new MFP result has the source and its sidelobes removed from the ambiguity surface. The remaining sources can then be identified. This technique is applied to simulated data as well as Swell-Ex 96 data with a small number of snapshots. [Work supported by ONR.]

4:15

1pUW13. Sound from a light aircraft for underwater acoustic inversions. Eric M. Giddens, Fernando Simonet, Thomas R. Hahn, and Michael J. Buckingham (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California, San Diego, 8820 Shellback Way, La Jolla, CA 92093-0238)

Experiments are being conducted in shallow (30 m) water off La Jolla, CA, to investigate the potential usefulness of the sound from a single-engine, propeller-driven, light aircraft for performing underwater acoustic inversions. The sound signature of the aircraft contains harmonics between 50 Hz and 1 kHz, which return the low-frequency geoacoustic properties of the seabed. A microphone approximately 1 m above the surface monitors the sound in air, a seven-element vertical array detects the acoustic arrivals underwater and a single, buried hydrophone receives the signals in the sediment. Aircraft overflights have been made at altitudes between 33 m and 330 m, yielding the altitude-dependence of the peak levels received underwater. Using the vertical array, the reflection coefficient of the seabed is being measured as a function of grazing angle. From the reflection coefficient, the critical angle of the sea floor and hence the sound speed in the sediment are inferred. The sound speed in the sediment should also be available directly from the Doppler shift on the buried hydrophone. These techniques and the available data sets will be discussed in the presentation. [Work supported by ONR.]

4:30

1pUW14. An AUV underwater acoustic array. Gerald L. D'Spain, Paul A. Lepper, and Richard Zimmerman (Marine Physical Lab, Scripps Inst. of Oceanogr., La Jolla, CA 93940-0704)

An eight-element hydrophone array has been installed within the shroud of an Odyssey IIb AUV. Each array channel is digitized and recorded with 10-kHz bandwidth by an autonomous, PC-104+ data acquisition system. To determine the effects of the AUV body itself on the acoustic field measured by the array, tones at low (200–800 Hz) and mid- (2–8 kHz) frequencies were transmitted in a large water tank to the AUV by a source continuously varying in azimuth but fixed range. Bartlett beamforming for source azimuth using both plane wave replicas and replicas calculated using a 2-D time domain, finite-difference code that accounts for the scattering from the two glass spheres in the AUV shows that (1) not accounting for scattering results in 2- to 6-dB signal gain degradation, but (2) scattering moves the ambiguity surface sidelobes closer to the main lobe location, thus improving localization capability. The effect of scattering from the AUV on an azimuthally isotropic noise field, estimated by averaging over pings at all azimuths, is a decrease in spatial coherence at low frequencies over that in free space, thus increasing detection capability. [Work supported by ONR.]

4:45

1pUW15. Tabu optimization for matched field inversion. Zoi-Heleni Michalopoulou and Urmi Ghosh-Dastidar (Dept. of Mathematical Sci., New Jersey Inst. of Technol., Newark, NJ 07102, michalopoulo@adm.njit.edu)

Matched field processing is a powerful tool for source localization and geoacoustic inversion. Because of significant environmental and geometry uncertainties, however, matched field processing usually involves multiparameter searches. To facilitate these searches, global optimization techniques such as genetic algorithms and simulated annealing have been successfully employed. In this work, a different approach, tabu, is implemented for optimization in matched field inversion. Tabu is a technique relying on the use of memory; it searches for the global maximum of the objective function through a navigation process that avoids already revisited models, also making use of aspiration criteria and diversification for faster convergence. The tabu performance in localization and geoacoustic inversion is demonstrated through experimentation with both synthetic and real (SWellEX 96) data. The approach is shown to provide reliable estimates in an efficient manner. [Work supported by ONR.]

5:00

1pUW16. Determining low-frequency source location from acoustic phase measurements. Travis L. Poole (MIT/WHOI Joint Prog. in Oceanogr. and Oceanogr. Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, tycho54@excite.com) and George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

For low-frequency cw sound sources in shallow water, the time rate-of-change of the measured acoustic phase is well approximated by the time rate-of-change of the source-receiver separation distance. An algorithm for determining a locus of possible source locations based on this idea has been developed. The locus has the general form of a hyperbola, which can be used to provide a bearing estimation at long ranges, and an estimate of source location at short ranges. The algorithm uses only acoustic phase data and receiver geometry as input, and can be used even when the source frequency is slightly unstable and/or imprecisely known. The algorithm has been applied to data from low-frequency experiments (20–300 Hz), both for stable and unstable source frequencies, and shown to perform well. [Work supported by ONR and WHOI Academic Programs Office.]

5:15

1pUW17. A simple method of passive source range measurement. Zeng Juan, Gao TianFu, Guo ShengMing, and Chen Yan (Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC)

In this paper, a simple method of passive range measurement is discussed. The receiver system can be either hydrophone or horizontal array. Unlike earlier methods of passive source range measurement, we mainly consider the phase information of the passive source, rather than the amplitude. The method only uses the simplest principle range=velocity time. The key of the method is how to acquire the partial deviation of the phase with respect to radial frequency and measurement variables. In physics, the former is equal to time, which can be obtained from the spectrum of the source. The latter is equal to group velocity, which can be evaluated according to the environmental variable, such as water depth, sound profiles, bottom quality, etc. In fact, measurement variables can be substituted by the sound speed in water over the bottom; the error is no more than 8 percent.

MONDAY EVENING, 2 DECEMBER 2002 CORAL ISLAND 1 AND 2, 7:00 TO 9:00 P.M.

Session 1eID

Interdisciplinary: Tutorial on Architectural Acoustics

Ralph T. Muehleisen, Chair

Department of Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder, Colorado 80309

1eID Tutorial on architectural acoustics. Neil Shaw (Menlo Scientific Acoustics, Inc., Topanga, CA 90290) Rick Talaske (The Talaske Group, Inc., Oak Park, IL 60301), and Sylvio Bistafa (Dept. of Mechanical Engineering, Polytechnic School, University of Sao Paulo, Brazil)

This tutorial is intended to provide an overview of current knowledge and practice in architectural acoustics. Topics covered will include basic concepts and history, acoustics of small rooms (small rooms for speech such as classrooms and meeting rooms, music studios, small critical listening spaces such as home theatres) and the acoustics of large rooms (larger assembly halls, auditoria, and performance halls).

Session 2aAAa

Architectural Acoustics: Scattering Topics in Room Acoustics

Richard H. Campbell, Cochair
Bang-Campbell Associates, Box 47, Woods Hole, Massachusetts 02543

Carlos Alejandro Bidondo, Cochair
Freire 3766, Capital Federal, Buenos Aires CP 1429, Argentina

Chair's Introduction—8:00

Invited Paper

8:05

2aAAa1. The rediscovery of diffuse reflection in room acoustics prediction. Bengt-Inge Dalenbäck (CATT, Mariagatan 16A, SE-41471 Gothenburg, Sweden, bid@catt.se)

Around 1986 ''hybrid methods'' appeared where specular ray tracing was used to speed up image source validation and were implemented in a range of software (sometimes in the form of specular cone tracing). Unfortunately, these methods neglected diffuse reflection and room acoustics prediction suffered. Consequently, most software that started out with such methods have today either changed their algorithms to incorporate diffuse reflection or use them only to predict the early part of the echogram. A possible reason for the popularity of these methods was the more detailed point-receiver echogram, but already in 1980, Kuttruff published an *Acustica* paper clearly indicating the importance of handling diffuse reflection. Nevertheless, software exist today, and are presented in journals and doctoral theses, that neglect diffuse reflection. One reason could be that diffuse reflection does not *always* need to be taken into account since when rooms act *geometrically mixing* or have an *even absorption distribution* a specular-only prediction may suffice. However, if none of these conditions are met the RT may be severely overestimated while, on the other hand, Sabine/Eyring formulas will severely underestimate the RT. This paper presents examples where neglecting diffuse reflection leads to large prediction errors, especially regarding the RT.

Contributed Papers

8:25

2aAAa2. Perceptibility of double-slope reverberation decays. Derrick P. Knight (Jaffe Holden Acoustics, Inc., 114A Washington St., Norwalk, CT 06854, dknight@jhacoustics.com), Yasushi Shimizu, and Rendell R. Torres (Rensselaer Polytechnic Inst., Troy, NY 12180)

No concert hall has a perfectly diffuse field, although many are close enough that their decay is perceived as linear. In recent years, concert hall acousticians have taken steps to ensure more exaggerated double-sloped (nonlinear) decays in their concert halls by using coupled volumes. Some acousticians feel that a coupled volume gives a hall a balance between clarity (subjectively speaking) and reverberance. However, there have been no studies done to determine when a nonlinear decay becomes perceptibly different from a linear decay. This work seeks to identify the threshold of perception for nonlinear decays. Nonlinear impulse responses of different lengths are generated by first computing uncoupled impulse responses of a concert hall and a coupled volume in CATT-Acoustic. The two linear impulse responses are convolved in MATLAB. These convolved impulse responses are manipulated to systematically vary the degree of nonlinear decay. The various nonlinear impulse responses are then convolved with anechoic signals with different temporal characteristics and presented to listeners for evaluation. From these evaluations, a criteria is derived to determine when a nonlinear decay becomes audibly different from a linear decay to a listener for various representative signals.

8:40

2aAAa3. Effects of size and density of diffusers on scattering coefficients measured in a 1:10 reverberation chamber. Jin Yong Jeon and Sung Chan Lee (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea)

The degree of diffusion, or scattering coefficient, in surface materials has been known to be one of the most important aspects of the acoustical qualities of concert halls. It has also been recognized that one of the best methods, which can reduce the errors in calculating the reverberation time and other acoustic parameters through computer modeling, is to calculate the scattering coefficient of surface materials. Based on the suggested ISO method, which measures the random-incidence scattering coefficient of surfaces in a diffuse field, the scattering coefficients of different size and density of wooden hemispheres were measured in a 1:10 reverberation room. As a result, the 17.5 cm hemisphere (real size) has the maximum scattering coefficient. It was also found that the scattering coefficient was growing higher when the coverage area increased from the center of the base plate and when the diffuse density increased until the density reached about 50%. Ceramic tiles designed by the calculations of scattering coefficients have been installed for the sidewalls of a 400-seat concert hall.

8:55

2aAAa4. Scattering and/or diffusing elements in a variety of recently completed music auditoria. Ronald L. McKay (McKay Conant Brook, Inc., 5655 Lindero Canyon Rd., Ste. 325, Westlake Village, CA 91362, rmckay@mcbinc.com)

Architectural elements which provide effective acoustic scattering and/or diffusion in a variety of recently completed auditoria for music performance will be presented. Color slides depicting the various elements will be shown. Each will be discussed with respect to its acoustic performance and architectural logic. Measured time-energy reflection patterns will be presented in many cases.

Invited Paper

9:10

2aAAa5. Room acoustic computer modeling: The effects of scattering on transient power flow. Richard H. Campbell (Elec. and Computer Eng., Worcester Polytechnic Inst., 100 Institute Rd., Worcester, MA 01609)

The CATT© room acoustic modeling program allows the computation of acoustic power projected to three mutualy perpendicular planes at a point in space integrated over a specified time period. By selecting successive time periods it is possible to observe the oscillatory nature of the transient power flow from first arrival to the long-term settled value during diffuse-field reverberant decay. The computation is normalized, and displayed as a fraction of the total power. Introducing scattering on the room surfaces results in a significantly smaller peak amplitude of power flow oscillation. Several cases are shown involving different room sizes and shapes and different orientations of source and receiver. Conclusions are made on the number of mean-free path intervals required to ''settle'' the computer-modeled acoustic field.

Contributed Papers

9:30

2aAAa6. Midfrequency modeling of coupled-room performance spaces by computation of transmission coefficients of apertures and arrays of apertures. Jason E. Summers, Rendell R. Torres, and Yasushi Shimizu (Prog. in Architectural Acoust., Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180)

The shapes of the energy decays in a coupled-room system are functions of the energy transmitted by the coupling apertures relative to the irradiation strengths on the boundary surfaces of the rooms. At lower frequencies, the effective coupling areas depart from the geometrical aperture areas as the transmission coefficients of the apertures depart from unity. In order to better predict energy decays in coupled-room systems, transmission coefficients of apertures and arrays of apertures are computed. Power transmission coefficients are calculated for a number of source positions and numerically integrated to yield a random-incidence transmission coefficient. Computed aperture transmission coefficients are compared with analytical frequency-domain solutions for the transmission by circular and rectangular apertures. High-frequency statistical and geometrical models are modified to be applicable at midfrequencies by adjusting the coupling areas in the models according to their computed transmission coefficients. From these models the significance of nongeometrical aperture transmission on the reverberant decays in coupled-room systems is evaluated. Finally, the validity and low-frequency limits of these midfrequency models are evaluated by comparison with measurements of energy decay in actual coupled-room systems. [Research supported by the Bass Foundation.]

9:45

2aAAa7. Scattering from faceted surfaces in optimized room acoustics computations. Rendell R. Torres (Prog. in Architectural Acoust., Rensselaer Polytech. Inst., 110 8th St., Troy, NY 12180-3590), U. Peter Svensson (Norwegian Univ. of Sci. and Technol. (NTNU), Trondheim, Norway), and Nicolas de Rycker (Rensselaer Polytech. Inst., Troy, NY 12180-3590)

To minimize the computational demands of including scattering in auralization, it is appropriate to study how many orders of scattering are necessary. For this purpose, studying edge diffraction is especially appropriate as an elementary form of surface scattering. In a previous study [Torres *et al.*, J. Acoust. Soc. Am. **109**, 600–610 (2001)], it was found that higher orders and combinations of edge diffraction components were not usually as significant as first-order diffraction components. The primary reason was that the reference geometry (a large concert-hall stagehouse) was conservatively composed of large flat walls with dimensions larger than most wavelengths of interest. In that case, significant edge-diffractions occurred at relatively low frequencies (below about 150 Hz). Other realistic reflecting surfaces in rooms, however, also include smaller-scale surface irregularities, e.g., facets for which higher-frequency wavelengths are typically a similar order or larger. This study examines a smaller test geometry consisting of reflector panel arrays similar to those found in concert halls, and we compare computations with various orders of diffraction. Studies of diffraction order are done to determine when inclusion of higher orders is necessary or may be neglected for applications such as interactive auralization.

10:00

2aAAa8. Room acoustics auralization studies of laterally incident boss-model scattering. Rendell R. Torres (Prog. in Architectural Acoust., Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180-3590, rrtorres@rpi.edu), Mendel Kleiner (Chalmers Univ. of Tech., SE-41296 Gothenburg, Sweden), and Georgios Natsiopoulos (Akustikon, SE-411 02 Gothenburg, Sweden)

Computation of room impulse responses by commercial auralization programs typically employs simplified Lambert models to account for nonspecular scattering. However, Lambert scattering models, developed primarily for diffuse lighting computation, have limited validity in acoustics. Instead, in this study, surface scattering from hemispherical bosses on walls is computed with numerical models based on exact classical solutions. Scattering impulse responses for auralization are calculated with various configurations of bosses on infinite planes on both sides of a binaural receiver (approximately 7–10 m away). The physical sound field is analyzed along with the resulting subjective effects of varying boss size and scatterer density.

10:15

2aAAa9. Acoustic diffusers III. Alejandro Bidondo (A. B. Ingenieria de Sonido, Argentina)

This acoustic diffusion research presents a pragmatic view, based more on effects than causes and 15 very useful in the project advance control process, where the sound field's diffusion coefficient, *sound field diffusivity* (SFD), for its evaluation. Further research suggestions are presented to obtain an octave frequency resolution of the SFD for precise design or acoustical corrections.

Session 2aAAb

Architectural Acoustics: Noise Isolation

Brandon Tinianov, Chair
Acoustical Laboratory, Johns Manville, 10100 West Ute Avenue, P.O. Box 625005, Littleton, Colorado 80162

Contributed Papers

10:45

2aAAb1. Prospects for a test procedure for rating floor toppings on joist floors. Alf Warnock (Natl. Res. Council Canada, M59, 1200 Montreal Rd., Ottawa, ON K1A 0R6, Canada, alf.warnock@nrc.ca)

Recently ASTM issued a new test method, E2179, for rating the reduction in impact noise when a floor topping is placed on a concrete slab. E2179 is essentially the same as ISO140-8. Measurements have shown that the reduction in impact noise obtained when a floor topping is placed on a concrete slab is not the same as when the topping is placed on a wooden subfloor supported on joists. To complicate matters, the improvement obtained for a given topping changes when the construction of the joist floor on which it is placed changes. Thus a standard test method with a rigorous definition of a reference joist floor might be created but the measured improvements would not necessarily be applicable to joist floors of different construction. The ISO and ASTM single number ratings for some joist floor systems with toppings do not always give good agreement because of the 8 dB rule that exists in ASTM E989 but is absent in ISO 717-2. This paper will review some of the activities in ASTM and ISO and present some measurement resuts obtained in several laboratories.

11:00

2aAAb2. Sound reduction by simple walls of brick and concrete. Rafael A. C. Laranja and Alberto Tamagna (UFRGS–DEMEC–GMAp, Rua Sarmento Leite, 425–Porto Alegre-RS-Brazil, 90050-110, raclaranja@hotmail.com)

Sound insulation can be difficult to forecast. In most cases it is necessary to take certain precautions in order to avoid degradation in the acoustical performance of walls. Knowing that sound transmission through walls depends on mass per unit area, bending stiffness, damping, mounting conditions, frequencies, etc., the sound transmission can be explained theoretically by several hypotheses. Even though sound transmission has been investigated for more than 100 years, there still remain a lot of issues to be solved. The purpose of this work is to present a discussion regarding the main analytical models used in sound reduction by simple walls, to help the engineer in noise control projects for industry or homes. Several model cases were analyzed by numerical estimation, and the best results were selected to perform experimental comparisons. Some comparative graphics are presented to compare numerical and experimental results. [Work supported by CAPES.] (To be presented in Portuguese.)

11:15

2aAAb3. Measurement and simulation of brick wall sound reduction index. Dinara X. da Paixao (Civil Eng., Federal Univ. of Santa Maria, Santa Maria, RS, Brazil, CEP: 97105-900), Elias B. Teodoro (Federal Univ. of Uberlandia, Uberlandia, MG, Brazil), and Samir N. Y. Gerges (Federal Univ. of Santa Catarina, Florianopolis, SC, Brazil)

This paper presents the procedures and results of an experimental, numerical simulation and analytical model of sound reduction index of solid brick wall. The solid brick used has the following dimensions: 22 $\times 10 \times 5$ cm and density of 1850 kg/m^3. The physic-mechanical characteristics: density, resistance, and elasticity modulus were measured using a set of small walls (0.60×0.63 m). One wall of $4.10 \times 3.20 \times 0.10$ m (length/height/thickness) was built between two reverberant rooms, with 60 and 63 cubic meters in volume, respectively. Two kinds of junctions were used between the wall and the concrete walls of the reverberant rooms: elastic junction (with a thin layer of rubber, sealed with silicon rubber on both sides and the top of the wall) and rigid junction (using the normal mortar). The sound reduction index was measured according to ISO 140-3 using airborne transmission. The brick characteristics were used as input data for the statistical energy analysis (SEA) model. The results from SEA, analytical, and measured models are compared.

11:30

2aAAb4. Active control of the headpiece of noise barriers. Michael Möser and Hyo-In Koh (Inst. of Tech. Acoust., Tech. Univ. Berlin, Einsteinufer 25, 10587 Berlin, Germany, m.moeser@tu-berlin.de)

A theoretical and numerical study of the sound field around an active controlled edge of a noise barrier is described. Studies about noise barrier with passive methods have shown improvement in reduction of the sound level in the range of middle and high frequencies. In this work the sound field on the surface of the headpiece is controlled in such a way that the tangential power transport parallel to their surface is lowered by means of secondary sound sources. First the sound field on the surface of the cylindrical headpiece is locally minimized. The effect of the control position and number of the control point on the sound field near the headpiece and in far field at various frequencies is discussed. The radiated power in the far field of the shadow region is globally minimized, which is the target to reduce, and compared to previous results. Afterwards the method of modal restructing is shown whereby each modal amplitude is controlled individually. Theoretical computations show improved levels in the shadow zone. Optimal active control methods are discussed and a practically oriented simulation is planned.

11:45

2aAAb5. The acoustical performance of building façades designed for warm and humid climates. Elcione Maria Lobato de Moraes (Univ. of the Amazonia, Brazil, elcione@hotmail.com)

The effects of noise on man are known to range from simple nuisances to illnesses and trauma. Studies of acoustic isolation provided by buildings in urban centers assume that external openings such as windows and doors remain closed. However, this paper presents results of a study of the acoustic isolation of airborne noise for different types of façade construction in the city of Belén, Brazil. The study recognizes that windows and doors are opened as a function of the local climatic conditions since the city is located in a warm and humid tropical area. The purpose of the investigation is to provide information to planners so that they can better control acoustic contamination of buildings located in these kinds of climates. It is concluded that acoustic isolation is affected by the façade construction and by openings in the façade.

Session 2aAB

Animal Bioacoustics: Amphibian Bioacoustics in Honor of Robert Capranica I

Peter M. Narins, Chair
Department of Physiological Science, University of California, Los Angeles, California 90095-1606

Chair's Introduction—8:00

Invited Papers

8:10

2aAB1. Bob meets bullfrogs: Love at first croak. Moise H. Goldstein, Jr. (Marine Biol. Lab., Woods Hole, MA 02543)

After graduating from UC Berkeley, young Bob took at job at Bell Labs (BTL) and he and his lovely wife Patricia moved east. Somehow he found himself in a ''wet'' biological laboratory hidden in a corner of the BTL Murray Hill campus. Here Bob found Larry Frishkopf and the author, a MIT summer visitor at BTL, doing unit recording from bullfrog auditory nerves. Bob was pleased to help with the project. Sometimes, we would speculate about the tie between the electrophysiology and bullfrog behavior. As summer ended, Bob competed for and won a BTL Ph.D. fellowship. Bob chose to do his dissertation research at MIT and to address that question. I was his preceptor. He developed a novel method, sound evoked vocalizations, and skillfully demonstrated the relation of neural coding and behavior in bullfrogs. Soon after competing his dissertation research, Bob accepted a faculty position at Cornell. It is a great pleasure for me to take part in honoring Bob's brilliant career.

8:25

2aAB2. Neural mechanisms of call recognition in Pacific treefrogs. Gary J. Rose (Dept. of Biol., Univ. of Utah, Salt Lake City, UT 84112)

Pacific treefrogs, Hyla regilla, use advertisement and encounter calls in their acoustic communication. These two call types differ primarily in the rate that pulses are repeated; at 17 °C, pulse repetition rates (PRRs) for advertisement and encounter calls are about 100 pulses/s and 30 pulses/s, respectively. Behavioral studies indicate that two discrete channels exist for processing these communication signals. Because sound pulses in these two call types are highly similar in spectral and temporal structure, the PRR selectivity of the advertisement filter is likely to result from an analysis of interpulse interval. Behavioral studies support the hypothesis that temporal integration might underlie this analysis. As a neural correlate to the behavioral studies, toral neurons have been recorded that respond selectively to PRRs found in the advertisement call. These cells fail to respond to encounter call PRRs. In support of the integration hypothesis, these units only respond after a threshold number of correct interpulse intervals have been presented; the average PRR is largely irrelevant. The interval-counting process can be reset by a single inappropriate interval. In the most selective cases, a single interval that is slightly longer or shorter than the optimal interval can reset the integration process. [Work supported by NIDCD.]

8:50

2aAB3. Nonlinearity in central auditory processing. Albert S. Feng (Beckman Inst., Univ. of Illinois at Urbana–Champaign, 405 N. Mathews, Urbana, IL 61801)

The pioneering work of Robert Capranica in sound communication in bullfrogs revealed that sound pattern recognition involves a nonlinear operation. He showed in 1965 that, for bullfrogs, although the peripheral auditory system exhibits nonlinear properties, nonlinearity of the central auditory system is necessary for sound perception. Researchers studying sound processing in anurans as well as in birds and mammals have since shown that central auditory processing underlying various perceptual tasks (e.g., localization, ranging, and pattern recognition) mostly involves nonlinear operations; these are time- or frequency-dependent or both. Signal detection and discrimination in noise is another example. Frogs exhibit spatial unmasking, a space-dependent phenomenon, i.e., the masking effect is reduced when signal and noise sources are spatially separated. Spatial unmasking allows frogs to communicate effectively by sound in a dense chorus. Recent physiological evidence showed that spatial unmasking is largely attributed to nonlinear binaural processing in the CNS. The auditory periphery contributes to spatial unmasking but the processing therein is largely linear and its role is more limited. [Work supported by NIH/NIDCD.]

9:15

2aAB4. A reexamination of the teleost swimbladder as an acoustic source. Michael L. Fine (Dept. of Biol., Virginia Commonwealth Univ., Richmond, VA 23284-2012, mfine@atlas.vcu.edu)

Classically swimbladders are considered pulsating resonant bubbles that are omnidirectional monopole sources capable of translating acoustic pressure to the ears. Swimbladder sounds are driven by sonic muscles (the fastest in vertebrates), yet high speed would seem unnecessary to excite a resonant structure. Recent studies in the oyster toadfish Opsanus tau and the weakfish Cynoscion regalis suggest that the classic generalizations may not apply to all fishes. The toadfish bladder is a low-Q inefficient (highly-damped) resonator that moves in a quadrupole fashion producing sound amplitude proportional to bladder velocity. Slow movements do not

produce audible sound. When the bladder is stimulated artificially, dominant frequency is determined by the waveform of the excitation stimulus and not fish size. The sound field, measured underwater, is mildly directional, and deflation of the bladder does not change auditory thresholds. The dominant frequency of weakfish sounds (pulses produced by individual muscle contractions) appears to be determined by timing parameters of muscle twitches and not the natural frequency of the bladder. [Work supported by NIH.]

9:40

2aAB5. Eighth-nerve regeneration, 20 years later. Andrea M. Simmons, Judith A. Chapman, and Alison Barnstable (Dept. of Psych., Brown Univ., Providence, RI 02912, Andrea_Simmons@brown.edu)

Over 20 years ago, Harold Zakon and Bob Capranica began a series of studies tracing the anatomical and physiological sequelae of damage to the eighth cranial nerve in ranid frogs. Their results indicated that the severed nerve could regenerate back into the brain, that proper anatomical connectivity was re-established, and that appropriate neural responses in target nuclei re-emerge. In an extension of these studies, patterns of gene expression in the crushed nerve and in target brainstem nuclei are being examined at various time points after damage. These molecular changes are correlated with the time course of recovery of normal vestibular function. Passive (head tilt) and active (response to rotation) vestibular function recover at different rates, with response to rotation still abnormal at 12 weeks after damage. This delayed recovery indicates a process of regeneration, rather than behavioral compensation. Differential display, a PCR-based technique, is used to elucidate which mRNAs are expressed at different time points following nerve crush. At periods less than 1 week after damage, mRNAs, both novel and previously identified, are differentially expressed in crushed versus sham-operated tissue. One gene product, expressed at 24 h post-crush, exhibits 91% coding region homology to human brain-specific protein, CGI-38.

10:05–10:25 Break

Contributed Papers

10:25

2aAB6. Auditory organs in *Xenopus laevis* and *Xenopus tropicalis*. Elba E. Serrano and Quincy A. Quick (Biol. Dept., NMSU, Las Cruces, NM 88003)

The genus *Xenopus* is comprised of over 20 genetically divergent species that occupy aquatic habitats such as lakes, rivers, and swamps south of the African Sahara. Like other anurans, advertisement calls are important signals for reproductive behavior and *Xenopus* have a vocal apparatus adapted for underwater sound production. The tetraploid *Xenopus laevis* is a well-established model organism for cell and developmental biology. *Xenopus tropicalis*, a diploid member of the *Silurana* group, is a newer *Xenopus* model for molecular genetics that is suited for transgenic studies. Both species are being used for investigations of inner ear organogenesis. Results presented here use light and confocal microscopy to examine the structure of the auditory organs of *X. laevis* and *X. tropicalis*. Images gathered from sectioned tissue stained with hematoxylin/eosin or microdissected organs labeled with Alexa 488 phalloidin (Molecular Probes) and propidium iodide illustrate the organization and innervation of the sensory epithelia of the sacculus, amphibian papilla and basilar papilla of the two species. The data show the similarities between the sensory fields and highlight the size differences of the organs in the two species. [Work supported by grant to EES (NIGMS, NIDCD) and awards to QQ (NASA NMSGC, NIGMS RISE).]

10:40

2aAB7. Auditory nerve recordings in *Xenopus laevis*. Darcy Kelley, Taffeta Elliot (Dept. of Biol. Sci., Columbia Univ., New York, NY 10027), and Jakob Christiasen-Dahlsgaard (Odense Univ.)

The South African clawed frog *X. laevis* communicates using a rich repertoire of underwater calls made up of clicks. Two calls, rapping and ticking, differ only by interclick interval [Tobias *et al.*, PNAS (1998)]. To begin to explore rate coding in the auditory system, extracellular recordings were obtained *in vivo* from the auditory nerve of frogs stimulated with both pure tones and recorded calls. To avoid problems of impedance matching, acoustic stimuli were delivered via vibration-excited probe directly to the frog's tympanum. Preferred frequencies across the population of 104 recorded fibers (taken from 5 females) fell into 2 or 3 groups: 1500, 600, and 200 Hz. Cells of all preferred frequencies fired in phase with clicks of the various calls. [Work supported by an NSF graduate fellowship to TME.]

10:55

2aAB8. Suppression of frog amphibian-papillar (AP) axons. Edwin R. Lewis (Dept. of EECS, Univ. of California, Berkeley, CA 94720, lewis@eecs.berkeley.edu), Pim van Dijk (Univ. Hospital Maastricht, The Netherlands), and Water M. Yamada (Univ. of Southern California, Los Angeles, CA 90089)

In response to white-noise stimuli, AP afferent axons of *Rana catesbeiana* and *R. esculenta* exhibit excitation, adaptation, and suppression simultaneously. After eigendecomposition of the second-order Wiener kernel, spectrotemporal properties of suppression and adaptation, in the individual axon, can be reconstructed by short-term averaging taken parallel to the main diagonal of the inhibitory subkernel. The results show suppressive spectra with deep, sharp notches in the vicinity of BEF. Spectral components outside the notch, on both sides of BEF, are effective in eliciting the suppression response; those very close to BEF are not. The results imply that the suppressive response (a negative dc shift in instantaneous spike rate) may occur with or without the presence of excitatory stimuli. It will be visible, however, only if there is background spike activity. The suppressive spectrum on the high side of BEF can extend beyond the highest AP BEF. In other words, AP axons can respond (negatively) to stimuli at frequencies beyond the highest BEF. The results further imply that a brief suppressive stimulus will be most effective if applied toward the end of a brief excitatory stimulus, or slightly after that excitatory stimulus has ended.

11:10

2aAB9. Amplitude modulation encoding in different auditory nuclei of the frog. G. Nikolay Bibikov (N. N. Andreyev Acoust. Inst., Schvernik St. 4, Moscow 117036, Russia)

Since R. R. Capranica published his pioneer articles dealing with the amplitude modulation (AM) encoding, we have explored the reproduction of sine AM in a thousand single units of the medullar, isthmal, and midbrain acoustical centers of the European frogs. Typically, it was phasic units that demonstrated the best phase-locking to the 80%–100% modulated tone. However, onset units usually did not reproduce modulation with low modulation indexes (10%–20%). Medullar tonic units moderately reproduced deep AM and scarcely reproduced the 10% AM. However, the population response to a simple AM tone demonstrated a reliable reproduction of 10% modulated signal and a slight tendency to increase phase-locking from the first to the last modulation period. In the superior olive and lateral lemniscus nuclei this tendency became more evident. In torus semicirdularis a large population of the units increase their phase-locking response during the short-term (tenth of second) and long-term (tens of seconds) adaptation. In many units a dramatic enhancement of the

mean firing rate in the sustained state was observed after noise addition. Morever, the amplitude of the sine modulation of the instantaneous spike rate was enhanced on some levels of the noise. [Work supported by the RFBR Grant No. 02-04-48236.]

11:25

2aAB10. Multiple components in distortion product otoacoustic emissions from the amphibian ear. Sebastiaan W. F. Meenderink and Pim Van Dijk (Dept. of Otorhinolaryngology and Head & Neck Surgery, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands)

Distortion product otoacoustic emissions (DPOAEs) are weak intermodulation distortions generated in the inner ear in response to two stimulus tones. Both the amphibian papilla and the basilar papilla in the inner of the frog may generate DPOAEs [Van Dijk and Manley, Hear. Res. **153**, 14–22 (2001)]. Here, we measured the level and phase of DPOAEs in the Leopard frog, *Rana pipiens*, in response to stimulus tones between 40- and 90-dB SPL. Results show that for stimulus tones in the amphibian papilla frequency range, two components contribute to DPOAEs. One component dominates for stimulus levels below about 70-dB SPL, while the other is most prominent at higher levels. The transition between both components is accompanied by a conspicuous phase change, and sometimes by a notch in the amplitude response curve. Similar results were obtained in the basilar papilla frequency range, but in addition a third component was present for stimuli below about 55-dB SPL. With the exception of this third component, our findings are remarkably similar to those in mammals, despite the structural differences between the mammalian and amphibian inner ear. [Work supported by NWO.]

TUESDAY MORNING, 3 DECEMBER 2002 CORAL GARDEN 2 AND 3, 8:00 A.M. TO 12:00 NOON

Session 2aAO

Acoustical Oceanography: Sensing the Basin Scale

Peter F. Worcester, Chair
Scripps Institute of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0225

Contributed Papers

8:00

2aAO1. Observations of internal tide variability in the far field of the Hawaiian Ridge: The far field component of the Hawaii Ocean Mixing Experiment (HOME). Brian Dushaw (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105, dushaw@apl.washington.edu), Peter Worcester, Matthew Dzieciuch (Univ. of California, San Diego, San Diego, CA 92093-0225), and Doug Luther (Univ. of Hawai'i at Manoa, Honolulu, HI 96822)

As part of the Hawaii Ocean Mixing Experiment (HOME), observations of internal tides in two regions on either side of the Hawaiian Ridge were obtained by tomography, thermistors, and CTD casts from FLIP. The tomographic observations detect radiation of low internal-tide modes in broad areas, while the thermistors and CTD casts measure the ''local'' internal-tide variability. These observations are used to estimate the amount of energy carried away from the Ridge by the internal tides, to estimate the relative energies of low- and high-mode internal tides, and to test numerical models of internal-tide generation. Barotropic currents and pressure were also measured by tomography, electromagnetic, and pressure sensors so that, with careful modeling, the energy lost from the barotropic tides at the Ridge can be determined. Thermistor data obtained on one mooring showed that the M2, mode-1 internal tide was mainly phase-locked and carried 1.3 kW/m of energy. Modes 2 and 3 had amplitudes comparable to mode-1, but they were not phase locked. Energy fluxes at three other moorings were 1.2, 2.0, and 6.7 kW/m. Energy fluxes obtained by tomography were 0 (1 kW/m) or less; the line-integral data are less susceptible to the interference effects in the outgoing internal waves.

8:15

2aAO2. Acoustic thermometry along an Arctic Ocean path. Peter Mikhalevsky, Brian Sperry (SAIC, 1710 SAIC Dr., McLean, VA 22102), and Alexander Gavrilov (P. P. Shirshov Inst. of Oceanology, Moscow 117851, Russia)

Acoustic thermometry has been shown to be a very effective technique for monitoring average heat content and average temperature in the Arctic Ocean and in particular in the Arctic Intermediate Water (AIW) layer. As part of the U.S./Russian Arctic Climate Observations using Underwater Sound (ACOUS) program a 14-month time series of acoustic transmissions were analyzed along a 1250 km propagation path that extended from the Franz Victoria Strait to the Lincoln Sea from Oct. 1998 through Dec. 1999. The receive array mooring in the Lincoln Sea was recovered in April 2001. Modal travel times were estimated after pulse compression processing and mode filtering of the vertical line array. The interarrival time between mode 1 and modes 2 and 3 show net cooling during the first several months followed by a dramatic warming of the AIW along the propagation path. This warming is consistent with direct CTD measurements made along a central Arctic transect performed by the USS Hawkbill during the Scientific Ice Exercise (SCICEX) 2000. [Work supported by ONR, NSF, the Civilian Research and Development Foundation, and the Ministry of Industry, Science and Technology of the Russian Federation.]

8:30

2aAO3. Long-range detection of hydroacoustic signals from large Antarctic icebergs. Jacques Talandier, Olivier Hyvernaud, Pierre F. Piserchia (Departement Analyse, Surveillance, Environnement du Commissariat a l'Energie Atomique, BP 12, 9680 Bruyeres-Le-Chatel, France), and Emile A. Okal (Northwestern Univ.)

T-waves are commonly observed on coastal seismographs of the French Polynesian Seismic Network (RSP), when an oceanic earthquake or an underwater explosion occurs, even for small events. *T*-waves are trapped in the underwater channel and can propagate at very long distances before being converted into seismic waves close to the coastal seismic stations. During the 2000/2001 Austral summer, coastal seismic stations of the RSP detected unique series of *T*-waves from Antarctica about 60 away in the frequency band 2–15 Hz. Some of them last a few minutes while other wavetrains last several hours; some are broadband while others feature prominent frequencies, occasionally accompanied by overtones. Most of the hydroacoutics sources are relocated using the RSP stations and some Antarctic seismographs. It is shown that observed waves

have a very long underwater path but may also propagate in the ice sheet. Satellite monitoring demonstrates that hydroacoustic source locations are very well correlated in space and in time with icebergs B-15B and B-17 moving off the Ross Ice Shelf. These two icebergs appeared after the Iceberg B15 broke from the Ross Ice Shelf in March 2000.

8:45

2aAO4. The dynamics of abyssal T-phases. Ralph A. Stephen, Deborah K. Smith (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Clare Williams (MIT/WHOI Joint Prog. in Oceanogr., Woods Hole, MA 02543)

The characteristics of earthquakes, as revealed by T-phase observations, have the potential to provide important constraints on physical models of crustal processes under the oceans. Although it has been postulated that some form of scattering at or near the seafloor is necessary to convert the compressional and shear body waves from earthquakes into the low grazing angle paths necessary for propagation in the ocean sound channel, there are T-phase observations that cannot be explained by seafloor scattering alone. Water depth above the epicenter, for example, should have a strong effect on T-phase excitation. We use the time domain finite-difference method combined with ray theory to demonstrate these issues and we compare the theory to a series of events that occurred near the mid-Atlantic Ridge at the Kane Fracture Zone (MARK) in 1999 and 2000. There is evidence in this data set which suggests that topographic steering of T-phase locations occurs. Earthquake energy appears to preferentially enter the sound channel at topographic highs and epicentral locations are biased toward shallow bathymetry.

9:00

2aAO5. Propagation of sound through a spicy ocean: Analysis. Walter Munk, Matthew Dzieciuch, and Daniel Rudnick (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093)

We derive some of the parameters for a simple model of spice and internal wave scatter, using the canonical ocean sound channel. A distant goal is for long-range acoustic transmissions to provide a measure of upper ocean stirring from abyssal ocean acoustic signatures.

9:15

2aAO6. Propagation of sound through a spicy ocean: A numerical experiment. Matthew Dzieciuch, Walter Munk, and Daniel Rudnick (Scripps Inst. of Oceanogr., Univ. of California, San Diego, CA 92093)

Using a closely sampled 1000 km hydrographic section in the eastern North Pacific, we separate the sound-speed finestructure into 2 component fields: (i) isopycnal tilt dominated by internal waves; and (ii) ''spicy'' (hot and salty) millifronts associated with upper ocean stirring. Scattering by the spicy millifronts is of the same order as internal wave scattering (the traditional view), and they both contribute to the penetration of sound into the shadow zone.

9:30

2aAO7. Travel time bias at sound reflection from an uneven surface: Implications for ocean remote sensing. Oleg A. Godin (CIRES, Univ. of Colorado and NOAA/Environ. Technol. Lab., Boulder, CO 80305) and Iosif M. Fuks (ZEL Technologies, LLC and NOAA/Environ. Technol. Lab., Boulder, CO 80305)

Travel time is the acoustic quantity which is most frequently used to infer physical properties of the ocean or its bottom from acoustic measurements. Ocean boundaries typically have a complicated shape with smaller-scale features normally described as random roughness. The roughness translates into fluctuations of reflected acoustic arrivals. In this paper, statistical properties of acoustic travel time are considered in homogeneous and a vertically-stratified 3-D ocean with rough boundaries and interfaces. On average, the rough surfaces can be either horizontal planes, corresponding to the ocean surface, or deterministic curved surfaces, corresponding to the sea floor or interfaces within the ocean bottom. It is shown that mean acoustic travel time differs from the travel time in an average medium. In particular, in agreement with Fermat's principle, small cross-range slopes of the reflecting surface always decrease ray travel time. The travel time bias is studied using the ray and the adiabatic normal mode theories. Implications are analyzed of the travel time bias on interpretation of measurements made with echo sounders, inverted echo sounders, and tomography systems employing surface- and/or bottom-reflected arrivals. [Work supported by ONR.]

9:45

2aAO8. Comments on "ray chaos" and ocean acoustic tomography. Brian Dushaw (Appl. Phys. Lab., Univ. of Washington, 1013 N.E. 40th St., Seattle, WA 98105) and John Colosi (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Recent publications describe ''ray chaos'' in the context of long-range acoustic propagation in the ocean. This work is relevant to long-range tomography, which relies on the identification of specific ray paths with pulse arrivals. However, the ''ray chaos'' work has mainly had a theoretical or numerical focus. Of necessity, artificial assumptions have been made as to internal-wave spectrum, infinite acoustic frequency, background sound speed profile, etc. The discussions of ''ray chaos'' have rarely incorporated the results of the long-range propagation experiments that have been conducted over the past 15 years. In spite of this disassociation between theory/numerics and experiment, discussions involving ''ray chaos'' have been critical of tomography/thermometry. One claim is that acoustic rays employed for purposes of tomography are an inappropriate description of the acoustic sampling associated with measured arrival patterns. Such criticism has occasionally implied that tomography cannot be employed for ocean studies at ranges larger than about 1 Mm as a result of ''ray chaos'' and other issues. Some aspects of the ''chaos'' view are obviously correct, while other aspects require a more rigorous test of modeling by experiment. However, the oceanographic measurements that result from using either ''classical'' or ''chaotic'' rays are practically indistinguishable.

10:00–10:15 Break

10:15

2aAO9. Discrepancies between ocean-acoustic fluctuations in parabolic-equation simulations and estimates from integral approximations. Michael D. Vera (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0225) and Stanley M. Flatté (Univ. of California, Santa Cruz, Santa Cruz, CA 95064)

Analytic, line-integral approximations to the acoustic path integral have been used to estimate the magnitude of internal-wave-induced fluctuations in a signal traveling through the ocean. These approximations for the bias and variance of travel time, the length scale of acoustic coherence in depth, and the spreading in time of acoustic intensity peaks are compared, in this discussion, to values from simulations that used the standard parabolic equation. Two different temperate-latitude sound-speed profiles were used in simulated 250 Hz acoustic propagations with a maximum range of 1000 km. The sound speed was perturbed by internal waves conforming to the Garrett–Munk (GM) spectral model with strengths of 0.5, 1, and 2 times the GM reference energy level. Though predictions of

the travel-time variance were largely successful, the other quantities examined did not correspond to simulation values. Calculated biases deviated from parabolic-equation results at ranges beyond a few hundred kilometers. The predicted depth-coherence lengths at 1000 km were significantly shorter than those extracted from the simulations. The estimated magnitudes of pulse spreading at 1000 km were much greater than the differences in widths between intensity peaks from simulations with and without internal-wave perturbations.

10:30

2aAO10. Calculations of ocean-acoustic fluctuations for use in tomography of internal waves. Stanley M. Flatté and Michael D. Vera (Phys. Dept., Univ. of California, Santa Cruz, CA 95064)

In order to use observations of the bias and variance of travel time, the length scale of acoustic coherence in depth, and the spreading of the acoustic pulse as information about the internal-wave field in the ocean (all for identifiable rays at long range), a reliable method of calculation of these quantities in the presence of a specific internal-wave model must be made. Attempts to use analytical methods having proven unreliable for long range, except for the variance of travel time; there are two possibilities: (1) restrict tomography of identifiable rays to travel-time variance; or (2) use simulation by parabolic equation to calculate expected values. The former is restricted and the latter is computer intensive. The needed resources for the latter will be discussed.

10:45

2aAO11. Limitations on perturbation theory applied to ocean acoustic inversion. B. Edward McDonald (Naval Res. Lab., Code 7145, Washington, DC 20375), Brian Sperry (SAIC, McLean, VA 22102) and Arthur Baggeroer (MIT, Cambridge, MA 02139)

Perturbation theory for ocean acoustic modal group speed responses to small environmental changes is investigated with regard to its applicability in ocean acoustic tomography. Assuming adiabaticity, the inverse problem for each vertical eigenmode is an integral equation whose kernel involves the eigenfunction and its frequency derivative. A new proof is given for the so-called ''third term problem'' which requires equivalence between two dissimilar forms of the integral equation. Numerical examples are given for the inversion kernel for four types of sound speed profiles, and then the parameter range (amplitude and scale size) in which perturbation theory is accurately examined. It is found that the range of validity is set not only by the amplitude of the perturbations, but also by their vertical scale size. [Work supported by ONR and Saclantcen.]

11:00

2aAO12. Long range nonlinear propagation in an ocean acoustic waveguide. Kaelig Castor, Peter Gerstoft, Philippe Roux, W. A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92037), and B. E. McDonald (Naval Res. Lab., Acoust. Div., Washington, DC 20375)

The Nonlinear Progressive Wave Equation (NPE) [McDonald and Kuperman, J. Acoust. Soc. Am. **81**, 1406–1417 (1987)] is an approximation of the Euler equations for nonlinear compressional waves in an inviscid fluid and, actually, is the nonlinear time domain counterpart of the frequency domain linear parabolic wave equation (PE) for small angle propagation. Simulations using a NPE code were used to study both harmonic (high frequency) and parametric (low frequency) generation. This code was coupled with a linear adiabatic normal mode program which allows range-dependent cases treatment to study propagation in shallow or deep water with longer range propagation paths. Included in the modeling are both shock dissipation and linear attenuation in the sediment layer. The results of these studies are presented.

11:15

2aAO13. Downslope measurements from a bottom mounted tomography source. Kevin D. Heaney (ORINCON Industries, 4350 N. Fairfax Dr., Ste. 470, Arlington, VA 22203), Brett Castille, Arthur Teranishi, and Daniel Sternlicht (ORINCON Industries, San Diego, CA 92121)

Long-range tomography experiments seek to measure the temporal acoustic fluctuations due to thermal changes in the ocean structure. To map these changes to specific depths in the ocean, accurate travel times, and the ability to predict ray paths are required. Both of these become difficult in an environment that is bottom interacting (even though the water may be deep). To examine near source effects of interaction with the bottom for a tomographic source set on the bottom (at a depth of 800 m) measurements were taken off of Kauai. A short VLA was deployed from a small vessel and broadband recordings were taken at 6 ranges, from directly overhead to 55 km away. The issues to be addressed are the accuracy of the source timing and the relative strength of the bottom bounce and direct path. Evidence exists that interaction with the sea-floor can lead to up to a 0.5 s delay in the measured travel time, from that predicted.

11:30

2aAO14. Horizontal refraction and coherence of acoustic signals propagating over a long-range in the ocean. Alexander G. Voronovich, Vladimir E. Ostashev (NOAA/Environ. Technol. Lab., 325 Broadway, Boulder, CO 80305), and The NPAL Group[a)]

The paper is devoted to experimental and theoretical studies of horizontal refraction and coherence of acoustic signals recorded during the NPAL experiment with the use of the billboard acoustic array [The NPAL Group, J. Acoust. Soc. Am. **109**, 2384 (2001)]. Both ray- and mode representations of the acoustic field are used for signal-processing. In the first approach, signals at different pairs of hydrophones located at approximately the same depth are cross-correlated. This allows us to obtain the mean horizontal refraction angle, its time-dependence and variance, and the coherence radius of the sound wavefront. The variance and the coherence radius are also estimated using a theory developed. In this theory, the energy of the sound field scattered by internal waves spreads in horizontal directions according to a diffusion law. The diffusion coefficient is evaluated numerically for the Garret–Munk spectrum and canonical Munk profile. It is shown that the variance of the horizontal refraction angle and the coherence radius of the wavefront calculated theoretically agree qualitatively with their values obtained using the NPAL data. [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.

11:45

2aAO15. NPAL horizontal refraction: RAKE correlator estimates. Matthew Dzieciuch (Scripps Inst. of Oceanogr., Univ. of California, San Diego, CA 92093) and the NPAL Group[a)] (SIO-UCSD, APL-UW, WHOI)

A purpose of the NPAL billboard array data set was to measure the horizontal refraction of low-frequency (75 Hz), long-range (4000 km) timefronts. Simple beamforming yields time series that are too noisy for an accurate estimate of horizontal refraction variability. Since the data show partial horizontal coherence, a RAKE correlator is designed to account for the signal variance across the array and to improve the performance of the linear beamformer. [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.

Session 2aBB

Biomedical Ultrasound/Bioresponse to Vibration: History of Biomedical Ultrasound/ Bioresponse to Vibration

Lawrence A. Crum, Chair
Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105

Chair's Introduction—10:30

Invited Paper

10:35

2aBB1. Origins and evolution of the developments which led to echo-Doppler duplex color flow diagnostic methodology. Donald W. Baker (Dept. of Bioengineering (ret.), Univ. of Washington, Seattle, WA 98105 and 13706 94th Ave. NE, Kirkland, WA 98034-1842, jdwbaker@attbi.com)

Research efforts to develop instrumentation for animal physiologic research to better characterize the cardiovascular system in engineering terms ultimately evolved for application on man and led to the Pacific Northwest becoming the current foci of the medical ultrasound industry. This presentation will trace the events from my being a student in Electrical Engineering to heading the Cardiovascular Instrument Development Program originally begun by Dr. Robert Rushmer in 1957. This narrative will range from early instruments for measurements on research animals to their development for noninvasive use on man. The instruments covered will be the transit time flow-meter, CW Doppler, pulsed Doppler, duplex scanner, color flow mapping. The role of collaboration in both engineering and many specialties of medicine will be demonstrated. Many of the original instruments have been in the Smithsonian Museum of American History and will in the near future be on permanent exhibit there.

Session 2aEA

Engineering Acoustics: Air Acoustics Devices, Techniques and Measurements

Zemar M. D. Soares, Cochair
Electroacoustics Laboratory, INMETRO, Av. N. S. das Gracas 50, Xerem, Rio de Janeiro 25250-020, Brazil

Gilles A. Daigle, Cochair
Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada

Chair's Introduction—8:00

Contributed Papers

8:05

2aEA1. Vibration balanced miniature loudspeaker. David E. Schafer, Mekell Jiles, Thomas E. Miller, and Stephen C. Thompson (Knowles Electronics, 1151 Maplewood Dr., Itasca, IL 60143)

The vibration that is generated by the receiver (loudspeaker) in a hearing aid can be a cause of feedback oscillation. Oscillation can occur if the microphone senses the receiver vibration at sufficient amplitude and appropriate phase. Feedback oscillation from this and other causes is a major problem for those who manufacture, prescribe, and use hearing aids. The receivers normally used in hearing aids are of the balanced armature-type that has a significant moving mass. The reaction force from this moving mass is the source of the vibration. A modification of the balanced armature transducer has been developed that balances the vibration of its internal parts in a way that significantly reduces the vibration force transmitted outside of the receiver case. This transducer design concept, and some of its early prototype test data will be shown. The data indicate that it should be possible to manufacture transducers that generate less vibration than equivalent present models by 15–30 dB.

8:20

2aEA2. Measurement and numerical simulation of the changes in the open-loop transfer function in hearing aid as a function telephone handset proximity. Gilles A. Daigle and Michael R. Stinson (Inst. for Microstructural Sci., Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

The presence of a nearby object (telephone handset, cupped hand, etc.) can cause acoustical feedback to occur in a hearing aid. The object reflects or scatters additional sound energy to the microphone position causing the open-loop transfer function (OLTF) to increase. Feedback can occur when the OLTF$>$0 dB. To investigate this problem, measurements of the OLTF were made for three hearing aids (BTE, ITC, ITE) mounted on a KEMAR manikin. A telephone handset, positioned initially in a typical user posi-

tion, was translated to positions between 0 and 100 mm away from the pinna, repeatibly, using a linear translation system. Changes of up to 15 dB or more were observed as the handset moved, particularly for positions within 20 mm of the pinna. In parallel, numerical simulations were made using a boundary element method. Computed changes in OLTF were consistent with the measured changes.

8:35

2aEA3. Derivation of moving-coil loudspeaker parameters using acoustical testing techniques. Brian E. Anderson (Dept. of Phys., Brigham Young Univ., Provo, UT 84602, loudspeakerdesign@hotmail.com) and Timothy W. Leishman (Brigham Young Univ., Provo, UT 84602)

A novel acoustical method of measuring small-signal moving-coil loudspeaker parameters has recently been developed. This technique involves the use of a plane wave tube to measure acoustical properties (e.g., reflection and transmission coefficients) of a driver under test (DUT). From this data, small-signal parameters are derived using curve-fitting techniques. Current loudspeaker parameter measurement techniques require measurement of the electrical impedance of the DUT. This paper will discuss the acoustical measurement apparatus, system modeling (via equivalent circuits), and a comparison of measured parameters to those derived using electrical techniques.

8:50

2aEA4. New acoustic test facility at Georgia Tech. Van Biesel and Kenneth Cunefare (Georgia Inst. of Technol., Atlanta, GA 30332, van.biesel@me.gatech.edu)

Georgia Tech's Integrated Acoustics Laboratory (IAL) is a state of the art research facility dedicated to the study of acoustics and vibration. The centerpiece of the laboratory is a 24 ft×24 ft×20 ft full anechoic chamber, which has been in operation since 1998. The IAL is currently expanding to include a reverberation room and hemi-anechoic chamber, designed and built by Acoustic Systems. These two chambers will be joined by an 8 ft×8 ft transmission loss opening, allowing for a detailed measurement and analysis of complex barriers. Both chambers will accommodate vehicles and similarly large structures. The reverberation room will have adequate volume for standardized absorption measurements. Each chamber will be equipped with dedicated multichannel data acquisition systems and instrumentation for the support of simultaneous research in all areas of the laboratory. The new test chambers are funded by a grant from the Ford Motor Company and are planned to be completed and fully functional by 1 January 2003.

9:05

2aEA5. PC interface for a stepped filter octave band analyzer. Miguel A. Horta and Marco A. Vazquez (Instituto Politecnico Nacional, Mexico)

Spectrum analyzers are a basic tool for an acoustic engineer. Since almost anyone in a classroom has access to a PC or notebook computer, a low cost alternative for measuring many different parameters with a spectrum analyzer (such as reverberation time, transmission loss, or equivalent noise level) is a digital interface that processes the signal combined with the proper software for the computer. This is a design of a spectrum analyzer that does such a task. The external interface sequentially processes the signal in an octave band filter bank and converts it to data for the computer to display. Depending on the software, the data can be used to measure any of the signal's desired parameters.

9:20

2aEA6. Secondary microphone calibration: Advantages on the use of the constant envelope sweeps. Zemar M. D. Soares, Walter E. Hoffmann (Electroacoustics Lab., INMETRO, Av. N. S. das Gracas 50, Xerem RJ, 25250-020, Brazil, zmsoares@inmetro.gov.br), and Swen Muller (INMETRO, Xerem RJ, 25250-020, Brazil)

With the objective of reducing the costs for the accreditation of secondary laboratories in the electroacoustical area without impairing the quality of the calibration of microphones, the Laboratory of Electroacoustics of INMETRO (Brazil) has investigated the advantages of using sweeps with user-defined spectral distribution and constant temporal envelope to obtain the impulse response between sound source and microphone. In applications in which the sensibility of the microphone in the free field is the objective, the use of anechoic chambers is fundamental. However, they can be substituted by applying a windowing of the impulse response to isolate the direct sound. The same technique can be used to separate reflections in the Jig for calibration of microphones in pressure field, as proposed by IEC61094-5. The calibration process presented here is based on FFT techniques, using a special sweep as excitation signal. The sweep is custom tailored in a way that its energy contents compensates the background noise spectrum. This way, the excitation signals' signal-to-noise ratio becomes independent of frequency, while the sweep keeps an almost constant temporal envelope which contains the maximum possible energy. (To be presented in Portuguese.)

9:35

2aEA7. Free-field calibration of measurement microphones at frequencies up to 80 kHz. Allan J. Zuckerwar and Gregory C. Herring (NASA Langley Res. Ctr., M.S. 493, Hampton, VA 23681, a.j.zuckerwar@larc.nasa.gov)

Civil-aviation noise-reduction programs, that make use of scaled-down aircraft models in wind tunnel tests, require knowledge of microphone pressure (i.e., not free-field) sensitivities beyond 20 kHz—since noise wavelengths also scale down with decreasing model size. Furthermore, not all microphone types (e.g., electrets) are easily calibrated with the electrostatic technique, while enclosed cavity calibrations typically have an upper limit for the useful frequency range. Thus, work was initiated to perform a high-frequency pressure calibration of Panasonic electret microphones using a substitution free-field method in a small anechoic chamber. First, a standard variable-frequency pistonphone was used to obtain the pressure calibration up to 16 kHz. Above 16 kHz, to avoid spatially irregular sound fields due to dephasing of loudspeaker diaphragms, a series of resonant ceramic piezoelectric crystals was used at five specific ultrasonic frequencies as the free-field calibration sound source. Then, the free-field sensitivity was converted to a pressure sensitivity with an electrostatic calibration of the reference microphone (an air condenser type), for which the free-field correction is known. Combining the low- and high-frequency data sets, a full frequency calibration of pressure sensitivity for an electret microphone was generated from 63 Hz to 80 kHz.

9:50

2aEA8. Broadband self-calibrating micromachined microphones with integrated optical displacement detection. Neal A. Hall, Wook Lee, and F. Levent Degertekin (G. W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Love Bldg., 771 Ferst Dr., Rm. 320, Atlanta, GA 30332-0105)

An optical displacement detection method for micromachined microphones is described and experimental results are presented. The microphone membrane is fabricated on a transparent substrate and the back electrode is patterned in the form of diffraction fingers. This structure forms a phase sensitive diffraction grating, providing the displacement sensitivity of an optical interferometer. The diffraction fingers are also used for electrostatic actuation, providing sensitivity adjustment and self-calibration capabilities. Optically semitransparent coatings are also employed to create Fabry–Perot resonant cavities that enhance the optical detection sensitivity by an order of 10 dB. The high electrical sensitivity provided by optical displacement detection relaxes requirements on me-

chanical sensitivity, and small microphone membranes on the order of 200 μm with vacuum-sealed and air-sealed cavities are used to fabricate microphones with a flat response from dc to over 200 kHz. The optical detection and electrostatic actuation capabilities are demonstrated on fully integrated devices with aluminum microphone membranes micromachined on quartz substrates and bonded to microfabricated silicon photodiodes. [Work supported by DARPA.]

10:05

2aEA9. Passive subtractive beamformer applied to line sound sources. Mitsunori Mizumachi and Satoshi Nakamura (ATR, 2-2-2 Hikaridai, ''Keihanna Sci. City,'' Kyoto 619-0288, Japan)

Speech is an attractive interface for mobile equipment if it is clearly received. A microphone array aids in reducing noise and enhancing speech. A small-scale microphone array generally adopts the subtractive beamforming technique, which constructs sharp notches in a beampattern in either an active or a passive way, so that target sound sources are assumed as point sound sources. However, the shapes of actual sound sources are various and not always points. In particular, the mouth needs to be modeled as line sound source or plane sound source. Therefore, beamformers should target line sound sources. A passive subtractive beamformer is proposed to expand width of a sharp notch by combining several 2ch subtractive beamformers. In multiple sound source conditions, it is necessary to avoid the effect of spatial aliasing against nontarget sound sources. Then, a hybrid technique is applied to realize an optimal connection of the quadruple, double, and single subtractive beamformer for the low, middle, and high frequency region. The feasibility of the proposed subtractive beamformer is confirmed by performance evaluation in suppressing the signals from line sound sources. [Work supported by the Telecommunications Advancement Organization of Japan.]

TUESDAY MORNING, 3 DECEMBER 2002 CORAL GARDEN 1, 8:30 TO 11:50 A.M.

Session 2aED

Education in Acoustics: Development of Acoustics Programs in Latin America

Daniel R. Raichel, Cochair
2727 Moore Lane, Fort Collins, Colorado 80526

Moyses Zindeluk, Cochair
Mechanical Engineering Department, COPPE, University of Rio de Janeiro, Caixa Postal 68503, Rio de Janeiro 21945-000, Brazil

Chair's Introduction—8:30

Invited Papers

8:35

2aED1. Education in acoustics and vibration at UFSC—Brazil. Samir N. Y. Gerges (Dept. of Mech. Eng., Federal Univ. of Santa Catarina, Cx.P. 476, Florianopolis, SC, Brazil)

In the 1970s, Brazil invested heavily on postgraduate program of all areas, especially in acoustics and vibration. Several universities achieved benefits from these investments, namely the Federal University of Santa Catarina (UFSC), the Federal University of Rio de Janeiro (UFRJ), and the Federal University of Santa Maria (UFSM). Part of the undergraduate and postgraduate studies at the Mechanical Engineering Department (EMC) of the Federal University of Santa Catarina relates to vibration and noise. On the undergraduate level an optional course called Noise Control, totaling 54 hours, is offered, which covers basic acoustics and noise control concepts. In the postgraduate program Master's and Doctorate degrees students can attend courses and pursue studies in the area of noise and vibration. This area of concentration is supported by a well equipped laboratory consisting of two reverberation chambers, with a third one, for transmission loss measurements, under construction, together with a hemianechoic room and equipment for the measurement and analysis of noise and vibration. Part of this laboratory, the Industrial Noise Laboratory, is accredited by the Brazilian authorities for measurements of and research on hearing protectors. [Work supported by the Federal Government, and industry.]

8:55

2aED2. Education in acoustics in Argentina. Federico Miyara (Acoust. and Electroacoust. Lab., Natl. Univ. of Rosario, Riobamba 245 bis, 2000 Rosario, Argentina, fmiyara@fceia.unr.edu.ar)

Over the last decades, education in acoustics (EA) in Argentina has experienced ups and downs due to economic and political issues interfering with long term projects. Unlike other countries, like Chile, where EA has reached maturity in spite of the acoustical industry having shown little development, Argentina has several well-established manufacturers of acoustic materials and equipment but no specific career with a major in acoustics. At the university level, acoustics is taught as a complementary—often elective—course for careers such as architecture, communication engineering, or music. In spite of this there are several research centers with programs covering environmental and community noise, effects of noise on man, acoustic signal processing, musical acoustics and acoustic emission, and several national and international meetings are held each year in which results are communicated and

discussed. Several books on a variety of topics such as sound system, architectural acoustics, and noise control have been published as well. Another chapter in EA is technical and vocational education, ranging between secondary and postsecondary levels, with technical training on sound system operation or design. Over the last years there have been several attempts to implement master degrees in acoustics or audio engineering, with little or no success.

9:15

2aED3. Efforts regarding acoustical education for architectural students at the Universidad Peruana de Ciencias Aplicadas (Peruvian University of Applied Sciences), UPC. Jorge Moy (Universidad Peruana de Ciencias Aplicadas, Av. Prol. Primavera 2390, Surco, Lima, Peru)

The lack of knowledge in acoustics among the vast majority of Peruvian architects results in acoustical problems in buildings from a lack of such considerations in the design stage. This paucity of knowledge on the part of the architects may be attributed to the lack of emphasis on the role of acoustics in most architectural curricula in Peruvian universities. A purpose in this paper is to present a brief report on last year's efforts to implement courses in Architectural Acoustics and Noise Control for architecture students at UPC.

9:35

2aED4. Acoustics lecturing in Mexico. Sergio Beristain (ESIME, IPN, IMA, Mexico, sberista@hotmail.com)

Some thirty years ago acoustics lecturing started in Mexico at the National Polytechnic Institute in Mexico City, as part of the Bachelor of Science degree in Communications and Electronics Engineering curricula, including the widest program on this field in the whole country. This program has been producing acoustics specialists ever since. Nowadays many universities and superior education institutions around the country are teaching students at the B.Sc. level and postgraduate level many topics related to acoustics, such as Architectural Acoustics, Seismology, Mechanical Vibrations, Noise Control, Audio, Audiology, Music, etc. Also many institutions have started research programs in related fields, with participation of medical doctors, psychologists, musicians, engineers, etc. Details will be given on particular topics and development.

9:55

2aED5. Engineering acoustics: A pioneer undergraduate program at Rio de Janeiro. Roberto A. Tenenbaum and Moyses Zindeluk (Acoust. and Vib. Lab., Mech. Eng. Dept., Federal Univ. of Rio de Janeiro, CP 68503, 21945-970 Rio de Janeiro, Brazil, moyses@ufrj.br)

Acoustics, essentially a multidisciplinary subject, still has in Brazil a small but increasing number of professionals with a solid background to deal with various aspects of this area. Since 1970 the faculty of the Acoustics and Vibration Laboratory, COPPE/UFRJ, offers graduate (M.Sc and D.Sc) programs, and some undergraduate courses in acoustics, vibration, and signal processing. In January 2000, this group launched a formal undergraduate engineering acoustics program in the Mechanical Engineering Department of the Federal University of Rio de Janeiro. After three years of mechanical engineering, with a firm foundation in physics, applied mathematics, and engineering basics, the undergraduate student may elect to take the engineering acoustics program for the remaining two years. In this program, a wide number of courses are offered, including basic acoustics, room acoustics, signal processing, musical acoustics, machine diagnosis, etc. Approximately 30 different courses may be chosen from. However, the student is not completely free, since the courses selected must fit within a subject concentration profile, e.g., noise control or musical acoustics. In this paper the programs curriculum are presented and its impact on the students is discussed. A first evaluation of the qualifications achieved by the graduate students in the area is also presented.

10:15–10:25 Break

10:25

2aED6. On helping Latin American countries in education in acoustics. Daniel R. Raichel (Douglas Eilar & Assoc., Encinitas, CA 92024-3130 and the Grad. Ctr., CUNY)

The science and applications of acoustics are just as important in Latin America as they are in North America and elsewhere. However, resources in academia are harder to come by in nearly all of the Central American and South American nations; and therefore it would behoove U.S. and European acousticians to help their Latin-American counterparts in achieving their goals of quality education in acoustics, particularly in architectural acoustics, noise control, biomedical usages of ultrasound, signal analyses, and measurement techniques. Among the means of helping are scholarly exchanges, more support by the U.S. government for such exchanges (particularly through Fulbright programs—it is unfortunate that the Fulbright Senior Specialist Program does not recognize acoustics as being one of the environmental sciences), collaboration on research projects, long-term equipment loans and/or outright donations, etc. Advice by experienced practitioners in establishing or improving acoustics laboratories can optimize equipment selection and development of the curriculum.

10:45

2aED7. Acoustics: A branch of engineering at the Universidad Austral de Chile (UACh). Victor Poblete, Jorge P. Arenas, and Jorge Sommerhoff (Inst. of Acoust., Universidad Austral de Chile, P.O. Box 567, Valdivia, Chile, vpoblete@uach.cl)

At the end of the 1960s, the first acousticians graduating at UACh had acquired an education in applied physics and musical arts, since there was no College of Engineering at that time. Initially, they had a (rather modest) four-year undergraduate program, and most of the faculty were not specialized teachers. The graduates from such a program received a sound engineering degree and they were skilled for jobs in the musical industry and sound reinforcement companies. In addition, they worked as sound engineers and producers. Later, because of the scientific, industrial and educational changes in Chile during the 1980s, the higher education system

had massive changes that affected all of the undergraduate and graduate programs of the 61 universities in Chile. The UACh College of Engineering was officially founded in 1989. Then, acoustics as an area of expertise was included, widened and developed as an interdisciplinary subject. Currently, the undergraduate program in acoustics at UACh offers a degree in engineering sciences and a 6-year professional studies in Civil Engineering (Acoustics), having two main fields: Sound and Image, and Environment and Industry.

Contributed Papers

11:05

2aED8. Acoustic Engineering program at the Universidad Austral de Chile (UACh). Jorge Sommerhoff, Victor Poblete, and Jorge P. Arenas (Inst. of Acoust., Universidad Austral de Chile, P.O. Box 567, Valdivia, Chile, jsommerh@uach.cl)

From the beginning of the acoustics program at UACh in 1968, the studies of Acoustic Engineering have been modified and developed according to the vision and human resources of its developers. Three different stages of growth can be seen. When the program began, it was totally aimed at forming skilled professionals in audio and recording. In this way, the professional title given was Sound Engineer. At that time, each applicant was required to have ''good musical hearing,'' which had to be demonstrated through a special musical audition test. The second stage was characterized by the incorporation of acoustics subjects which allowed students, with no musical abilities, to competently work on acoustic engineering activities not related to music. Then, the professional title was changed to Acoustic Engineer. Thus, job opportunities were diversified and access was allowed by all types of students. In the last stage, the study plan was modified as a response to the new vision and requirements of the globalized world in which the environmental component has a great importance. In this work the development of a program that dates from 35 years ago is presented and justified.

11:20

2aED9. Using dynamic means to teach a dynamic subject: A new concept for a graduate vibrations course in Mexico. Salvador Echeverria-Villagomez (Apdo. Postal 1-100, C.P. 76000, Queretaro, Mexico)

Mechanical vibrations is a subject that belongs in every undergraduate curricula in mechanical engineering, and to many graduate programs in mechanical design and manufacturing. It is a subject in which many disciplines come in, from mechanics, through materials properties, to mechanisms and special mathematics tools. It is also a most interesting discipline from many points of view: from the theoretical models used to analyze the dynamical behavior of systems, through the modern numerical and computational tools, to the experimental techniques used to measure and test their performance. The paper presents a concept to teach vibrations and acoustics. This concept provides a framework for using the mentioned tools, which is homologous in structure to the subject being taught, thus enhancing the power and effectiveness of teaching and training in the field. The main characteristics of the concept are their systemic approach, fluid logic, use of conceptual virtual maps and dynamic means.

11:35

2aED10. Acoustic outsourcing: New employment possibilities for the specialists. Patricia Perez, Heriberto Rios, Armando Andrade, and Mario Ramirez (Laboratorio de Desarrollo Tecnologico en Bioingenieria, ESIME-IPN, Mexico)

The need for companies to be more competitive has led them to resort to training, external consultantship, continuous improvement programs, but with the aim of achieving maximum productivity, the big companies go even further: they are opting to focus on their high-priority activities, leaving some nonstrategic functions in the hands of third parties (organizations or individuals). Acoustic outsourcing presents immense business opportunities for the specialists in this area when offering services or completing a production process that the company carries out in an internal way but that is not its main function or activity. Outsourcing contemplates a serious long term commitment between the two parties; a kind of strategic alliance, all with the purpose of increasing efficiency and the quality of the products that the company develops, besides solving acoustic problems related to the production stage. (To be presented in Spanish.)

TUESDAY MORNING, 3 DECEMBER 2002 CORAL SEA 1 AND 2, 8:00 TO 11:45 A.M.

Session 2aMU

Musical Acoustics: Analysis, Synthesis, Perception and Classification of Musical Sounds

James W. Beauchamp, Chair
School of Music, Department of Electrical and Computer Engineering, University of Illinois, Urbana, Illinois 61801

Invited Papers

8:00

2aMU1. Spectral modeling, analysis, and synthesis of musical sounds. Sylvain Marchand and Myriam Desainte-Catherine (LaBRI, Univ. of Bordeaux 1, 351 cours de la Liberation, F-33405 Talence cedex, France)

Spectral models provide general representations for sound well-suited for expressive musical transformations. These models allow us to extract and modify perceptually-relevant parameters such as amplitude, frequency, and spectrum. Thus, they are of great interest for the classification of musical sounds. A new analysis method was proposed to accurately extract the spectral parameters for the model from existing sounds. This method extends the classic short-time Fourier analysis by also considering the derivatives of the sound signal, and it can work with very short analysis windows. Although originally designed for stationary sounds with no noise, this method shows excellent results in the presence of noise and it is currently being extended in order to handle nonstationary sounds as

well. A very efficient synthesis algorithm, based on a recursive description of the sine function, is able to reproduce sound in real time from the model parameters. This algorithm allows an extremely fine control of the partials of the sounds while avoiding signal discontinuities as well as numerical imprecision, and with a nearly optimal number of operations per partial. Psychoacoustic phenomena such as masking are considered in order to reduce on the fly the number of partials to be synthesized.

8:20

2aMU2. Easily extensible unix software for spectral analysis, display, modification, and synthesis of musical sounds. James W. Beauchamp (School of Music and Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801, j-beauch@uiuc.edu)

Software has been developed which enables users to perform time-varying spectral analysis of individual musical tones or successions of them and to perform further processing of the data. The package, called SNDAN, is freely available in source code, uses EPS graphics for display, and is written in ANSI C for ease of code modification and extension. Two analyzers, a fixed-filter-bank phase vocoder (''pvan'') and a frequency-tracking analyzer (''mqan'') constitute the analysis front end of the package. While pvan's output consists of continuous amplitudes and frequencies of harmonics, mqan produces disjoint ''tracks.'' However, another program extracts a fundamental frequency and separates harmonics from the tracks, resulting in a continuous harmonic output. ''monan'' is a program used to display harmonic data in a variety of formats, perform various spectral modifications, and perform additive resynthesis of the harmonic partials, including possible pitch-shifting and time-scaling. Sounds can also be synthesized according to a musical score using a companion synthesis language, Music 4C. Several other programs in the SNDAN suite can be used for specialized tasks, such as signal display and editing. Applications of the software include producing specialized sounds for music compositions or psychoacoustic experiments or as a basis for developing new synthesis algorithms.

8:40

2aMU3. Analysis-synthesis of musical sounds by hybrid models. S. Ystad (CNRS, Laboratoire de Mecanique et d'Acoustique 31, Chemin Joseph Aiguier, 13402 Marseille cedex 20, France)

Analysis-synthesis consists of constructing synthetic sounds from natural sounds by algorithmic synthesis methods. The models used for this purpose are of two kinds: physical models which take into account the physical characteristics of the instrument and signal models which take into account perceptual criteria. By combining physical and signal models hybrid models can be constructed taking advantage of the positive aspects of both methods. In this presentation I show how hybrid models can be adapted to specific instruments producing both sustained and plucked sounds. In these cases signal models are used to model the nonlinear source signal. The parameters of these models are obtained from perceptual criteria such as the spectral centroid or the tristimulus. The source signal is further injected into the physical model which consists of a digital wave guide model. The parameters of the physical model are extracted from the natural sound by analysis based on linear time-frequency representations such as the Gabor and the wavelet transforms. The models which will be presented are real-time compatible and in the flute case an interface adapted to a traditional flute which pilots a hybrid model will be described.

9:00

2aMU4. Recent developments in automatic classification of musical instruments. Bozena Kostek (Sound & Vision Eng. Dept., Gdansk Univ. of Technol., Narutowicza 11/12, 80-952 Gdansk, Poland)

In this paper recent developments in automatic classification of musical instrument domain are presented. Issues related to automatic classification of music are data representation of musical instrument sounds, automatic musical sound recognition, musical duet separation, music recognition, etc. These problems belong to the so-called Musical Information Retrieval domain. The best developed is the automatic recognition of individual musical sounds. In rich literature on this subject many references can be found. Another issue deals with music information retrieval understood as searching for music-related features such as song titles, etc. A query-by-humming can be also cited as one of the MIR topics. The most difficult problem that deals with automatic recognition of multipitch excerpts still remains unsolved, however, recently some approaches to this issue can be found in the literature. Some of the mentioned problems were subjects of the research carried out at the Sound & Vision Department of the Gdansk University of Technology. The developed solutions in the domain of automatic classification of individual sounds, duet separation, and music recognition will be presented as examples of possible case-studies in the MIR domain. The proposed approach was evaluated on musical datebases created at the Department. [Work supported by KBN, Grant No. 4 T11D 014 22.]

9:20

2aMU5. The timbre model. Kristoffer Jensen (Dept. of Datalogy, Univ. of Copenhagen, 2100 Copenhagen, Denmark, http://www.diku.dk)

A timbre model is proposed for use in multiple applications. This model, which encompasses all voiced isolated musical instruments, has an intuitive parameter set, fixed size, and separates the sounds in dimensions akin to the timbre dimensions as proposed in timbre research. The analysis of the model parameters is fully documented, and it proposes, in particular, a method for the estimation of the difficult decay/release split-point. The main parameters of the model are the spectral envelope, the attack/release durations and relative amplitudes, and the inharmonicity and the shimmer and jitter (which provide both for the slow random variations of the frequencies and amplitudes, and also for additive noises). Some of the applications include synthesis, where a real-time application is being developed with an intuitive gui, classification, and search of sounds based on the content of the sounds, and a further understanding of acoustic musical instrument behavior. In order to present the background of the model, this presentation will start with sinusoidal A/S, some timbre perception research, then present the timbre model, show the validity for individual music instrument sounds, and finally introduce some expression additions to the model.

9:40

2aMU6. Hypersignal analyses of orchestral instrument signals as correlated with perception of timbre. Roger A. Kendall (Music Cognition and Acoust. Lab., Schoenberg Hall, UCLA, Los Angeles, CA 90024)

Experiments were conducted to assess the relationships among signal analyses and timbral perception across the playing range of bassoon, trombone, tenor saxophone, alto saxophone, soprano saxophone, French horn, violin, oboe, flute, clarinet, and trumpet. Spectral analyses employed Hypersignal using 9th order Zoom FFT on 22.05 samples/s signals. Spectral centroid and spectral flux measures were calculated. Perceptual experiments included similarity scaling and identification at various pitch chroma across the playing range of the instruments. In addition, a pilot experiment assessing the interaction of pitch chroma and timbre was conducted where timbral judgements were made across, rather than within, pitch chroma. Results suggest that instruments with relatively low tessitura produce higher centroid ranges since the larger air column yeilds a large number of vibrational modes. In contrast, higher tessitura instruments, using smaller air columns, produce fewer modes of vibration with increasing pitch chroma, to the point that the centroids coverge near Bb5. Perceptual data correspond to the spectral, resulting in less specificity among instruments at their higher tessituras. It is suggested that spectral centroid, which maps strongly near A4 in the majority of studies, must be viewed with caution as a predictor of timbre at tessitura extremes.

10:00

2aMU7. A confirmatory analysis of four acoustic correlates of timbre space. Stephen McAdams, Anne Caclin, and Bennett K. Smith (Ircam–CNRS, 1 pl. Igor Stravinsky, F-75004 Paris, France)

Exploratory multidimensional scaling studies of musical instrument timbres generally yield two- to four-dimensional perceptual spaces. Acoustic parameters have been derived that correlate moderately to highly with the perceptual dimensions. In a confirmatory study, two three-dimensional sets of synthetic, harmonic sounds equalized for fundamental frequency, loudness, and perceived duration were designed. The first two dimensions corresponded to attack time and spectral centroid in both sound sets. The third dimension corresponded to spectral flux (variation of the spectral centroid over time) in the first set and to the energy ratio of odd to even harmonics in the second set. Group analyses of dissimilarity judgments for all pairs of sounds homogeneously distributed in each space revealed a two-dimensional solution for the first set and a three-dimensional solution for the second set. Log attack time and spectral centroid were confirmed as perceptual dimensions in both solutions. The even/odd energy ratio was confirmed as a third dimension in the second set. Spectral flux was not confirmed in the first set, suggesting that this parameter should be re-examined. Analyses of individual data sets tested for differences across listeners in the mapping of acoustic parameters to perceptual dimensions. [Work supported by the CTI program of the CNRS.]

10:20–10:30 Break

Contributed Papers

10:30

2aMU8. Piano string modeling: From partial differential equations to digital wave-guide model. J. Bensa (CNRS, Laboratoire de Mecanique et d'Acoustique 31, Chemin Joseph Aiguier, 13402 Marseille cedex 20, France), S. Bilbao (Stanford Univ., Stanford, CA), R. Kronland-Martinet (CNRS, 13402 Marseille cedex 20, France), and Julius O. Smith III (Stanford Univ., Stanford, CA)

A new class of partial differential equations (PDE) is proposed for transverse vibration in stiff, lossy strings, such as piano strings. While only second-order in time, it models both frequency-dependent losses and dispersion effects. By restricting the time-order to 2, valuable advantages are achieved: First, the frequency-domain analysis is simplified, making it easy to obtain explicit formulas for dispersion and loss versus frequency; for the same reason, exact bounds on sampling in associated finite-difference-schemes (FDS) can be derived. Second, it can be shown that the associated FDS is ''well posed'' in the sense that it is stable, in the limit, as the sampling period goes to zero. Finally, the new PDE class can be used as a starting point for digital wave-guide modeling [a digital wave-guide factors one-dimensional wave propagation as purely lossless throughout the length of the string, with losses and dispersion lumped in a low-order digital filter at the string endpoint(s)]. We perform numerical simulations comparing the finite-difference and digital wave-guide approaches, illustrating the advantages of the latter. We examine a procedure allowing the resynthesis of natural string vibration; using experimental data obtained from a grand piano, the parameters of the physical model are estimated over most of the keyboard range.

10:45

2aMU9. The wave digital piano hammer. Stefan D. Bilbao, Julius O. Smith III (Ctr. for Computer Res. in Music and Acoust., Dept. of Music, Stanford Univ., Stanford, CA 94305), Julien Bensa, and Richard Kronland-Martinet (S2M-LMA-CNRS, Marseille, France)

For sound synthesis purposes, the vibration of a piano string may be simply modeled using bidirectional delay lines or digital waveguides which transport traveling wavelike signals in both directions. Such a digital wave-type formulation, in addition to yielding a particularly computationally efficient simulation routine, also possesses other important advantages. In particular, it is possible to couple the delay lines to a nonlinear exciting mechanism (the hammer) without compromising stability; in fact, if the hammer and string are lossless, their digital counterparts will be exactly lossless as well. The key to this good property (which can be carried over to other nonlinear elements in musical systems) is that all operations are framed in terms of the passive scattering of discrete signals in the network, the sum of the squares of which serves as a discrete-time Lyapunov function for the system as a whole. Simulations are presented.

11:00

2aMU10. Musical sound analysis/synthesis using vector-quantized time-varying spectra. Andreas F. Ehmann and James W. Beauchamp (Univ. of Illinois at Urbana–Champaign, 5308 Music Bldg., 1114 W. Nevada St., Urbana, IL 61801)

A fundamental goal of computer music sound synthesis is accurate, yet efficient resynthesis of musical sounds, with the possibility of extending the synthesis into new territories using control of perceptually intuitive

parameters. A data clustering technique known as vector quantization (VQ) is used to extract a globally optimum set of representative spectra from phase vocoder analyses of instrument tones. This set of spectra, called a Codebook, is used for sinusoidal additive synthesis or, more efficiently, for wavetable synthesis. Instantaneous spectra are synthesized by first determining the Codebook indices corresponding to the best least-squares matches to the original time-varying spectrum. Spectral index versus time functions are then smoothed, and interpolation is employed to provide smooth transitions between Codebook spectra. Furthermore, spectral frames are pre-flattened and their slope, or tilt, extracted before clustering is applied. This allows spectral tilt, closely related to the perceptual parameter ''brightness,'' to be independently controlled during synthesis. The result is a highly compressed format consisting of the Codebook spectra and time-varying tilt, amplitude, and Codebook index parameters. This technique has been applied to a variety of harmonic musical instrument sounds with the resulting resynthesized tones providing good matches to the originals.

11:15

2aMU11. The syrinx: Nature's hybrid wind instrument. Tamara Smyth and Julius O. Smith III (Ctr. for Computer Res. in Music and Acoust., Stanford Univ., Stanford, CA 94305)

Birdsong is commonly associated with the sound of a flute. Although the pure, often high pitched, tone of a bird is undeniably flutelike, its sound production mechanism more closely resembles that of the human voice, with the syringeal membrane (the bird's primary vocal organ) acting like vocal folds and a beak acting as a conical bore. Airflow in the song bird's vocal tract begins from the lungs and passes through two bronchi, two nonlinear vibrating membranes (one in each bronchial tube), the trachea, the mouth, and finally propagates to the surrounding air by way of the beak. Classic waveguide synthesis is used for modeling the bronchi and trachea tubes, based on the model of Fletcher [J. Acoust. Soc. Am. (1988, 1999)]. The nonlinearity of the vibrating syringeal membrane is simulated by finite-difference methods. This nonlinear valve, driven by a steady pressure from the bronchi, generates an oscillatory pressure entering the trachea.

11:30

2aMU12. Electrophysiological correlates of musical timbre perception. Anne Caclin (Ircam–CNRS, 1 pl. Igor Stravinsky, F-75004 Paris, France), Elvira Brattico (Univ. of Helsinki, Helsinki, Finland), Bennett K. Smith (Ircam–CNRS, F-75004 Paris, France), Mari Tervaniemi (Univ. of Helsinki, Helsinki, Finland), Marie-Hilhne Giard (Inserm U280, F-69424 Lyon 03, France), and Stephen McAdams (Ircam–CNRS, F-75004 Paris, France)

Timbre perception has been studied by deriving a multidimensional space of the perceptual attributes from listeners' behavioral responses. The neural bases of timbre space were sought. First, a psychophysical timbre dissimilarity experiment was conducted. A three-dimensional space of 16 synthetic sounds equalized for fundamental frequency, loudness, and perceived duration was designed. Sounds varied in attack time, spectral center of gravity, and energy ratio of odd/even harmonics. Multidimensional scaling revealed a three-dimensional perceptual space with linear or exponential relations between perceptual and physical dimensions. Second, in an electrophysiological experiment, the mismatch negativity (MMN) component of event-related potentials was recorded. The MMN is elicited by infrequently presented sounds differing in one or more dimensions from more frequent ones. Although elicited without the focus of attention, it correlates with the subjects' behavioral responses, revealing the neural bases of preattentive discrimination. Eight sounds were chosen within the perceptual space. Changes along individual and combined dimensions elicited an MMN response. MMN latency varied depending on the dimension changed. In addition, preliminary analyses tend to show an additivity of the MMN waves for some pairs of dimensions. These results shed light on the neural processes underlying the perceptual representation of multidimensional sounds. [Work supported by the CTI program of the CNRS.]

TUESDAY MORNING, 3 DECEMBER 2002 CORAL ISLAND 1 AND 2, 8:00 A.M. TO 12:00 NOON

Session 2aPA

Physical Acoustics: Bubbles, Drops and Foams I

R. Glynn Holt, Chair

Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Chair's Introduction—8:00

Invited Papers

8:05

2aPA1. Radiation pressure of standing waves on liquid columns and small diffusion flames. David B. Thiessen, Mark J. Marr-Lyon, Wei Wei, and Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814)

The radiation pressure of standing ultrasonic waves in air is demonstrated in this investigation to influence the dynamics of liquid columns and small flames. With the appropriate choice of the acoustic amplitude and wavelength, the natural tendency of long columns to break because of surface tension was suppressed in reduced gravity [M. J. Marr-Lyon, D. B. Thiessen, and P. L. Marston, Phys. Rev. Lett. **86**, 2293–2296 (2001); **87**(20), 9001(E) (2001)]. Evaluation of the radiation force shows that narrow liquid columns are attracted to velocity antinodes. The response of a small vertical diffusion flame to ultrasonic radiation pressure in a horizontal standing wave was observed in normal gravity. In agreement with our predictions of the distribution of ultrasonic radiation stress on the flame, the flame is attracted to a pressure antinode and becomes slightly elliptical with the major axis in the plane of the antinode. The radiation pressure distribution and the direction of the radiation force follow from the dominance of the dipole scattering for small flames. Understanding radiation stress on flames is relevant to the control of hot fluid objects. [Work supported by NASA.]

8:30

2aPA2. Foam rheology in the wet and dry limits. J. Gregory McDaniel, R. Glynn Holt (Aerosp. and Mech. Engr. Dept., Boston Univ., 110 Cummington St., Boston, MA 02115), and Iskander Sh. Akhatov (Bashkir State Univ., 450000 Ufa, Russia)

Understanding the rheological behavior of wet foams is important as a basic problem in fluid physics, and as a practical problem in many industries. This lecture will describe research into the wet and dry limits of foam rheology by a relatively new experimental technique in which foam drops are acoustically levitated and driven into motion. In the dry limit, the drops behave as viscoelastic solids. The effective moduli of the foam are estimated by observing the resonances of the drops and matching them to an analytical model for viscoelastic sphere vibrations. Analytical explorations of the wet limit have proceeded by considering the dynamics of a single bubble in a volume of liquid determined by the foam's void fraction. The linearized result is a wave equation, from which the natural frequencies and mode shapes of wet foam drops are determined. Relationships between this wave equation and those of classical effective medium theories will be described. [Work supported by NASA.]

8:55

2aPA3. Airborne chemistry single cell level. Staffan Nilsson, Peter Viberg, Peter Spegel, Sabina Santesson (Tech. Analytical Chemistry, Lund Univ., P.O. 124, SE-221 00 Lund, Sweden), Eila Cedergren, Eva Degerman, Tomas Johansson, and Johan Nilsson (Lund Univ., Lund, Sweden)

A miniaturized analysis system for the studying of living cells and biochemical reactions in microdrops was developed. Cell studies were performed using single adipocytes in 250-nL drops. Continuous flow-through droplet dispensers, developed in-house, were used for additions to the levitated droplet. Addition of b-adrenergic agonists stimulates the lipolysis in the adipocytes, leading to free fatty acid release and a consequent *p*H decrease of the surrounding buffer, a change that can be easily followed using a *p*H-dependent fluorophore continuously monitored by fluorescence imaging detection. An analytical method using capillary electrophoresis and nanospray mass spectrometry for measurement of the cAMP level in activated single adipocytes are now being developed for future use in combination with the levitation technique. The levitation approach was also employed for the screening of nucleation conditions for macromolecules. Here, the acoustic levitator offers a simplified way to determine the main features of the phase diagram (i.e., precipitation diagram). Using the droplet dispensers, different types and amounts of precipitation agents are injected into the levitated drop, allowing a systematic search for nucleation conditions that is not possible using standard crystallization methods. Once the precipitation diagram has been obtained, optimization using standard methods is employed to grow the crystals.

9:20

2aPA4. The role of bubbles and cavitation in the production of thermal lesions from high-intensity focused ultrasound. Ronald A. Roy, R. Glynn Holt, Xinmai Yang, and Patrick Edson (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, ronroy@bu.edu)

Rapid hyperthermia resulting in tissue necrosis is a key physical mechanism for focused ultrasound surgery (FUS). At therapeutic intensities, tissue heating is often accompanied by cavitation activity. Although it is well known that bubbles promote mechanical damage, *in vitro* and *in vivo* experiments have shown that under certain conditions bubble activity can double the heating rate. With a view towards harnessing bubbles and cavitation for useful clinical work, we report the results of *in vitro* experiments and modeling for the dynamic and thermal behavior of bubbles subjected to 1-megahertz ultrasound at mega-pascal pressures. The dominant bubble-related heating mechanism depends critically on the bubble size distribution which, in turn, depends on insonation control parameters (acoustic pressure, pulse duration), medium properties (notably dissolved gas concentration), and bubble-destroying shape instabilities. The evidence points to a range of control parameters for which bubble-enhanced FUS can be assured. [Work supported by DARPA and the U.S. Army.]

9:45–10:00 Break

Contributed Papers

10:00

2aPA5. The effect of dissolved gas concentration on bubble-enhanced heating in tissue-mimetic materials. Xinmai Yang, Ronald A. Roy, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

Bubble-enhanced heating is a key mechanism to cause tissue damage in ultrasound surgery. We have conducted experiments in an agar-based tissue phantom. We found that the difference of air concentration in the tissue phantom has a small but measurable effect on the enhanced heating. Notably, high air concentration samples exhibit very good repeatability. We have passively monitored broadband acoustic emissions from the bubbles in order to determine if diagnostic information could be gleaned from such signals. Finally we investigate the effect of bubble size distribution on bubble-enhanced heating by employing bubble-based contrast agents to control the initial bubble size distribution. [Work supported by DARPA and the U.S. Army.]

10:15

2aPA6. A theoretical model for bubble enhanced ultrasound heating due to time-dependent bubble size distributions. Yang Xinmai, R. Glynn Holt, Patrick Edson, and Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, xmyang@bu.edu)

Substantial *in vitro* and *in vivo* evidence shows that cavitation activity can affect tissue heating in focused ultrasound surgery and acoustic hemostasis applications. In particular, the heating rate in tissue increases significantly after cavitation sets in. Exploitation of this phenomenon for clinical use requires knowledge of, among other parameters, the time-dependent bubble size distribution sustained during insonation. Difficulties associated with the measurement of bubble sizes during *in vitro* or *in vivo* experiments call for a theoretical approach to the problem. We will present a theoretical model that estimates the time-dependent distribution of bubble equilibrium radii. Shape instability thresholds and rectified diffusion thresholds bound asymptotic bubble size distributions, and the instan-

taneous size distributions are governed by growth rates. The temperature rise caused by such bubble activity is calculated and compared with experimental data. [Work supported by DARPA and the U.S. Army.]

10:30

2aPA7. Low-power, cylindrical, air-coupled acoustic levitation/concentration devices: Symmetry breaking of the levitation volume. Gregory Kaduchak, Aleksandr S. Kogan, Christopher S. Kwiatkowski, and Dipen N. Sinha (Los Alamos Natl. Lab., MS D429, Los Alamos, NM 87545)

A cylindrical acoustic device for levitation and/or concentration of aerosols and small liquid/solid samples (up to several millimeters in diameter) in air has been developed [Kaduchak *et al.*, Rev. Sci. Instrum. **73**, 1332–1336]. It is inexpensive, low-power, and, in its simplest embodiment, does not require accurate alignment of a resonant cavity. It is constructed from a cylindrical PZT tube with thickness-to-radius ratio $h/a \sim 0.03$. The novelty of the device is that the lowest-order breathing mode of the tube is tuned to match a resonant mode of the interior air-filled cylindrical cavity. A high-Q cavity results that is driven very efficiently; drops of water in excess of 1-mm diameter are levitated for approximately 100 mW of input electrical power. The present research addresses modifying the different spatial configurations of the standing wave field within the cavity. By breaking the cylindrical symmetry, it is shown that pressure nodes can be localized for collection or separation of aerosolds or other particulate matter. Several different symmetry-breaking configurations are demonstrated. It is shown that experimental observations of the nodal arrangements agree with theoretical predictions.

10:45

2aPA8. Static and oscillatory response measurements of acoustically levitated foam drops. Li Liu, J. Gregory McDaniel, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

Small samples of aqueous foam of varying gas volume fraction are acoustically levitated in an ultrasonic field. The drops are subjected to both static and time-varying pressures. Normal mode frequencies and inferred rheological properties (yield stress, shear modulus) for foams as a function of gas volume fraction will be presented. We compare the experimental results to recent theoretical descriptions of such modal oscillations [McDaniel and Holt, Phys. Rev. E **61**, 2204 (2000); McDaniel, Akhatov, and Holt, Phys Fluids **14**, 1886 (2002)]. [Work supported by NASA.]

11:00

2aPA9. Transport parameters for pulsed ultrasonic waves propagating in an aluminum foam. Arnaud Tourin, Arnaud Derode, Victor Mamou, Mathias Fink (LOA, ESPCI 10, rue Vauquelin 75005 Paris, France), John Page, and Michael L. Cowan (Univ. of Manitoba, Winnipeg, MB R3T 2N2, Canada)

Aluminum foams have now been studied for many years in large part because of their applications as light-weight elastic materials (e.g., car bumpers, aerospace engineering applications). The pore size and the spatial distribution of the pores govern the mechanical behavior of the foam and can vary enormously depending on the method of manufacturing. Thus, new methods for the nondestructive characterization of these materials are needed. We present here a set of experimental ultrasonic methods in a range of frequencies where the ultrasonic waves are multiply scattered in the medium. In this regime, the propagation is described by ultrasonic transport parameters which are related to the microstructure of the foam. The diffusion coefficient and the absorption mean free path have been determined in pulse transmission experiments by fitting the solution of the diffusion equation to the average intensity, the so-called time of flight distribution. To more fully characterize the medium, the transport mean path and the diffusion coefficient have been measured in backscattering experiments using the static and dynamic coherent backscattering effects. For both methods, the properties of the sample interfaces have been taken into account.

11:15

2aPA10. Low frequency cavitation erosion. Sally J. Pardue and Gautam Chandekar (Dept. of Mech. Eng., Tennessee Technolog. Univ., Box 5014, Cookeville, TN 38505, spardue@tntech.edu)

Damage of diesel engine piston sleeve liners due to cavitation of the coolant fluid can be severe. Coolant fluid additives are used to inhibit cavitation damage, and are evaluated by industry suppliers using ASTM G32-98 Standard Test Method for Cavitation Erosion Using Vibratory Apparatus. The ASTM G32-98 test procedure uses an ultrasonic horn at 20 kHz to vibrate a test button in the coolant. The test button mass loss and surface appearance are studied to sort the performance of new coolant additives. Mismatch between good lab performers and actual engine test runs has raised concerns over the current lab test. The frequency range of the current test has been targeted for investigation. A low frequency, less than 2000 Hz, test rig was built to explore the cavitation damage. The test system did produce cavitation on the surface of the test button for a period of 36 h, with minimal mass loss. The test rig experienced cyclic fatigue when test times were extended. The work is now focusing on designing a better test rig for long duration tests and on developing numerical models in order to explore the effects of cavitation excitation frequency on surface erosion.

11:30

2aPA11. A novel cavitation probe design and some preliminary measurements of its application to megasonic cleaning. Lawrence A. Crum (Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105) and Gary Ferrell (SEZ America, Inc., Mountain View, CA 94043)

An initial prototype design for a cavitation probe that uses the property of a collapsing cavitation bubble to produce visible photons (sonoluminescence) has been designed and constructed. These light emissions can be easily detected within a small, finite volume and thus this probe provides a direct means of measuring the cavitation density (activity/per unit volume) within a cavitating fluid and the delivery of ultrasonic energy at an engineered surface. As a result, ultrasonic methods treating a surface can be directly monitored and controlled in real-time, leading to the ability to improve and predict the performance of the resulting structure. This probe provides the potential for constructing a real-time monitor of ultrasonic/megasonic cleaner efficiency and effectiveness. In addition, because the entire three-dimensional cavitation field can be measured with this probe, it can also serve as a useful tool in ultrasonic/megasonic cleaner design. A real-time cavitation-density measuring device would have great utility in the semiconductor cleaning industry and thus this probe provides considerable promise for commercial development. A description of the probe will be presented as well as some preliminary data on cavitation density within a commercial megasonic cleaner. [Work supported in part by the NSF.]

11:45

2aPA12. Bubble dynamics in an acoustic flow field. Dmitry V. Voronin, Georgij N. Sankin (Lavrentyev Inst. of Hydrodynamics, Prosp. Acad. Lavrentyeva, 15, Novosibirsk 630090, Russia, voron@hydro.nsc.ru), Robert Mettin (Universitaet Goettingen, 37073 Goettingen, Germany), Vyacheslav S. Teslenko (Lavrentyev Inst. of Hydrodynamics, Novosibirsk 630090, Russia), and Werner Lauterborn (Universitaet Goettingen, 37073 Goettingen, Germany)

Dynamics of interaction between cavitational bubbles is investigated when a complex of a compression and a rarefaction pulse passes through a liquid with pre-existing micro bubbles. Cavitation was generated experimentally with the help of electromagnetic generator of a flat and a convergent acoustic pulse (2-μs duration, 1–20 MPa) having the form of a hollow sphere segment. A modeling was performed within the frame of two-dimensional axisymmetric nonstationary approach on the basis of conservation laws for a model of an ideal compressible liquid. A thermodynamic flow field was computed both in liquid and inside bubbles. Behind the rarefaction wave the microbubbles begin to grow and generate secondary compression shocks, the amplitude of which may exceed that of

the incident pulse under certain conditions. It is shown that the process of bubble interaction within a cluster is accompanied by bubble coalescence, fragmentation, and collapse of the initial bubble or its fragments. Simultaneously, high temperature spots appear in the bubble compressing by the secondary wave. Adiabatic heating of gas either inside a bubble or near the neck between a bubble and its fragment may result in sonoluminescence, also observed in experiments. [Work supported by ASA, DAAD, and RFBR.]

TUESDAY MORNING, 3 DECEMBER 2002 CORAL GALLERY FOYER, 8:00 TO 11:45 A.M.

Session 2aPP

Psychological and Physiological Acoustics: General Topics in Psychological Acoustics

Jont B. Allen, Cochair
Mimosa Acoustics, 382 Forest Hill Way, Mountainside, New Jersey 07092

Rodrigo Ordoñez, Cochair
Department of Acoustics, Aalborg University, Fredrik Bajers Vej 7 B4, DK-9220 Aalborg, Denmark

Contributed Papers

8:00

2aPP1. Attentional focus and the method of adjustment revisited. Charles S. Watson, Gary R. Kidd, and Soriya V. Pok (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, watson@indiana.edu)

In past reports we have described a technique by which listeners may be trained to focus their auditory attention on a particular spectral-temporal region of a complex acoustic stimulus, using the psychophysical method of adjustment. Previous work will be reviewed, and the results of a new experiment will be described, in which listeners were trained under both the adjustment method and a standard adaptive tracking technique to detect changes in the frequency of a tone within a nine-tone sequence. Under some circumstances, the adjustment procedure can enable listeners to learn to detect very small changes within a few minutes, whereas several hours of training under adaptive methods may be required to achieve the same detection or discrimination performance. Other differences between the two methods will be described. [Work supported by the NIH/NIDCD.]

8:15

2aPP2. Psychophysical analysis of sound and vibration in the cabin of passenger aircrafts. Volker Mellert, Ingo Baumann, Nils Freese, Roland Kruse, Reinhard Weber (Dept. of Phys., Oldenburg Univ., 26111 Oldenburg, Germany), Hermann Remmers, and Michael Bellmann (ITAP GmbH, 26129 Oldenburg, Germany)

The vibroacoustics within the fuselage of several types of aircrafts is recorded with microphones, ear-related devices, and accelerometers at different locations of passengers seats, and the workplace of the cabin and cockpit crew. The signals are analyzed according to standard psychoacoustic and vibration parameters. The requirements for the reproduction of the signals in a ground-based test-bed (e.g., mock-up) are identified. Results are reported on how well test facilities at ground meet real-flight conditions. [Work supported by the European Community (www.heace.org).]

8:30

2aPP3. Distinguishing sound from noise—The significance of attention and noise sensitivity. George Dodd (Acoust. Res. Ctr., Univ. of Auckland, Private Bag 92019, Auckland, New Zealand, g.dodd@auckland.ac.nz)

There is often a large discrepancy between the accuracies of physical measurements and the precision ascribed to subjective responses to sound. Consequently the criteria by which sound and noise are assessed can appear somewhat loose. In previous work it was proposed that there is a need to formalize definitions for noise and non-noise sounds in order to render subjective reactions more readily quantifiable. This is necessary to give greater recognition to the significance of differences in individuals responses, and also to put criteria for environmental sound on a more scientific basis. This presentation reviews our research looking for physiological responses which correlate with a person's attention to sound, and presents results from our study of noise sensitivity. In this work noise sensitivity is defined as being a tendency to be distracted by sound and is viewed as a stable characteristic of people differing between individuals and distinct from noise annoyance experienced at a particular time. The results of assessing noise sensitivity by self-assessment questionaires and other measures are presented and how they relate to individual listening habits.

8:45

2aPP4. Spectral pattern and harmonic relations as factors governing the perceptual cohesion of low-numbered components in complex tones. Brian Roberts (School of Psych., Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, UK, B.Roberts@Bham.ac.uk) and Jeffrey M. Brunstrom (Loughborough Univ., Loughborough, Leicestershire LE11 3TU, UK)

Mistuning a harmonic changes its pitch more than expected and also increases its salience. Both effects can be used to explore the auditory organization of complex tones. In this study we extend our findings indicating that the effects of a spectral pattern on grouping are not restricted to harmonic relations. Stimuli were either harmonic ($F0 = 200$ Hz) or frequency shifted by 25% of $F0$. Component 1 or 2 was replaced by one of a set of sinusoidal probes in the same spectral region. In experiment 1, listeners adjusted the frequency of a pure tone to match the probe pitch. Inflections of the pitch-shift functions were close to the two expected values in the harmonic condition. In the shifted condition, they were close to the suboctave (225 Hz) and the frequency (450 Hz) of component 2. In experiment 2, listeners matched the probe loudness by adjusting the level of a tone of identical frequency. Loudness minima corresponded closely to the pitch-shift inflections. Experiment 3 showed that a pitch-shift inflection close to 450 Hz requires the presence of component 1. These results suggest that the first component of a shifted complex is grouped differently from the others, based on harmonicity rather than spectral spacing.

9:00

2aPP5. Sound localization: Effects of gender, aging, head related transfer function, and auditory performance. Pedro Menezes, Ilka Soares, Silvio Caldas Neto, and Mauricy Motta (UFPE, UNCISAL, Av. Prof. Moraes Rego, 25-Cidade Universitria, Recife, PE, Brasil. Cep. 50960.870, son@hotlink.com.br)

The sound localization resolution in 80 normal hearing subjects of both sexes will be compared to audiometric parameters, audiometry, timpanometry, and stapedian reflexes. Moreover, the pinna and concha length and width so as the interaural distance, ear–head angle, sex, and age were compared too, in a reverberating room. Three tones of square waves with 1, 2, and 3 kHz, were present with an intensity of 70-dB SPL in order of the speakers and sequence of frequencies randomized. The subjects were trained to indicate the origin of the sound in a control console with push-buttons representing the space disposition of the speakers, being 8 in the horizontal plane, 5 in the medial sagital plane, and 5 in the medial frontal plane. The identification of each speaker is done by pressing the respective push-button. The angle between the speaker's axis in the same plane is of 45, and the distance of 1 meter from the analyzed subject. The preliminary results from 50 subjects showed a better localization precision at 1 kHz frequency, without sex or age predominance. The medial sagital plane presented more errors, in all the frequencies, according to the specialized literature. More subjects are actually being and soon they will be available. (To be presented in Portuguese.)

9:15

2aPP6. Visual bias on sound location modulated by content based processes. Ilja Frissen and Beatrice de Gelder (Tilburg Univ., Tilburg, The Netherlands)

Ventriloquism refers to a perceptual phenomenon in which the apparent location of a sound source is displaced in the direction of a synchronous but spatially disparate visual stimulus. It is generally accepted that spatial and temporal proximity are factors facilitating crossmodal integration. Here we investigate whether content based processes could also play a role. In order to control for strategic factors, a psychophysical staircase method (Bertelson and Aschersleben, 1998) was adopted. Auditory stimuli were digital recordings of vowels (/i/ and /o/). Visual stimuli were digital pictures of talking faces articulating the same vowels, and a scrambled face. We ran eight concurrent staircases. In half of these the auditory stimuli were paired with the corresponding face, and in the other half with the scrambled face. Half the staircases started from the extreme left and the other from the extreme right. Presentation of staircases was randomized. Participants were asked to judge whether the sound was coming from the left, or from the right, of the median plane. On the staircases with a face stimulus, reversals started to occur significantly earlier than with a nonface. Thus, a ''realistic'' stimulus pairing enhances crossmodal integration.

9:30–9:45 Break

9:45

2aPP7. Auditory localization in the horizontal plane with single and double hearing protection. Douglas S. Brungart (Air Force Res. Lab., Wright–Patterson AFB, OH 45433), Alexander J. Kordik (Sytronics, Inc., Dayton, OH 45432), and Brian D. Simpson (Veridian, Dayton, OH 45431)

Although most occupational noise problems can be adequately addressed with either earplugs or earmuffs, some extreme noise environments require listeners to wear both earplugs and earmuffs at the same time. However, little is known about the impact that double hearing protection has on sound localization. In this experiment, listeners wearing single and double hearing protection were asked to localize pink noise signals originating from 24 evenly spaced loudspeakers in the horizontal plane. In the single hearing protection conditions, localization accuracy was severely degraded with a short (250 ms) stimulus, but only modestly degraded with a continuous stimulus that allowed listeners to make exploratory head movements. In the double hearing protection conditions, localization accuracy was near chance level with the short stimulus and was only slightly better than chance with the continuous stimulus. A second experiment showed that listeners wearing double hearing protection were routinely unable to identify the lateral positions of stimuli originating from loudspeakers located at $\pm 45°$ in azimuth. The results suggest that double hearing protection reduces the air-conducted signals in the ear canals to the point that bone and tissue conducted signals disrupt the interaural difference cues listeners normally use to localize sound. [Work supported by AFOSR.]

10:00

2aPP8. Standardization of infrasounds and low-frequency noises for health benefits on humans. Zbigniew Damijan (Structural Acoust. and Biomed. Eng. Lab., Staszic Univ., AGH Krakow, Poland), Ryszard Panuszka (Staszic Univ., AGH Krakow, Poland), and James McGlothlin (Purdue Univ., West Lafayette, IN)

Annoyances from low-frequency noises and infrasound are effects of using technologies, transportation, manufacturing equipment, and large air-condition systems. Present developed procedures for evaluations of annoyance of low-frequency noise (LFN) on humans and occupational health are based on international and national standards. Comparisons show that there is a large difference between permitted values of acoustic pressure levels used. These standards are based on levels of thresholds heard by the human auditory system and subjective observations of impacts from vibrations of infrasound waves on the human body. A newly discovered phenomenon shows a follow-up effect in the brain and is an important reason to check and investigate attenuation of infrasound, especially near and below 10 Hz. Previous investigations showed that nonlinear reaction is observed according to the external infrasound field pressure on the human body. New studies need to investigate electrical reactions of the human brain and electrodermal reactions by influences of infrasound. [Work supported by the Kosciuszko Foundation, Inc., an American Center for Polish Culture, with Funding provided by the Alfred Juzykowski Foundation and KBN Warsaw.]

10:15

2aPP9. Gap detection and location in the precedence effect. Liang Li and Bruce A. Schneider (Dept. of Psych., Univ. of Toronto at Mississauga, Mississauga, ON L5L 1C6, Canada)

The nature of the precedence effect was investigated by introducing a gap into (1) both the leading and lagging sounds, (2) the lagging but not the leading sound, and (3) the leading but not the lagging sound. When a 50-ms gap was introduced into both sounds with an onset asynchrony equal to the delay between the leading and lagging sounds, the gap was perceived to occur only on the leading side as long as the delay between leading and lagging sounds did not exceed approximately 15 ms, even though the precedence effect itself broke down when the delay between the leading and lagging sounds exceeded approximately 9 ms. Gaps presented only in lagging sounds were always heard as occurring in the source position of the leading sound, but no gaps were perceived when the gap occurred only in the leading sound, rather, the listener heard a noise burst from the position of the lagging (suppressed) sound. The present results indicate that gaps in the lagging sound are perceived as belonging

to the leading sound, whereas gaps in the leading sound release the lagging sound from ''echo suppression,'' indicating that higher-order (top–down) processes are involved in the precedence effect.

10:30

2aPP10. Artificial environment mapping from acoustic information. Rodolfo Martinez (CIIDIR, IPN, OAXACA, Mexico, rodolfo_mc@yahoo.com) and Sergio Beristain (Acoustics Lab., ESIME, IPN, Mexico)

Living creatures have the capacity to build maps of the surrounding world from their hearing or ultrasonic perception systems. This mapping is generally assumed to allow survival within a complex environment. When this ability is applied to artificial life or artificial intelligence, the problem becomes very complex. This paper describes some reference models for several species.

10:45

2aPP11. Pitch and loudness memory in musicians and nonmusicians. Peter Bailey and Stuart Dobinson (Dept. of Psych., Univ. of York, York YO10 5DD, UK, pjb1@york.ac.uk)

The finding that pitch and loudness traces decay at different rates [S. Clement, L. Demany, and C. Semal, J. Acoust. Soc. Am. **106**, 2805–2811 (1999)] is one of several results indicating that the processes of pitch and loudness memory may be distinct. A speculation raised by Clement *et al.*, among others, is that these specialized memory subsystems might be differently influenced by musical experience. To explore this hypothesis, difference limens (DLs) for the fundamental frequency (DLF) and intensity (DLI) of complex tones were measured for groups of musically-experienced and musically-naive participants, using a roving-standard, 2-interval procedure in which the duration and content of the interstimulus interval (ISI) within a trial were manipulated. In the first experiment ISIs were silent and 0.5 s or 2.0 s in duration. For both groups DLs increased with ISI; DLFs were smaller for musicians than nonmusicians at both ISIs, but DLIs were smaller for musicians only when ISI=0.5 s, and did not differ when ISI=2.0 s. In a second experiment the ISI included interpolated tones. DLFs were larger for nonmusicians than musicians, but DLIs were similar for both groups. The results suggest that musical experience has different effects on memory fo rpitch and loudness.

11:00

2aPP12. Effect of envelope lowpass filtering on consonant and melody recognition. Arthur P. Lobo, Felipe Toledos, Philip C. Loizou (Dept. of Elec. Eng., Univ. of Texas, Dallas, Richardson, TX 75083), and Michael F. Dorman (Arizona State Univ., Tempe, AZ 75827)

Recent work [Smith *et al.*, Nature **416**, 87–90 (2002)] has shown that the speech envelope contains fine temporal information which is used in pitch perception and spatial localization. This study was performed on normal hearing subjects. In this paper, we investigated the effect of lowpass filtering of the envelope on consonant and melody recognition in subjects using the Clarion cochlear implant. The subjects were originally fitted with the simultaneous analog stimulation (SAS) speech processing strategy, a strategy known to provide fine time-envelope information. The consonants and instrumental music were bandpass filtered into seven channels and the envelope of each channel was lowpass filtered with cutoff frequencies ranging between 100 and 1200 Hz. Initial results on the consonant recognition task showed that some subjects performed equally well for all envelope cutoff frequencies. On the melody recognition task, some subjects performed best at a particular envelope cutoff frequency. Results for the full set of subjects who participated in this study will be presented. [Work supported by NIH.]

11:15

2aPP13. Visual speech recalibrates auditory speech identification. Paul Bertelson (Universite Libre de Bruxelles, Bruxelles, Belgium), Jean Vroomen, and Beatrice de Gelder (Tilburg Univ., The Netherlands)

Exposure to spatially incongruent auditory and visual inputs produces both immediate crossmodal biases and aftereffects. But for event identification, rather than localization, only biases have been demonstrated so far. Taking the case of incongruent audiovisual speech, which produces the well-known McGurk bias effect, we show that, contrary to earlier reports (e.g., Roberts and Summerfield, 1981), aftereffects can be obtained. Exposure to an ambivalent auditory token from an /aba–ada/ continuum combined with the visual presentation of a face articulating /aba/ (or /ada/) increased the tendency to interpret test auditory tokens as /aba/ (or /ada/). The earlier results that were taken as disproving the possibility of visual recalibration of auditory speech identification were obtained with exposure to nonambiguous auditory tokens that (as we confirm in another experiment) create an auditory contrast effect in a direction opposite that of recalibration, and presumably masked the effect of recalibration.

11:30

2aPP14. Factors affecting frequency discrimination by poor readers. Peter Bailey, Maggie Snowling (Dept. of Psych., Univ. of York, York YO10 5DD, UK, pjb1@york.ac.uk), Yvonne Griffiths, and Nick Hill (Univ. of Essex, Colchester CO4 3SQ, UK, ymg@essex.ac.uk)

Some of the experiments on frequency discrimination by groups of poor readers have shown impairments relative to normal-reading control groups, while other experiments have shown no reliable group differences. It remains uncertain to what extent these differences in outcome are attributable to individual differences in the severity of a sensory processing deficit associated with dyslexia, or to use different psychophysical procedures that make different demands of higher-level cognitive processes (such as attention and memory) which may be compromised by dyslexia. To explore these issues, pure-tone frequency difference limens (DLFs) were measured for groups of dyslexic adults and normal-reading controls in conditions incorporating a range of procedural manipulations. Dyslexic and control participants' DLFs did not differ reliably when the procedure involved four-interval trials, but dyslexics' DLFs were larger than controls using two-interval trials. The *relative* difference between dyslexic and control participants' DLFs found using two-interval trials did not differ systematically across conditions involving a fixed or roving standard frequency, long or short duration stimuli, long or short interstimulus intervals, or interstimulus intervals that were either silent or included interpolated tones. The results suggest no obvious link between elevated DLFs and impaired short-term pitch memory in these dyslexic participants.

Session 2aSAa

Structural Acoustics and Vibration: Analysis, Measurements, and Control of Structural Intensity

Sabih I. Hayek, Chair
Department of Engineering Science and Mechanics, Pennsylvania State University, State College, Pennsylvania 16802-6812

Chair's Introduction—8:00

Invited Papers

8:05

2aSAa1. The intensity potential approach. Jean Louis Guyader and Michael Thivant (LVA, INSA de Lyon, 69621 Villeurbanne, France)

Sound intensity vectors can be decomposed into irrotational and curl components. The irrotational part describes intensity propagation, and the curl component vortices and near field of sources. To predict power flow one can limit the problem to irrotational intensity, that is to say to the determination of intensity potential. The resulting equation is analogous to heat transfer, permitting one to use standard heat transfer solvers to predict acoustic power flow. The intensity potential approach is presented for acoustic propagation from sources in cavities with apertures. The modelization of absorbing materials through thermal convection factor is discussed. Finally, comparison with exact prediction and experimental results are presented.

8:35

2aSAa2. Active control of structural intensity and radiated acoustic power from an infinite point-excited submerged Mindlin plate. Jungyun Won and Sabih Hayek (Active Vib. Control Lab., Dept. of Eng. Sci. and Mech., 212 EES Bldg., Penn State Univ., University Park, PA 16802)

In this paper, the active vibrational structural intensity (VSI) in, and the radiated acoustic power from an infinite elastic plate in contact with a heavy fluid is modeled by the Mindlin plate theory. The plate is excited by a point force, which generates a vector-active VSI field in the plate. The resulting acoustic radiation generates an active acoustic intensity (AI) in the fluid medium. The displacement, shear deformation, VSI vector map, radiated acoustic pressure, and the AI vector map are computed. One, two, or four synchronous point controllers are placed symmetrically with respect to the point force on the plate. Minimization of either the structural intensity at a reference point or the total radiated acoustic power is achieved. Below coincidence, a significant portion of the point force input power is trapped in the plate in the form of VSI. The total radiated power is calculated by use of the input power from the source, the controllers, and the VSI. Above coincidence, a significant portion of the input source power is leaked to the fluid in the form of AI, so that the acoustic radiated power is equal to the input power from the source and the controllers.

Session 2aSAb

Structural Acoustics and Vibration: Vibration and Noise Control

J. Gregory McDaniel, Chair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Contributed Papers

9:15

2aSAb1. Active control of cantilever-beam vibration. M. Roman Serbyn (Morgan State Univ., Baltimore, MD 21251, rserbyn@morgan.edu)

A bang–bang control system previously developed for the stabilization of a rigid platform [ISA Trans. **21**, 55–59 (1982)] has been adapted to the problem of reducing flexural vibrations of a beam. The electromechanical system develops an appropriate control signal for the actuator from samples of the disturbance by analog and digital signal processing using integrated circuits. The effectiveness of this approach is predicated upon the sampling rate being much higher than the maximum vibration frequency to be silenced. It is also robust with respect to the waveform of the disturbance. Noise reductions of 10–20 dB have been achieved, depending on the bandwidth of the noise. The cantilever, chosen because of its mechanical and theoretical simplicity, provides a good foundation for the study of more complex structures, like airfoils and nonrigid platforms. In both experimental and analytical investigations the emphasis has been on

the optimization of control parameters, particularly with regard to the application of the cancellation signal. Reduction in size and cost of the control unit is possible by incorporating the latest technological advances in electronic and electromechanical devices, such as FPGA boards and MEMS components.

9:30

2aSAb2. An active structural acoustic control approach for the reduction of the structure-borne road noise. Hugo Douville, Alain Berry, and Patrice Masson (Dept. of Mech. Eng., Universite de Sherbrooke, 2500 Boul. Universite, Sherbrooke, QC J1K 2R1, Canada)

The reduction of the structure-borne road noise generated inside the cabin of an automobile is investigated using an Active Structural Acoustic Control (ASAC) approach. First, a laboratory test bench consisting of a wheel/suspension/lower suspension A-arm assembly has been developed in order to identify the vibroacoustic transfer paths (up to 250 Hz) for realistic road noise excitation of the wheel. Frequency Response Function (FRF) measurements between the excitation/control actuators and each suspension/chassis linkage are used to characterize the different transfer paths that transmit energy through the chassis of the car. Second, a FE/BE model (Finite/Boundary Elements) was developed to simulate the acoustic field of an automobile cab interior. This model is used to predict the acoustic field inside the cabin as a response to the measured forces applied on the suspension/chassis linkages. Finally, an experimental implementation of ASAC is presented. The control approach relies on the use of inertial actuators to modify the vibration behavior of the suspension and the automotive chassis such that its noise radiation efficiency is decreased. The implemented algorithm consists of a MIMO (Multiple-Input–Multiple-Output) feedforward configuration with a filtered-X LMS algorithm using an advanced reference signal (width FIR filters) using the Simulink/Dspace environment for control prototyping.

9:45

2aSAb3. Active control of noise radiated through rectangular plates using piezeletric patches. Danuza Cristina Santana, Marcus Antonio Viana Duarte, and Domingos Alves Rade (School of Mech. Eng., Federal Univ. of Uberlandia, P.O. Box 593, CEP 38400-902 Uberlandia, MG, Brazil)

Due to problems caused by noise in industrial environment and in human daily life, techniques of noise control have received increasing attention from engineers and researchers lately. More recently, the use of piezeletric elements as sensors and/or actuators in noise and vibration control systems has been extensively investigated. The main advantage of the use of such devices is that they can be easily integrated to the mechanical system with little added mass and relatively high control authority. The present paper addresses a technique of active control of sound transmitted through a rectangular, thin, simply supported plate by employing multiple piezeletric patches bonded to the plate's surface. A harmonic plane wave incident on one side of the plate is considered to be the primary noise source. Aiming at minimizing the noise transmitted to the other side of the plate, bending motion is induced through the piezeletric patches so that the plate behaves as a secondary sound source. The paper brings the development of the system mathematical model which enables to obtain the spatial distribution of sound pressure radiated through the plate in the far field. An optimal control technique providing the voltage control signals for the activation of the piezoelectric patches is presented, based on the minimization of a cost function representing the mean square integral of the sound pressure radiated in a semi-sphere in far field. It is also proposed a methodology for the optimal placement on the piezelectric patches using Genetic Algorithms. (To be presented in Portuguese.)

10:00

2aSAb4. Mechanical realization of passive scalar transfer functions. Pierre E. Dupont and Wenyuan Chen (Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, pierre@bu.edu)

There are typically an infinite number of mass-spring-damper systems that share the same input–output characteristics as described by a drive-point accelerance (passive) transfer function. The realization problem of solving for one or more of these mechanical systems arises in several application areas. These include the scaled acoustic testing of complex ship systems as well as the design of mechanical filters. The theory for scalar transfer functions involving no damping or proportional damping is well known. The family of solutions includes realizations consisting of masses connected entirely in parallel or in series; and algorithms are available for computing the components values. The assumption of proportional damping is purely for mathematical convenience, however, and does not usually concur with the reality of experimental data. This talk addresses the mechanical realization problem for scalar accelerance transfer functions with arbitrary viscous damping. A time domain approach is utilized to obtain a parameterized description of the solution set. Numerical methods are presened which can be used to search the solution set for realizations satisfying criteria relating to ease of fabrication, e.g., involving the fewest components. Laboratory experiments are included to validate the approach. [Work supported by ONR.]

10:15

2aSAb5. A vibration technique for measuring railroad rail stress. Vesna Damljanovic and Richard L. Weaver (Dept. of Theoret. & Appl. Mech., Univ. of Illinois, 104 S. Wright St., Urbana, IL 61801, damljano@uiuc.edu)

Longitudinal rail stress, related to constrained thermal contractions and expansions, leads to broken and buckled rails, and consequent service delays and derailments. There is a broad consensus on the need for cost-effective and reliable methods for the measurement of rail stress. Vibration techniques for assessing rail stress, based on the effect of longitudinal force on the free vibrations of beams, have long been proposed. It is well understood that compressive stresses decrease the flexural frequencies, while tensile stresses increase them. Past efforts attempting to use this for measurements of stress in railroad rails have failed due to an inability to control or adequately measure other parameters, most particularly the placement and stiffness of the supports. We are developing a new method in which the influence of support parameters should be minimal. Scanned laser vibrometry measurements of vibration fields, followed by a comparison with guided wave theory for the complex cross section of the rail, promises to allow the stress to be determined with the requisite precision. Here we report on the status of this work. [Work supported by Association of American Railroads and the Transportation Research Board.]

10:30

2aSAb6. Theoretical and experimental study of vibration, generated by monorail trains. Samuil A. Rybak (N. N. Andreev Acoust. Inst., Moscow 117036, Russia), Sergey A. Makhortykh (Russian Acad. of Sci., Pushchino, Moscow Region 142290, Russia), and Stanislav A. Kostarev (Lab. of Acoust. and Vib. Tunnel Assoc., Moscow 107217, Russia)

Monorail transport as all other city transport vehicles is the source of high noise and vibration levels. It is less widespread than cars or underground transport but its influence in modern cities enhances. Now in Moscow the first monorail road with trains on tires is designed, therefore the problem of vibration and noise assessments and prediction of its impact on the residential region appears. To assess the levels of generated vibration a physical model of interaction in the system wagon-tire-road coating-

viaduct-soil has been proposed and then numerically analyzed. The model is based on the known from publications facts of automobile transport vibration and our own practice concerning underground trains vibration generation. To verify computer simulation results and adjust model parameters the series of measurements of noise and vibration near experimental monorail road was carried out. In the report the results of calculations and measurements will be presented and some outcomes of possible acoustical ecologic situation near monorail roads will be proposed.

TUESDAY MORNING, 3 DECEMBER 2002 GRAND CORAL 3, 8:30 TO 11:30 A.M.

Session 2aSC

Speech Communication: Vowel Acoustics and Perception (Poster Session)

Michael J. Kiefte, Cochair
Human Communication Disorders, Dalhousie University, 5599 Fenwick Street, Halifax, Nova Scotia B3H 1R2, Canada

Gonzalo Corvera, Cochair
Monte Athos 116, 11000, D.F. Mexico

Contributed Papers

All posters will be on display from 8:30 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:30 a.m.

2aSC1. Causes and consequences of the unequal distribution of vowels in the vowel space. Peter Ladefoged (Phonet. Lab., Linguist., UCLA, Los Angeles, CA 90095-1543)

A set of possible vocal tract shapes for vowels was generated, using a vocal tract model that operates in terms of the two tongue shape parameters determined by Harshman, Ladefoged, and Goldstein (1975) and one parameter of lip opening. Each of the tongue shape parameters was varied through 15 equal steps. Many of the combinations of the two parameters produced impossible distortions of the tongue or tongue shapes associated with consonants, but 147 of the 225 vowel shapes were humanly possible. These are 147 equally spaced vocal tract shapes as defined by the two parameters. Each of these tongue shapes was combined with seven degrees of lip rounding. When the first two formants of these 1029 vowels were plotted, some parts of this formant space are more densely populated than others. There are few vowels with low $F1$ and low $F2$. The difference among front vowels, such as those in heed, hid, head, had, can be made simply by varying the vocal tract shape, but the back vowels, such as those in hawed, hood, who'd, require added lip rounding. In the world's languages there is often an asymmetry in the height of front and back vowels.

2aSC2. Realization of the English [voice] contrast in $F1$ and $F2$. Elliott Moreton (Dept. of Cognit. Sci., Krieger Hall, Johns Hopkins Univ., Baltimore, MD 21218)

Before a [voice] coda, $F1$ is higher for monophthongs but lower for /aI/ than before [+voice]. We test the hypothesis that this is due to local hyperarticulation before voiceless obstruents. Experiment 1, with 16 American English speakers, found the /aI/ pattern of more peripheral $F1$ and $F2$ in the offglides /oI eI aU/ as well, showing that it is part of the realization of [voice] rather than a historical property of /aI/. Some of the $F2$ increase in /aI oI eI/ cannot be accounted for by articulatory raising alone, but must be ascribed to fronting. The diphthong nuclei tended to change in the same direction as the offglides. Experiments 2 and 3, each with a different 16 American English speakers, collected ''tight''-''tide'' (Exp. 2) or ''ate''-''aid'' (Exp. 3) judgments of a synthetic stimulus in which offglide $F1$, offglide $F2$, and nuclear duration were varied independently. ''Tight'' and ''ate'' responses were facilitated by lower $F1$, by higher $F2$, and by shorter nuclei. Log-linear analysis showed that the three factors contributed independently, and that $F2$ was a stronger cue than $F1$ in terms of logits per Bark. Thus [voice] is correlated with, and cued by, peripheralization of diphthong offglides.

2aSC3. Influence of formant transitions and linguistic relevance on vowel imitation. Gautam K Vallabha and Betty Tuller (Ctr. for Complex Systems & Brain Sci., Florida Atlantic Univ., 777 Glades Rd., Boca Raton, FL 33431, vallabha@walt.ccs.fau.edu)

Speakers are unable to imitate the formants of isolated self-produced vowels accurately. The pattern of bias directions cannot be ascribed solely to articulatory or perceptual noise, and is different for each speaker [G. K. Vallabha and B. Tuller, J. Acoust. Soc. Am. **110**, 2657 (2001)]. The present experiment examines whether the bias pattern persists when the vowels are embedded in a CV syllable. A dialectically uniform group of male and female American English [AE] speakers produced vocalic sounds in three conditions: (1) they read a list of /hVd/ words 10 times each, (2) they imitated /V/ and /dV/ targets containing monophthongal AE vowels and (3) they imitated /V/ and /dV/ targets that were systematically distributed over the $F1\times F2$ space. All targets were self-produced (hence perfectly reproducible in principle) and were imitated 10 times each. Results will be discussed in terms of the bias and variability of AE versus non-AE targets, the imitation bias of /dV/ vs /V/ targets, and whether the bias pattern differences among subjects are reduced by the dialect uniformity. [Work supported by NIMH.]

2aSC4. Context effects on the perception of vowel spectral properties. Michael Kiefte (School of Human Commun. Disord., Dalhousie Univ., 5599 Fenwick St., Halifax, NS B3H 1R2, Canada) and Keith R. Kluender (Univ. of Wisconsin, Madison, WI 53706-1696)

Previous studies [M. Kiefte and K. R. Kluender, J. Acoust. Soc. Am. **57**, 55–62 (2001)] have shown that long-term stationary spectral characteristics have a very strong influence on the perception of specific cues in monophthongs. For example, F_2 will be perceptually neglected if a precursor sentence is filtered by a one-pole filter corresponding to the center

frequency and bandwidth of the target vowel. The present work continues this research and presents results from experiments using nonspeech precursors. The first set of experiments use speechlike stimuli that have been LPC vocoded from a time-reversed sentence having similar spectro-temporal properties to natural speech. The second set of experiments use FM-formant glides as precursors to test multiple hypotheses including the time course of these effects over varying precursor durations. The second set of experiments also test the hypothesis that the duration of the precursor plays a role in the magnitude of the context effect on the perception of the following vowel.

2aSC5. Acoustic correlates of English glottal [t] allophone. Rina Kreitman (Dept. of Linguist., Cornell Univ., 203 Morrill Hall, Ithaca, NY 14850, rk83@cornell.edu)

This experiment investigates the acoustics of American English coda [t], which is often glottalized. A list of monosyllabic and bisyllabic words with matching vowel quality in target syllables across pairs was devised. The matching pairs were controlled for lexical frequency from the CELEX database. The words were read in isolation and in a frame sentence, where the target word was at the end of an intonational phrase. Results show that there is glottalization on the vowel preceding [t], which is interpreted as laryngeal coarticulation from the glottal allophone of [t]. Additionally, results show that for some speakers the proportion of the vowel that is glottalized is longer in sentences than in isolated words. Results of a cepstral measure of harmonics-to-noise-ratio (Hillenbrand *et al.*, 1994) are different from what was found for Ju–hoansi glottalized consonants (Miller-Ockhuizen, 2002), but similar to results for glottalized vowels. That is, Gamnitude of R1 increased toward the end of the vowel. Since glottalization, as measured via a low H1–H2 (NiChasaide and Gobl, 1997) and aperiodicity seen through waveform inspection, does occur on the vowel, this suggests an articulatory mechanism for glottalization different from constriction of the glottis, such as false vocal-fold contraction (Fujimura and Sawashima, 1971).

2aSC6. Effects of frequency shifts on vowel category judgments. Catherine M. Glidden, Peter F. Assmann, and Terrance M. Nearey (School of Human Development, Univ. of Texas, Dallas, Box 830688, Richardson, TX 75083)

To investigate the effects of fundamental frequency ($F0$) and formant frequency shifts on vowel identification, a high-quality vocoder (''STRAIGHT'') was used to process the syllables ''bit'' and ''bet'' spoken by an adult female talker. From these two endpoints a nine-step continuum was generated by interpolation of the time-varying spectral envelope. Upward and downward frequency shifts in spectral envelope (scale factors of 0.75, 1.0, or 1.33) were combined with shifts in $F0$ (scale factors of 0.5, 1.0, or 1.25). Downward frequency shifts generally resulted in malelike voices whereas upward shifts were perceived as childlike. Matched frequency shifts, in which $F0$ and spectral envelope (i.e., formant frequencies) were shifted in the same direction, had relatively little effect on phoneme boundaries. Mismatched frequency shifts, in which $F0$ was modified independently of spectral envelope or *vice versa*, resulted in systematic boundary shifts. The changes in the identifications functions were qualitatively consistent with predictions of a model trained using acoustic measurements derived from a database of naturally spoken vowel tokens from men, women, and children. The empirical and modeling results are consistent with the idea that vowel boundary shifts are a consequence of listeners sensitivity to the statistical structure of natural speech.

2aSC7. Context dependencies in vowel identification in ablated CVC syllables. Jim Talley (Dept. of Linguist., Univ. of Texas, Austin, TX 78712, jtalley@acm.org)

In previously reported work [Talley, J. Acoust. Soc. Am. **108**, 2601 (2000)], novel results from a new perceptual study of human vowel identification under ablation conditions were discussed. That study, which used ten American English (AE) vowels in each of four simple CVC consonantal contexts, found highly significant effects of ablation condition and consonantal context on vowel identifiability. However, little insight was available at the time regarding the specifics of the vowel–context interactions. This paper extends that work providing detailed analysis of vowel identification sensitivity relative to consonantal context under differing ablation conditions.

2aSC8. Lexical duration effects in Japanese function particles. Setsuko Shirai (Dept. of Linguist., Univ. of Washington, Box 354340, Seattle, WA 98195-4340)

It is well known that the lexical status, i.e., content versus function, influences vowel quality and duration in English. However, it is not clear if function particles in Japanese are reduced in a way that is equivalent to English. Previous studies [for example, N. Campbell, in *Speech Production and Linguistic Structure,* 1992] point out that Japanese particles tend to be long because of phrase-final lengthening effects and that function particles and content words tend to have different segmental content. Therefore, any study of the reduction in function particles must control for word and phrase position and must compare segmentally matched tokens. In this study function particles containing three vowels, [e, a, o], are compared with equivalent syllables in content words controlling for segmental context, the number of syllables, and for word and phrase position. For example, a content word [kiga] is matched with a token containing a function particle [ki-ga]. For 11 speakers of Tokyo Japanese, there was a significant effect for lexical category [$F(1, 426)$, $p<0.01$]. When split for vowel, [a] maintained a significant effect, while the other two vowels only maintained a trend towards reduced duration in function particles.

2aSC9. Speech perception based on spectral peaks versus spectral shape. James M. Hillenbrand (Speech Pathol. and Audiol., Western Michigan Univ., Kalamazoo, MI 49008) and Robert A. Houde (RIT Res. Corp., Rochester, NY 14623)

Some spectral details are more intimately associated with the transmission of phonetic information than others, and a good deal of phonetic perception research has involved drawing inferences about the nature of the spectral representations that mediate phonetic recognition by conducting listening experiments using speech signals which are contrived in such a way as to retain only some characteristics of the speech signal, while purposely removing or distorting other spectral details. The present study was designed to address one aspect of this problem having to do with the relative contributions of spectral envelope peaks versus the detailed shape of the spectral envelope. The problem was addressed by asking listeners to identify nonsense syllables that were generated by two structurally identical source-filter synthesizers, one of which constructs the filter function based on the detailed spectral envelope shape, while the other uses a coarse estimate that is constructed entirely from the distribution of peaks in the envelope. Results suggest that nearly as much phonetic information is conveyed by the relatively coarse peaks-only representation method as by the method that preserves the fine details of the original envelope. However, there is a modest but reliable increase in the transmission of phonetic information when the detailed envelope shape is preserved.

2aSC10. Temporal characteristics of Chinese-accented English: Preliminary findings. Bruce L. Smith (Dept. of Commun. Disord., Univ. of Utah, 390 S. 1530 E., Rm. 1201 Behavioral Sci. Bldg., Salt Lake City, UT 84112-0252), Ann R. Bradlow, and Tessa Bent (Northwestern Univ., Evanston, IL 60208)

This study investigated several temporal features of English to determine the extent of their occurrence in the speech of talkers of Chinese-accented English who had relatively limited experience with spoken English. Specifically, the extent to which these speakers produced the following temporal contrasts was examined: (1) tense versus lax vowel duration, (2) vowel duration before voiced versus voiceless consonants, and (3) vowel and consonant duration in sentence-final versus nonfinal position. Preliminary data from sentences produced by eight non-native and eight native talkers indicates that the native English speakers and the Chinese-accented talkers did not differ in the extent to which they realized the inherent duration difference between tense and lax vowels. However, the native English subjects tended to show substantially greater vowel lengthening before voiced versus before voiceless consonants than the Chinese-accented talkers. In addition, while the two groups did not differ significantly in the extent to which they lengthened sentence-final consonants relative to nonfinal consonants, the native English talkers showed greater sentence-final vowel lengthening than the Chinese-accented talkers. When group differences for a given temporal parameter were found, 1 or 2 of the non-native subjects typically fell within the range of performance shown by the native speakers.

2aSC11. Duration and rate effects on American English vowel identification by native Danish listeners. Terry L. Gottfried (Dept. of Psych., Lawrence Univ., Appleton, WI 54912) and Ocke-Schwen Bohn (Aarhus Univ., DK-8000 Aarhus C, Denmark)

Native listeners alter their identification of American English vowel contrasts according to speaking rate, apparently making judgments about the relative duration of vowels. Berman *et al.* [J. Acoust. Soc. Am. **105**, 1402 (1999)] created a series of spectral continua (varying $F1$ and $F2$) from ''beat'' to ''bit,'' ''pat'' to ''pet,'' and ''cot'' to ''cut,'' also varying syllable duration according to natural speech endpoints. These syllables were inserted into natural speech sentence contexts of two rates (normal and fast). Berman *et al.* found that longer syllable duration led to more long vowel (''beat,'' ''pat,'' ''cot'') responses; faster rate contexts also led to more long vowel responses. In the present research we tested native speakers of Danish on their use of duration and rate context in identifying these English vowels. Danish listeners were significantly affected by the duration of syllables in their vowel identification. However, despite vowel duration being phonemic in Danish, native Danish listeners were not significantly affected by rate context in their identification of these English vowels. This might be explained by perceived sufficiency of spectral information in these English vowel contrasts for Danish listeners, or by the lack of rate dependent vowel processing in their native language. [Work supported by Danish-American Fulbright Commission.]

2aSC12. Effects of perceptual assmilation on the production of English vowels by native Japanese speakers. Takeshi Nozawa (Kansai Univ. of Intl. Studies, 1-18 Aoyama Shjimi-cho Miki, Hyogo 673-0521, Japan) and Elaina M. Frieda (Auburn Univ., Auburn, AL 36849)

In our previous studies, it was found that English vowel contrasts that were difficult for Japanese speakers to discriminate were for the most part identified with the same Japanese vowels. This finding complies with one the Perceptual Assimilation Model (Best, 1995) postulates. The present study investigated how perceptual assimilation affects production of English vowels by Japanese speakers. Three experienced Japanese learners in Columbus, Ohio, and five inexperienced Japanese learners in Kobe, Japan produced English vowels, repeating after two different native speakers' utterances. Each subject heard the same token twice in a different order, so they each produced four tokens of each English vowel. The results support what our previous studies have found, and show the effects of perception on production. The production of the vowels difficult to discriminate tended to be close to each other within subjects' production vowel space. The production of the vowels identified as good exemplars of Japanese vowels were more stable than those identified as poor exemplars. Thus, the subjects /i/ tokens were almost always identified as /i/, whereas their /æ/ tokens were often identified as /ɛ/, /ɑ/, or /ʌ/. Experienced learners' utterances were far more correctly identified.

2aSC13. On the relationship between perception and production of American English vowels by native speakers of Japanese: A pair of case studies. Kanae Nishi (Speech & Hearing Sci., City Univ. of New York–Grad. Ctr., 365 Fifth Ave., New York, NY 10016) and Catherine L. Rogers (Univ. of South Florida, Tampa, FL 33620)

In a previous study, 15 American English (AE) vowels in /hVd/ words recorded by native speakers of Japanese (J) were presented in pairs to native speakers of AE. From these data, native-perceived vowel spaces of J-accented AE were obtained (Nishi, 2001). In the present study, two male native speakers of Japanese who had served as speakers in the previous study listened to the 15 AE vowels produced by two male native speakers of AE in a similarity rating task. Their perceptual data were analyzed using multidimensional scaling. J-perceived vowel spaces of AE vowels produced by AE speakers were compared to AE-perceived vowel spaces of J-produced AE vowels. Vowels produced by both AE and J speakers were also subjected to acoustic analysis. The acoustic vowel spaces obtained were then compared to the AE and J perceptual vowel spaces. Results revealed considerable differences between the two J speakers in terms of their perception of AE vowels. These differences were found to be strongly correlated with the speakers' vowel spaces as perceived by AE listeners, as well as with their acoustic vowel spaces.

2aSC14. Assimilation and discrimination of Canadian French vowels by English-speaking adults. Linda Polka (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, linda.polka@mcgill.ca), Paola Escudero (Utrecht Univ., The Netherlands), and Shelly Matchett (McGill Univ., Montreal, QC H3G 1A8, Canada)

According to the Perceptual Assimilation Model (PAM) discriminability of non-native contrasts depends on the perceived similarity to native phonetic categories [Best (1994)]. With respect to vowel contrasts, it has been claimed that PAM predictions hold so long as context-specific realizations of a given contrast are considered [Strange *et al.* (2001); Levy (2002)]. To further assess these claims, English adults were tested on perception of Canadian French tense high vowels, /i/, /y/, /u/ and their respective context-conditioned variants [I], [Y], and [U]. Adults completed 2 tasks with both citation and sentence context stimuli; stimuli for both tasks were natural /bVs/ tokens, produced by male and female talkers. For each vowel, subjects completed an identification and rating task (using English vowel response categories) which provided data to assess the assimilation pattern for four non-native contrasts: /i-y/, /y-u/, /I-Y/, and /Y-U/. Discrimination was assessed for each contrast using a categorical AXB task, in which each token is produced by a different talker. The ability to predict discrimination differences from assimilation data was assessed. The findings were examined together with acoustic analysis to determine whether acoustic differences and assimilation differences are equally predictive of relative discriminability.

2aSC15. The identification and discrimination of English vowels produced by native Mandarin speakers. Yang Chen (Univ. of Wyoming P.O. Box 3311, Laramie, WY 82072) and Michael Robb (Univ. of Connecticut, U85, Storrs, CT 06269)

Individuals who speak an Asian language as their first language (L1) are reported to show phonetic inaccuracies in their production of English spoken as a second language (L2) [Flege, 1989]. Phonetic inaccuracies are

assumed to arise from the phonetic differences between the two languages, whereby L1 competes with L2 production [Flege *et al.*, 1997]. A small database is currently available in regarding the influence of Mandarin on the production of American English vowels. The present investigation is to perform a detailed acoustic and perceptual evaluation of American English tense–lax vowel pairs produced by 40 native Mandarin speakers. The Mandarin language does not contain lax vowels, therefore, examining the production of American English tense–lax pairs will provide a critical assessment of a native Mandarin-speaking individual's ability to differentially produce English vowels. The first hypothesis to be tested will be that Mandarin speakers will demonstrate indistinctive acoustic tense–lax vowel differentiation in their English productions. The second hypothesis to be tested will be that Mandarin speakers will be judged perceptually as less successful in differentially producing the adjacent tense–lax English vowel pairs as compared to the American speakers. The phonetic influences of L1 (Mandarin) on L2 (English) will be discussed.

2aSC16. Improving English vowel perception and production by Spanish-speaking adults. Karen Stenning and Donald G. Jamieson (Natl. Ctr. for Audiol., Univ. of Western Ontario, London, ON N6G 1H1, Canada)

This study investigated the effects of perceptual and production training on the abilities of adult native speakers of Spanish to identify and produce the English vowels /i, I, e, ɛ, æ/. Testing was performed prior to and following an average of 12 h of perceptual training in a category inclusion task, then again following an average of 7.5 h of production training involving visual feedback of vowel formant ($F1$ and $F2$) values. A lagged control group of participants who were delayed in starting training showed no improvement in perception and production skills during the control period but changed equivalently during training. The mean improvement in perceptual identification accuracy for /i, I, e, ɛ/ was 17%. The mean improvement in the intelligibility of participants productions of /I, ɛ, æ/ following training was 11%. For both perception and production, most improvement occurred during perceptual training.

2aSC17. Production and perception of vowels in Karitiana. Didier Demolin (Free Univ. of Brussels, 50 av. F. D. Roosevelt, 1050 Brussels, Belgium) and Luciana Storto (Universitade de Sao Paolo, S.P., Brazil)

The main acoustic features of vowels in Karitiana, a language of the Tupi stock spoken in Brazil are examined. This language has 5 vowel qualities [i,e,a,o,ɨ] which can be oral (short and long) and nasal (short and long). This vowel system has no high back vowel [u]. The main characteristics are that short oral vowels have a wider distribution than long oral vowels; nasal vowels are centralized when compared to oral vowels; length measurements show that central vowels are very short; whatever their quality short nasal vowels have similar length; the difference between short and long nasals is less than in the oral dimension. The dynamic characteristics of nasal vowels where there is a characteristic rising movement of $F2$ are also examined. Finally a perception test was carried out to understand how Karitiana vowels are perceived by native speakers. The test used a set of 58 synthetic stimuli which covers the vowel space. The main results show that speakers easily recognize peripheric vowels while central vowels are less salient. Karitiana speakers did not identify any stimuli in the area of the high back vowel.

2aSC18. Effects of hearing status and perturbation with a bite block on vowel production. Jennell C. Vick, Joseph S. Perkell, Harlan Lane, Melanie Matthies, Majid Zandipour, Ellen Stockmann, Frank Guenther, and Mark Tiede (Speech Commun. Group, Res. Lab. of Electron., MIT, Cambridge, MA 02139)

This study explores the effect of hearing status on adaptation to a bite block in vowel productions of normal hearing (NH) adults and adults who use cochlear implants (CI). CI speakers are tested prior to and following experience with the implant. Different sized bite blocks (BB) are used to create unusual degrees of mandibular opening for vowel productions in an /hVd/ context (had, head, heed, hid, and hod). Four conditions are elicited from each NH and CI speaker: (1) no BB with hearing (CI processor on), (2) no BB with no hearing (NH speakers with masking noise and CI speakers with processor off), (3) BB with no hearing, and (4) BB with hearing. Prior to fitting with the implant, CI speakers are tested without hearing in two conditions: (1) no BB and (2) BB. Spectra of the vowel productions are analyzed for dispersion of tokens in the $F1$–$F2$ plane in the four conditions. Pilot results support the hypothesis that prior to fitting, CI users are less able to adapt to perturbations than NH speakers and that experience with a CI improves adaptation. The current study is exploring this result further with larger groups of subjects. [Research supported by NIH.]

2aSC19. Mothers exaggerated acoustic-phonetic characteristics in infant-directed speech are highly correlated with infant's speech discrimination skills in the first year of life. Hue-Mei Liu, Patricia K. Kuhl, and Feng-Ming Tsao (Ctr. for Mind, Brain and Learning, CMBL 357988, Univ. of Washington, Seattle, WA 98195, liumei@u.washington.edu)

In addition to the well documented suprasegmental features of infant-directed speech, recent studies demonstrate that critical acoustic-phonetic features are exaggerated at the level of segments in infant-directed speech (IDS) [Kuhl *et al.* (1997); Liu *et al.* (2000); Burnham *et al.* (2002)]. The speech directed to infants contains acoustically more extreme vowels when compared to adult-directed speech, resulting in an expanded vowel space. The relationship between the exaggerated characteristics of IDS and infants' speech perception skills has not previously been examined. The present study tested the hypothesis that there is a correlation between characteristics of speech input and speech perception sensitivity for individual infants in the first year of life. Thirty-two Mandarin Chinese mothers' speech samples were recorded in infant-directed and adult-directed speech conditions. The acoustic analysis demonstrated that Mandarin mothers modify their speech when speaking to their infants, amplifying the important acoustic cues. More importantly, the acoustic-phonetic characteristics in a particular mother's infant-directed speech are significantly correlated with her infant's speech perception. This result supports the view that hyperarticulation of the phonetic units in infant-directed speech facilitates infant's language development. [Work supported by NIH grant (HD37954) and Talaris Research Institute.]

2aSC20. Effect of age and context on vowel area. Megan M. Hodge (Dept. of Speech Pathol. & Audiol., Univ. of Alberta, Rm. 2-70 Corbett Hall, Edmonton, AB T6G 2G4, Canada, megan.hodge@ualberta.ca)

Vowel quadrilateral area, based on first and second formant measures of the corner vowels, has been shown to be correlated positively with intelligibility scores in studies of adults with and without dysarthria [Weismer *et al.*, 2001] and children with and without dysarthria [Higgins and Hodge, 2002]. This study compared vowel areas of three groups of typical talkers (3 year-olds, 5-year-olds, and women) using a log Hz scale in two different speaking conditions. The first speaking condition was production of multiple tokens of isolated /hV/ syllables for each corner vowel, and the second was production of single word items containing the four corner vowels taken from a childrens' test of intelligibility [Hodge, 1996]. An interaction of age with phonetic context was found. Vowel areas for the 3-year-olds did not differ between the two conditions and were the largest of the three groups. The 5-year-olds' and women's vowel areas were of similar size and both were smaller in the word condition. The women

showed the greatest effect of phonetic context with a significantly smaller mean vowel area in the word than isolated vowel condition. [Work supported by Glenrose Rehabilitation Hospital and Canadian Languaage and Literacy Research Network.]

2aSC21. Rhythmic patterns in the speech of developmental apraxia of speech and articulation-disordered children. Maria Passadakes, Fredericka Bell-Berti (Dept. of Speech, Commun. Sci., and Theatre, St. John's Univ., Jamaica, NY 11439, bellf@mail.stjohns.edu), Joanne Paoli (St. John's Univ., Jamaica, NY 11439), and Carole Gelfer (William Paterson Univ., Wayne, NJ 07470)

Many disorders, particularly those of neurological origin, involve disturbances in timing control. Research on children's speech is needed to better understand children's acquisition of phonological processes, and how it goes awry. Studies investigating timing aspects of the speech of developmentally apraxic and articulation-disordered children are lacking. The purpose of this research is to describe both phrase-final lengthening and compensatory shortening patterns of apraxic and articulation-disordered children, as well as normal children, under 5 years old.

2aSC22. Guidelines for the acoustical conditioning of classrooms in educational buildings. Jorge Alberto Mastroizzi, Carmen Montes, Susana Amura, and Maria Amelia Mastroizzi (Universidad Argentina John F. Kennedy, Gabinete de Investigacin y Vinculacin Technolgica, Rivarola 139 37, (1015) Buenos Aires, Argentina)

The purpose of this work is to identify deficiencies in acoustical conditions within classrooms. Taking the classroom as a representative educational space, acoustical requirements will be determined as a function of the grade level of the students being taught. Sources of noise that affect the classrooms will be examined, assessed, and corrective measures proposed.

TUESDAY MORNING, 3 DECEMBER 2002 GRAND CORAL 1, 7:55 A.M. TO 12:00 NOON

Session 2aUW

Underwater Acoustics: Geoclutter and Boundary Characterization I

Charles W. Holland, Cochair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

John R. Preston, Cochair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

Chair's Introduction—7:55

Contributed Papers

8:00

2aUW1. The Boundary Characterization 2001 Experiment. Charles Holland, Kevin LePage, Chris Harrison (SACLANT Undersea Res. Ctr.), Paul Hines, Dale Ellis, John Osler, Dan Hutt (DRDC-A), Roger Gauss, Redwood Nero (Naval Res. Lab.), and John Preston (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

The weakest link in performance prediction for naval systems operating in coastal regions is the environmental data that drive the models. In shallow water downward refracting environments, the seabed properties and morphology often are the controlling environmental factors. In the Boundary 2001 Experiment, seabed, surface, and biologic scattering, seabed reflection, propagation, reverberation, and ambient noise data were collected in order to develop and refine measurement techniques for key environmental model inputs. Both Rapid Environmental Assessment (REA) methods and high-resolution measurement techniques were employed from 0.1–10 kHz. Supporting oceanographic, geologic, and geophysical data were also collected. The experiment was conducted in May 2001 in two littoral regions: the New Jersey shelf and the Scotian Shelf. This paper provides an overview of the experiment objectives, hypotheses and conduct. [Research supported by NATO SACLANT Undersea Research Centre, ONR, and DRDC-A.]

8:15

2aUW2. Measurements of acoustic backscatter at shallow grazing angles at low kHz frequencies. Paul C. Hines, John C. Osler, and Darcy J. MacDougald (Defence R&D Canada Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, paul.hines@drdc-rddc.gc.ca)

The acoustic backscattering strength of the seabed has been demonstrated to be one of the key inputs required in sonar performance prediction models. Yet, direct measurement of acoustic scattering from the seabed at shallow grazing angles presents a considerable challenge in littoral waters. The DRDC Atlantic (formerly DREA) Wide Band Sonar system, which consists of a parametric transmitter and a superdirective receiver, is ideally suited to make this measurement. The system was used to measure backscatter as a function of grazing angle and azimuth on the Scotian Shelf off the coast of Nova Scotia and on the Strataform site off the coast of New Jersey. Interpretation of the data set is enhanced with swath bathymetry measurements made at one of the experimental sites. In this paper the experimental geometry is described and the backscatter measurements are presented and discussed in light of the swath bathymetry results.

8:30

2aUW3. Measurements of signal spread and coherence on the New Jersey Shelf and in the Straits of Sicily using time-forward and time-reversed signals. Roger C. Gauss and Richard Menis (Naval Res. Lab., Code 7144, Washington, DC 20375-5350, roger.gauss@nrl.navy.mil)

Mid-frequency shallow-water propagation measurements were made at a variety sites on the New Jersey Shelf and in the Straits of Sicily during three joint trials with the SACLANTCEN (Boundary Characterization: 2000–2002) in order to extract measures of signal spread and coherence, and to evaluate the spatial robustness of time-reversal techniques. The experiments had the NRV Alliance periodically transmitting a 1-s, 200-Hz-bandwidth LFM while it traversed an arc about a stationary platform, which captured these signals and transmitted back time-forward and a set of time-reversed versions of them, as well as transmitting its own version of the original signal. Analysis of normalized matched-filter data indicate that while the two-way time spreads were generally modest, marked decorrelation was observed and that time reversal did well in reconstructing the original impulse response of the LFMs. As the various stored time-reversed versions of the original signals corresponded to different source–receiver paths, a comparison of their signal correlations provided a mea-

sure of the spatial robustness of phase conjugation in the different shelf environments. That statistically the stored time-reversed signals often did as well as *in situ* time-reversed signals suggests the potential for bistatic applications of phase-conjugation concepts. [Work supported by ONR.]

8:45

2aUW4. Analysis of monostatic and bistatic reverberation measurements on the Scotian Shelf. Dale D. Ellis (DRDC Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada) and John R. Preston (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

During the Boundary 2001 sea trial, a number of long-range reverberation measurements were made at a site on the Scotian Shelf between Halifax and Sable Island. The water depth was about 80 m over a sandy bottom. The sources were SUS charges dropped from either the NATO research vessel Alliance or the Canadian research vessel CFAV Quest. The receiver was the SACLANTCEN 254-m towed array aboard Alliance. The data were analyzed in two array apertures and frequencies from 80 Hz to 1400 Hz. Model-data comparisons were made using the Generic Sonar Model, with bottom parameters being extracted using both a manual procedure and an automated procedure. Direct measures of the scattering and bottom properties were made by other researchers, including Hines, Holland, and Osler. The bottom topography was fairly smooth near the site, but at long ranges there were numerous scattering features. The polar plots of the data, and model-data differences, are compared with the bathymetric features in the area. [Work supported in part by ONR Code 32, Grant No. N00014-97-1-1034.]

9:00

2aUW5. Shallow-water reverberation highlights and bottom parameter extractions from the STRATAFORM. John R. Preston (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804) and Dale D. Ellis (DRDC Atlantic, Dartmouth, NS B2Y 3Z7, Canada)

Together with SACLANTCEN, the authors recently participated in the Boundary Characterization Experiment to measure shallow-water bottom reverberation in the STRATAFORM off New Jersey. SUS charges were used as monostatic sources. The receivers were horizontal arrays. Data were analyzed in bands from 160–1500 Hz. The STRATAFORM is known to have benign surface morphology but contains many buried river channels. Highlights of the reverberant returns are discussed that include returns from over the shelf break. Some comparisons in reverberation characteristics between SUS and coherent pulses are noted. Another objective of these reverberation experiments was to quickly invert for bottom scattering and bottom loss parameters. An automated geo-acoustic parameter extraction method was used together with the Generic Sonar Model and a Jackson-Mourad model for scattering. After automatically adjusting bottom loss and scattering strength, good agreement is achieved between the diffuse reverberation data and model predictions in relatively flat areas. Model/data differences are generally correlated with bottom scattering features. Since reverberation typically lasts 10–20 s or more, extracted parameters apply over wide areas. Local bottom loss and backscattering measurements were made by Holland in these areas. A comparison with Holland's results is given. [Work supported by ONR Code 32, Grant No. N00014-97-1-1034.]

9:15

2aUW6. Scattering strength uncertainty. Chris H. Harrison (SACLANT Undersea Res. Ctr., Viale San Bartolomeo, 400, 19138 La Spezia, Italy, harrison@saclantc.nato.int)

A serious weakness in modeling shallow water reverberation is the uncertainty in bottom scattering strength and its angle-dependence. If the bottom scattering law is assumed to be a separable function of an incoming and outgoing angle it follows that the reverberation contains separable incoming and outgoing propagation terms. Thus the returning multipaths from a scattering patch are weighted directly by (the outgoing part of) the scattering law. This means that comparisons of reverberation and propagation angle-dependence on a vertical receiving array have the potential to reveal the scattering law directly. In this paper we discuss a reverberation experiment with complementary propagation measurements using a VLA and a broadband source to deduce scattering law angle-dependence and absolute scattering strength. The approach is justified by some analysis, and findings are compared with the numerical results of a new multistatic sonar model, SUPREMO. The experiment was conducted in a fairly flat bottomed part of the Mediterranean south of Sicily during BOUNDARY2002.

9:30

2aUW7. Measurements of mid-frequency boundary scattering on the New Jersey Shelf and in the Straits of Sicily. Edward L. Kunz and Roger C. Gauss (Naval Res. Lab, Code 7140, Washington, DC 20375-5350, edward.kunz@nrl.navy.mil)

During joint trials with the SACLANTCEN, direct-path monostatic scattering measurements were conducted at 18 sites across the New Jersey Shelf (Boundary Characterization 2001) and at 9 sites in the Straits of Sicily (Boundary Characterization 2002). Using combinations of short duration cw and LFM signals, both the mean (scattering strength) and statistical (probability density function) characteristics of the bottom- and surface-interaction zones were measured at each site. Bottom-zone scattering strength results show that many sandy and rocky sites exhibited a generally flat dependence on grazing angle (over 20–70 deg) and a moderate dependence ($\sim$5 dB) on frequency over 2.5–5 kHz. In contrast, the measured surface-zone scattering strengths exhibited a strong dependence on grazing angle (and a mild dependence on frequency), consistent with scattering from the rough air–sea interface. Using physics-based scattering models, coupled with the supporting environmental measurement results, estimates of both the relative contributions of different bottom-zone scattering mechanisms (water-sediment interface, sediment volume and, at low grazing angles, near-bottom fish) and geophysical quantities (such as bottom roughness spectral parameters) were derived and will be presented along with the acoustic data results. [Work supported by ONR.]

9:45

2aUW8. Geoacoustic characterization of seabed scattering experiment locations. John Osler and Blair Lock (Defence R&D Canada Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

Measurements of acoustic forward and backscattering have been made by DRDC Atlantic during collaborative sea-trials on the Scotian Shelf off the coast of Nova Scotia and on the Strataform site off the coast of New Jersey. In this paper, the geoacoustic properties and roughness parameters that are necessary to interpret and model the scattering measurements are presented. They have been determined using complementary *in situ* and acoustic techniques. The *in situ* measurements have been made using grab samples and a free fall cone penetrometer that has been fitted with a resistivity module. The probe provides two independent means of calculating the undrained shear strength, an empirical sediment classification and sediment bulk density. The acoustic measurements include inversions for geoacoustic parameters using the WARBLE [Holland and Osler, J. Acoust. Soc. Am. (2000)] and normal incidence sediment classification [Hines and Heald, Proc. Inst. Acoust. (2001)] techniques. At the scattering experimental locations, these measurements have been combined with surveys using commercial equipment: sidescan sonar, multibeam bathymetry, and subbottom profilers to characterize the seabed.

10:00–10:15 Break

10:15

2aUW9. Bistatic reverberation modeling for range-dependent waveguides. Kevin D. LePage[a] (Naval Res. Lab., Washington, DC 20375-5350, kevin.lepage@nrl.navy.mil)

The BiStaR bistatic reverberation model has been developed at SACLANTCEN to model general reverberation scenarios in range-dependent waveguides. Features of the model include coherent propagation to and from the scattering patch, a coherent scattering patch model

which includes anisotropy in the scatterer correlation length scales, and explicit inclusion of system parameters such as bandwidth and source and receiver directivity. The model is built upon the C-SNAP coupled mode model developed at SACLANTCEN by Ferla and Porter. In this talk the theoretical features of the model are presented and the characteristics of coherent and range-dependent bistatic reverberation predicted with the model are discussed. The model is also exercised to help interpret reverberation phenomena observed during the Geoclutter/Boundary Characterization Cruise conducted off the coast of New Jersey during the spring of 2001, including reverberation from beyond the shelf break and scattering from features buried beneath the sediment. [a] Previously at the SACLANT Undersea Research Centre, La Spezia, Italy.

10:30

2aUW10. Comparison of *in situ* compressional wave speed and attenuation measurements to Biot–Stoll model predictions. Barbara J. Kraft, Larry A. Mayer (Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, 24 Colovos Rd., Durham, NH 03824), Peter G. Simpkin (IKB Technologies Ltd., Bedford, NS B4B 1B4, Canada), and John A. Goff (Univ. of Texas Inst. for Geophys., Austin, TX 78759)

The importance of estimating acoustic wave properties in saturated marine sediments is well known in geophysics and underwater acoustics. As part of the ONR sponsored Geoclutter program, *in situ* acoustic measurements were obtained using *in situ* sound speed and attenuation probe (ISSAP), a device developed and built by the Center for Coastal and Ocean Mapping (CCOM). The location of the Geoclutter field area was the mid–outer continental shelf off New Jersey. Over 30 gigabytes of seawater and surficial sediment data was collected at 99 station locations selected to represent a range of seafloor backscatter types. At each station, the ISSAP device recorded 65 kHz waveform data across five acoustic paths with nominal probe spacing of 20 or 30 cm. The recorded waveforms were processed for compressional wave speed and attenuation. Experimental results are compared to predicted values obtained using the Biot–Stoll theory of acoustic wave propagation. Several methods are examined to estimate the required model parameters. The contribution of loss mechanisms to the effective attenuation is considered. [Research supported by ONR Grant No. N00014-00-1-0821 under the direction of Roy Wilkens and Dawn Lavoie.]

10:45

2aUW11. Modeling of reverberation in the East China Sea. T. W. Yudichak, D. P. Knobles (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029l, yudi@arlut.utexas.edu), P. Cable, Y. Dorfman (BBN Technologies, Mystic, CT 06355-3641), R. Zhang, Z. Peng, F. Li, Z. Li (Chinese Acad. of Sci., Beijing 100080, PROC), P. H. Dahl (Univ. of Washington, Seattle, WA 98105), J. H. Miller, and G. R. Potty (Univ. of Rhode Island, Narragansett, RI 02882)

Reverberant time series recorded in the East China Sea component of the Asian Seas International Acoustics Experiment are analyzed with the aid of a model of acoustic scattering from inhomogeneities in the seabed. Wideband sources deployed by the IOA were used to produce the time series, which were recorded on a thirty-two element VLA also deployed by the IOA. The model employs perturbation theory with the framework of a normal mode approach to evaluate scattering by fluctuations in the sound speed and density in the sediment volume as well as of the roughness of the water–sediment interface. The relative importance of surface and volume scattering at low frequencies in shallow-water environments, such as the East China Sea experimental location, is considered. Also, by incorporating this model in an environmental inversion scheme, information about the distribution of inhomogeneities may be gained. [Work supported by ONR.]

11:00

2aUW12. Effects of bathymetric and oceanographic variations on short range high frequency acoustic propagation in shallow water. Christian de Moustier (Ctr. for Coastal & Ocean Mapping, Univ. of New Hampshire, 24 Colovos Rd., Durham, NH 03824-3525), Seth Mogk, Melissa Hock, and Gerald D'Spain (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0205)

Detailed bathymetry of shoaling regions at two shallow water sites of the San Clemente Offshore Range in Southern California is combined with oceanographic data, taken with CTD, XBT, and bottom moored ADCPs during the winter and summer conditions, to evaluate the effects of environmental variability on acoustic transmission loss and ambient noise measurements made at 16 discrete frequencies ranging from 3 kHz to 30 kHz over 10 km path lengths. Over such short ranges, bottom slopes and tidal effects have the strongest influence on acoustic propagation. [Work supported by the U.S. Naval Oceanographic Office.]

11:15

2aUW13. A comparison of high frequency acoustic transmission data from a smooth water/sand interface with a composite poroelastic model. Marcia Isakson, Daniel Weigl, Erik Bigelow, and Nicholas Chotiros (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029, misakson@arlut.utexas.edu)

Reflection data taken from a smooth water/sand interface have been successfully modeled using a composite poroelastic model. The model was able to account for a decrease in reflectivity at subcritical angles while maintaining physically realistic input parameters. In this study, transmission data have been measured at a number of grazing angles including subcritical angles through a smooth water/sand interface. Because of the flat interface, there should be no contribution to transmission from Bragg scattering or similar scattering effects. The data will be compared with simulations based on a composite poroelastic model using Biot parameters determined from the reflection data. The fit will be considered for its accuracy in predicting the arrivals of the fast, evanescent, and slow waves. [Work supported by ONR, Ocean Acoustics.]

11:30

2aUW14. Separation of scattering mechanisms for the Asian Sea International Acoustics Experiment East China Sea reverberation measurements. Y. Dorfman, P. G. Cable (BBN Technologies), D. P. Knobles, T. W. Yudichak (Appl. Res. Labs., Univ. of Texas, Austin, TX), R. Zhang, Z. Peng, F. Li, Z. Li (Inst. of Ocean Acoust., Beijing, PROC), G. H. Miller, and G. R. Potty (Univ. of Rhode Island, Narragansett, RI)

Low-frequency monostatic reverberation data collected in the East China Sea on a VLA are analyzed to infer bottom scattering strength characteristics. The VLA was deployed and the data collected by researchers from the Institute of Ocean Acoustics, Beijing, China. Reverberation data originating from ranges from 3–15 km are analyzed using coherent array processing methods to determine scattering strengths as a function of frequency and angle at the scatterer. Subaperture processing is used to separate sea surface and bottom contributions and to gain physical insight into the scattering mechanisms responsible for the observed reverberation level. Modeled transmission loss obtained from analyses of measured forward propagation data is employed within the framework of the subaperture signal processing to enable the extraction of the scattering strength. [Work supported by ONR.]

11:45

2aUW15. Sensitivity of broadband propagation and reverberation to oceanographic and bottom variability in shallow water wave guides. Kevin D. LePage (Naval Res. Lab., Washington, DC 20375-5350, kevin.lepage@nrl.navy.mil)

In a previous paper the author derived closed form expressions for the average intensity of broadband time series averaged over oceanographic variability in range independent wave guides [LePage, J. Comput. Acoust. **4** (2001)]. Here, equivalent expressions are derived for the expected value of reverberation intensity. Examples are computed which show that for time series, oceanographic variability most strongly affects the earliest arrivals, while bottom variability most strongly affects the late arrivals. For reverberation, the relative sensitivity to bottom and oceanographic variability were explored using the new model. [Work supported under the ONR capturing uncertainty DRI.]

TUESDAY AFTERNOON, 3 DECEMBER 2002 GRAND CORAL 2, 1:30 TO 3:05 P.M.

Session 2pAAa

Architectural Acoustics and Noise: Music Buildings in Latin America

J. Christopher Jaffe, Chair
Jaffe Holden Acoustics, 114A Washington Street, Norwalk, Connecticut 06854

Chair's Introduction—1:30

Invited Papers

1:35

2pAAa1. The acoustics for speech of eight auditoriums in the city of Sao Paulo. Sylvio R. Bistafa (Dept. of Mech. Eng., Polytechnic School, Univ. of Sao Paulo, Sao Paulo, 05508-900, SP, Brazil, sbistafa@usp.br)

Eight auditoriums with a proscenium type of stage, which usually operate as dramatic theaters in the city of Sao Paulo, were acoustically surveyed in terms of their adequacy to unassisted speech. Reverberation times, early decay times, and speech levels were measured in different positions, together with objective measures of speech intelligibility. The measurements revealed reverberation time values rather uniform throughout the rooms, whereas significant variations were found in the values of the other acoustical measures with position. The early decay time was found to be better correlated with the objective measures of speech intelligibility than the reverberation time. The results from the objective measurements of speech intelligibility revealed that the speech transmission index STI, and its simplified version RaSTI, are strongly correlated with the early-to-late sound ratio C50 (1 kHz). However, it was found that the criterion value of acceptability of the latter is more easily met than the former. The results from these measurements enable to understand how the characteristics of the architectural design determine the acoustical quality for speech. Measurements of ST1-Gade were made as an attempt to validate it as an objective measure of ''support'' for the actor. The preliminary diagnosing results with ray tracing simulations will also be presented.

2:00

2pAAa2. Three halls for music performance in Chile. Jaime Delannoy, Carolina Heuffemann, Daniel Ramirez, and Fernando Galvez (Universidad Perez Rosales, Brown Norte 290 Nunoa, Santiago, Chile)

The primary purpose of this work was to investigate about the present acoustic conditions of used architectonic spaces in Santiago of Chile for orchestras of classic music performance. The studied halls were three: Aula Magna Universidad de Santiago, Teatro Municipal de Nunoa, and Teatro Baquedano. The used methodology was based on studies made by L. Beranek, M. Barron, among others, in concert halls worldwide. As it guides, for the measurement procedure, physical parameters RT, EDT, C50, C80, LF, BR, G, U50 were evaluated according to norm ISO 3382. On the other hand, it has been defined, to proposal way, a questionnaire of subjective valuation directed to musicians, specialized conductors, and listeners.

2:25

2pAAa3. The first vineyard concert hall in North America. Christopher Jaffe and Carlos Rivera (Jaffe Holden Acoustics, Inc., 114A Washington St., Norwalk, CT 06854)

The first vineyard or surround concert hall designed and built in the Western Hemisphere is the Sala Nezahualcoyotl in Mexico City. The Hall was completed in 1976 and is part of the Cultural Center at the Universidad Nacional Autonoma de Mexico. The hall was named after a Toltec poet, architect, and musician who lived in the 15th century and was the Renaissance man of his day. In order to provide the familiar traditional sound of the rectangular (shoebox) European Hall, the acoustic designers set the criteria for reverberation times through the frequency spectrum and the Initial Time Delay Gap at every seat in the house to match the measure-

ments taken at the Grosser Musik vereinssaal in Vienna and Boston Symphony Hall. In this paper we discuss the techniques used to create the traditional sound in a vineyard hall and the reaction of musicians and audiences to the completed facility. The Sala was the model for Suntory Hall in Japan which in turn spawned a number of vineyard halls in Japan. Most recently, the vineyard style seems to be appealing to more and more symphonic organizations in Europe and North America.

Contributed Paper

2:50

2pAAa4. The acoustic design of the Centro Nacional de las Artes in Mexico City. Rusell Cooper (Jaffe Holden Acoustics, Inc., 114A Washington St., Norwalk, CT 06854)

In this paper the acoustic design of the separate buildings housing the school of music, school of drama, and school of dance that opened in 1996 will be described. Spaces that JHA designed included practice rooms, studios, rehearsal rooms, black box, and concert hall. Details of room acoustic treatments, sound isolation measures, and venturi air flow will be illustrated. An overview of the entire project will also include the 500 seat multipurpose theater (with variable absorption systems) and the Alla Magna. Differences between the American and Mexican styles of consulting, importing of materials, installation, and commissioning will also be discussed.

TUESDAY AFTERNOON, 3 DECEMBER 2002 GRAND CORAL 2, 3:15 TO 5:00 P.M.

Session 2pAAb

Architectural Acoustics: Sound Recording and Studios in Mexico

Jose Negrete, Chair

Inst. Politecnico Nacional de Mexico, ESIME-Zacatenco, edif 1, Av. IPN, Mexico 07738, D.F. Mexico

Chair's Introduction—3:15

Invited Paper

3:20

2pAAb1. National Television of Chile—New headquarters building acoustic projects. Mario Huaquin (Natl. Television of Chile, Santiago, Chile, acustmhm@entelchile.net)

In the last 15 years TV stations in Chile have been incorporating in their facilities architectural acoustic and noise control approaches. This has been necessary as much for the technological advance, as for the necessity to achieve a better quality of sound that the listeners receive. In 1998, the National Television of Chile, with the sponsorship of the College of Architects of Chile, requested preliminary architectural designs in order to enlarge and to renovate its headquarters buildings in Santiago, Chile in stages. The Acoustic Project has been developed in an integral way, with three fundamental disciplines: noise and noise control; Machine rooms; vibrations and vibration control; Buildings, engines; architectural acoustics and acoustic comfort; TV studios and technical rooms. This presentation describes the Acoustic Project, phases I (1999), and II (2002), how it was possible to establish a common language with architects and engineers and the different specialties, to apply acoustic criteria and standards, the theoretical development and the projected acoustic solutions. (To be presented in Spanish.)

Contributed Papers

3:45

2pAAb2. Sound perception in the mixing and mastering processes of recorded music. Jose Javier Muedano Meneses (Acoust. Lab., ESIME, IPN, Mexico, jmued@hotmail.com)

This research work presents an analysis of the perceived musical sounds within the mixing process and computer mastering at the audio and video recording studio of the Acoustics Laboratory in ESIME, IPN, Mexico, working with the so-called plug-ins to process the audio signal. The analysis is made from a psychoacoustics standpoint. Samples analyzed include rock, tropical and classical music, and these experiences are part of the Audio course at the Acoustics option of the Electronics and Communications Engineers studies in the Mexican Polytechnic. (To be presented in Spanish.)

4:00

2pAAb3. Acoustic simulations of studio control rooms. Richard A. Moscoso (Seccion Fisica, Pontificia Universidad Catolica del Peru, Apartado 1761, Lima, Peru, rmoscoso@pucp.edu.pe) and Sylvio R. Bistafa (Univ. of Sao Paulo, Sao Paulo, 05508-900, SP, Brazil)

The results of studio-control-room computer simulations with a ray-tracing-type computer program will be presented. Although the validity of ray-tracing simulations of small rooms and at low frequencies may be questionable, the early time echogram at mid and high frequencies obtained by ray tracing provide essential information. Reflections up to 20 ms and frequencies above 500 Hz are responsible for the quality of the listening conditions in studio control rooms, particularly for the assessment of stereophony. Thus, ray-tracing simulations of small rooms for these ranges in the time and frequency domains are justifiable. Computer-

generated impulse responses were obtained for different studio-control-room design philosophies as applied to a basic rectangular room. From a simple scrutiny of the impulse responses the characteristics of the initial time delay region could be verified for each design. Traditional room acoustical parameters such as reverberation time and early decay time were analyzed as well as center time Ts, C5, and C20 for the assessment of the spatial impression, timbre, and transparency. The characteristics of each studio-control-room design philosophy will be discussed in light of these results.

4:15

2pAAb4. Brazilian professional recording studios: Analysis and diagnostics of their acoustical properties. Lineu Passeri, Jr. (Dept. of Technol., Faculty of Architecture and Urbanism, Univ. of Sao Paulo, 05424-970 Sao Paulo, SP, Brazil, nutmusic@uol.com.br) and Sylvio R. Bistafa (Univ. of Sao Paulo, Sao Paulo, SP, Brazil)

Acoustic measurements and design characteristics from a sample of Brazilian professional recording studios located in Sao Paulo and Rio de Janeiro will be presented and discussed. Noise levels, EDT (early decay time), RT20 and RT30 measured in the studios and control rooms will be presented and compared. From these results, the project solutions will be analyzed to check how each recording environment is compatible with its acoustical needs and with the contemporary recording technology. The main goal is to establish guidelines and objective acoustical parameters to be used in the design of new recording studios. (To be presented in Portuguese.)

4:30

2pAAb5. "A" weighting curve algorithm. Jose de Jesus Negrete Redondo, Pablo Roberto Lizana Paulin, Isabel Elena Romero Rizo, and Igmar Moreno Cervantes (Acoust. Lab., ESIME, IPN, Mexico, plizana98@hotmail.com)

The overall project was aimed to build a sound level meter, based on national and international standards. The instrument is being developed on a PC. For this section of the project, the sound signal is captured with a microphone, through an analog/digital converter to the computer. The weighted signal is obtained with a software which includes the equation for the ''A'' weighting curve to perform a convolution with the fast Fourier transfer of the sound signal. (To be presented in Spanish.)

4:45

2pAAb6. Indoor acoustic gain design. Justo Andrés Concha-Abarca (Universidad Tecnológica Vicente Pérez Rosales, Brown Norte 290, Ñuñoa, Santiago de Chile, Chile, justo.concha@upper.cl)

The design of sound reinforcement systems includes many variables and usually some of these variables are discussed. There are criteria to optimize the performance of the sound reinforcement systems under indoor conditions. The equivalent acoustic distance, the necessary acoustic gain, and the potential acoustic gain are parameters which must be adjusted with respect to the loudspeaker array, electric power and directionality of loudspeakers, the room acoustics conditions, the distance and distribution of the audience, and the type of the original sources. The design and installation of front of the house and monitoring systems have individual criteria. This article is about this criteria and it proposes general considerations for the indoor acoustic gain design.

TUESDAY AFTERNOON, 3 DECEMBER 2002 CORAL KINGDOM 2 AND 3, 1:00 TO 5:30 P.M.

Session 2pAB

Animal Bioacoustics: Amphibian Bioacoustics in Honor of Robert Capranica II

Albert S. Feng, Chair
Beckman Institute, University of Illinois, 405 North Mathews Avenue, Urbana, Illinois 61801

Invited Papers

1:00

2pAB1. Calling and plasticity in Pacific Treefrog choruses. Eliot A. Brenowitz (Depts. of Psychol. & Zool., Univ. of Washington, Box 351525, Seattle, WA 98195, eliotb@u.washington.edu) and Gary J. Rose (Univ. of Utah, Salt Lake City, UT 84112)

Male Pacific treefrogs (*Hyla regilla*) use advertisement and encounter calls to regulate intermale spacing within breeding choruses. Encounter calls are produced when a neighbor's advertisement calls exceed a threshold amplitude. These aggressive thresholds are plastic; males resume advertisement calling (i.e., accommodate) to repeated presentation of advertisement calls at amplitudes 4–8 dB above their aggressive threshold. Correspondingly, a male's aggressive threshold for the advertisement call is positively correlated with the maximum amplitude of neighbors advertisement calls measured at his position. The aggressive thresholds of males for the encounter call are also plastic but not correlated with the maximum call amplitude of their neighbors. This disjunction stems from the fact that males in stable choruses are not regularly exposed to suprathreshold encounter calls. Accommodation to one call type fails to significantly alter a male's aggressive threshold to the other call type, which suggests that the two call types are processed by discrete neural filters. Plasticity of aggressive thresholds and aggressive signalling appears to be important in balancing the costs and benefits of aggressive behaviors. In phonotaxic studies, females show a strong preference for the advertisement call over the encounter call. Prolonged encounter calling would therefore decrease a male's chance of mating.

1:25

2pAB2. Environmental influences on anuran sound communication. Mario Penna (Prog. of Physiol. and Biophys., Univ. of Chile, Casilla 70005, Correo 7, Santiago, Chile, mpenna@machi.med.uchile.cl)

Studies of anuran vocal behavior in the South American temperate forest may represent the southernmost influence of Robert Capranicas comparative approach to sound communication. Vocalizations of leptodactylid frogs in this region exhibit patterns of propagation characteristic for different microhabitats. Calls containing frequencies above about 1 kHz experience higher attenuations in bogs as compared to marshes, irrespective of the species native environment. A similar lack of optimal relationships between signal

structure and habitat properties for call propagation occurs in sound environments other than the temperate austral forest, as indicated by joint studies with colleagues from Cuba in Caribbean tropical forests and from Spain in European Mediterranean habitats. Anurans from southern Chile calling in bogs have peculiar acoustic adaptations: males call from burrows inside which the calls of neighboring cospecifics are amplified considerably. This effect potentially counteracts the constraints for signal propagation of sound-attenuating environments. The evidence presented indicates that anuran calls have not been subjected to environmental selection pressures, to the extent vocalizations of other vertebrates have. Dispersion of anurans is limited by availability of water sources, but males can choose locations favoring signal broadcast and reception in their relatively restricted breeding areas. [Work supported by FONDECYT Grant No. 1010569.]

1:50

2pAB3. Neural correlates of temporal pattern selectivity in anurans. H. Carl Gerhardt (Univ. of Missouri, Columbia, 215 Tucker Hall, Columbia, MO 65211)

Female anurans typically show strong phonotactic selectivity for synthetic signals having pulse-repetition (pulse) rates that are close to those in advertisement calls produced by conspecific males. Auditory neurons showing temporal selectivity for species-typical pulse rates (especially bandpass neurons) are commonly reported from the inferior colliculus and central nucleus of the auditory thalamus. However, a number of questions remain about the neural bases of temporal pattern selectivity. First, the proportion of neurons tuned to conspecific values varies enormously among species. Are these differences real or by-products of different methodologies? Second, behavioral studies indicate that some species do not evaluate the pulse rate *per se* but rather are selective to certain ranges of pulse duration and interpulse intervals. What is the role of bandpass (and other filter types) neurons in these species? Third, temporally selective neurons are often tuned to spectral peaks that are not emphasized in the advertisement call. Do these cells nevertheless play a role in temporal selectivity? Finally, recent studies indicate that extensive lesions of the auditory thalamus have little effect on phonotactic selectivity for conspecific pulse rates. What then is the role of temporally tuned neurons in this structure? [Work supported by NSF and NIH.]

2:15

2pAB4. Temporal call characters and vocal communication in the cricket frog (*Acris crepitans*). Walter Wilczynski (Dept. of Psych., Univ. of Texas, Austin, TX 78712) and Michael J. Ryan (Univ. of Texas, Austin, TX 78712)

During intermale vocal interactions, male cricket frogs change the temporal, frequency, and amplitude characteristics of their calls. In playback studies, males changed the temporal structure of their calls in the direction of higher aggression to a greater degree when presented with calls having a high aggression temporal structure compared to a low aggression structure (with frequency and amplitude held constant). Hearing a low aggression call from the same location prior to a high aggression call resulted in a smaller change in the male's call. Temporal structure degrades progressively between 4 and 16 m from the source over natural substrates. As distance-induced temporal degradation increases, males change temporal call characters progressively less. In sum, male cricket frogs respond to the temporal characteristics of other male calls by changing the temporal parameters of their own calls in a graded fashion. The magnitude of the change depends on the parameters of the challengers call, prior experience with calls from the challengers position, and temporal degradation (hence distance) of the challengers' call. Phonotaxis experiments show that female cricket frogs prefer the high aggression calls. The call changes therefore seem to represent males increasing call attractiveness in response to increased competition from neighboring males.

2:40

2pAB5. Dynamic bimodal signal evokes fighting behavior in a dart-poison frog. Peter M. Narins (Dept. of Physiological Sci., UCLA, Los Angeles, CA 90095-1606), Walter Hoedl, and Daniela S. Grabul (Univ. of Vienna, A 1090, Vienna, Austria)

As a neuroethologist, Bob Capranica strongly encouraged his students to study animals in their natural habitat in order to truly understand the behavioral rules underlying acoustic communication. This study was one of many inspired by those lessons. Male anuran vocalizations serve to attract conspecific females and to regulate male spacing through territorial interactions. In response to playback of a conspecific call, some male frogs have been reported to alternate their call with the perceived stimulus or shift their call-dominant frequency to avoid acoustic interference, or add call notes to signal an increased state of aggression. In some species, males orient toward the sound source followed by physically approaching the loudspeaker. Although natural fighting behavior between conspecific males has been observed in the field, it has not heretofore been possible to elicit with loudspeakers broadcasting sounds. In this playback study of the Brilliant-thighed dart-poison frog (*Epipedobates femoralis*) in French Guiana, we used an electromechanical model to provide realistic bimodal cues (acoustic and visual) to calling males. Our data suggest that agonistic behavior could be evoked in territorial males only when both acoustic and visual cues were presented simultaneously. [Work supported by grants from NIH (PMN) and Austrian Science Foundation (WH).]

3:05–3:30 Break

3:30

2pAB6. The role of call frequency modulation and the auditory papillae in phonotactic behavior in a dart-poison frog. Walter Hoedl (Inst. of Zoology, Univ. of Vienna, Althanstrasse 14, A-1090 Wien, Austria, walter.hoedl@univie.ac.at), Adolfo Amezquita (Universidade de Los Andes, Bogota, Colombia), and Peter M. Narins (Univ. of California, Los Angeles, Los Angeles, CA 90095)

Territorial males of the Brilliant-thighed dart-poison frog, *Epipedobates femoralis*, are known to present stereotypic phonotactic responses to the playback of conspecific and synthetic calls. Fixed site attachment and a long calling period within an environment of little microclimatic changes render this terrestrial and diurnal pan-Amazonian species a rewarding subject for frog bioacoustics. In field experiments at Aratai, French Guiana, we tested whether the prominent frequency modulation of the advertisement call notes is critical for eliciting phonotactic response. Substitution of the natural upward sweep by either a pure tone within the species frequency range or a reverse sweep did not alter the males' phonotactic behavior. Playbacks with advertisement calls embedded in high levels of either low-pass or high-pass masking noise designed to saturate either the amphibian (AP) or basilar papilla (BP) showed that male phonotactic behavior in this species is subserved by activation of the BP rather than the AP of the inner ear. [Work supported by grants from Austrian Science Foundation P 11565, P 15345 (WH), and NIH (PMN).]

3:45

2pAB7. Auditory grouping in the túngara frog: The roles of complex call components in what and where decisions. Hamilton Farris (Integrative Biol., C0930, Univ. of Texas, Austin, TX 78712), A. Stanley Rand (Smithsonian Tropical Res. Inst., Balboa, Panama), and Michael J. Ryan (Univ. of Texas, Austin, TX 78712)

Like humans, numerous animals across disparate taxa must identify and locate complex acoustic signals imbedded in multiple overlapping signals and ambient noise. A requirement of this task is the ability to group sounds into auditory streams in which sounds are perceived as emanating from the same source. Although comparatively few assays have demonstrated aspects of auditory grouping in nonhuman animals, Capranica and colleagues have revealed several excellent examples through their work with anuran bioacoustics. In this study, we build on their work by presenting evidence for auditory grouping in túngara frogs (*Physalaemus pustulosus*). The complex calls of *P. pustulosus* consist of two discrete components, which are commonly produced in multi-male choruses. By measuring the phonotactic responses of females to spatially segregated components, we show that, in contrast to humans, spatial cues play a limited role in grouping, as grouping occurs over wide angular separations. In addition, the presentation of spatially segregated call components allowed us to measure the behavioral significance of each component in the complex. We show that once grouped the separate call components are weighted differently in recognizing and locating the call, so-called ''what'' and ''where'' decisions, respectively.

4:00

2pAB8. Nonparallel coevolution of sender and receiver in the acoustic communication system of treefrogs. Johannes Schul and Sarah L. Bush (Biological Sci., Univ. of Missouri, 207 Tucker Hall, Columbia, MO 65211)

Advertisement calls of closely related species often differ in quantitative features such as the repetition rate of signal units. These differences are important in species recognition. Current models of signal/receiver co-evolution predict two possible patterns in the evolution of the mechanism used by receivers to recognize the call. (1) Classical sexual selection models (Fisher-Process, good-genes/indirect benefits, direct benefits models) predict that close relatives use qualitatively similar signal recognition mechanisms tuned to different values of a call parameter. (2) Receiver bias models (hidden preference, pre-existing bias models) predict that if different signal recognition mechanisms are used by sibling species, evidence of an ancestral mechanism will persist in the derived species, and evidence of a pre-existing bias will be detectable in the ancestral species. We describe qualitatively different call recognition mechanisms in sibling species of treefrogs. Whereas Hyla chrysoscelis uses pulse rate to recognize male calls, H. versicolor uses absolute measurements of pulse duration and interval duration. We found no evidence of either hidden preferences or pre-existing biases. The results are compared with similar data from katydids (Tettigonia sp.). The data are discussed with regard to current models of signal/receiver co-evolution.

4:15

2pAB9. Forty years of solitude: Using laboratory lines of *D. melanogaster* to study behavioral isolation. Christine R. Boake (Dept. of Ecol. & Evol. Biol., Univ. of Tennessee, Knoxville, TN 37996, cboake@utk.edu)

Population genetic models of speciation show that the initial population divergence can be very rapid when sexual signals are involved, and that speciation through sexual selection has a high probability compared to speciation through viability selection. The models also suggest that the exact nature of the changes in a signal system can be arbitrary. This raises empirical questions of whether behavioral divergence can be detected sooner than postzygotic isolation, and whether, in a multimodal signal system, certain signals are more likely to diverge than others. The early stages of behavioral isolation are being investigated by using stocks of D. melanogaster that have been in captivity for 40–50 years. Two lines that were part of a 1950s study of DDT resistance have begun to evolve behavioral isolation; however, postzygotic isolation is not detectable. The courtship signals of recently diverged populations can be compared to published reports of behavioral isolation between populations of D. melanogaster and between D. melanogaster and its close relative to learn whether signal divergence always follows the same trajectory.

4:30

2pAB10. How long do females listen? Assessment time for female choice in the gray treefrog, *Hyla versicolor*. Joshua J. Schwartz (Dept. of Biological Sci., Pace Univ., 861 Bedford Rd., Pleasantville, NY 10570)

A satisfactory understanding of the process of mate choice by female frogs requires that we know how sensitive females are to the variation in males' calls under natural conditions and what is the time scale or ''window'' over which females compare males. In natural choruses, gray treefrog females may sit near calling males for many minutes before approaching a particular individual to mate, while in laboratory-based tests they may approach a speaker following less than 30 s of exposure to broadcast calls. Females prefer long to short calls. In order to estimate ''assessment time'' of females in nature, calls were broadcast from four pairs of 360-deg speakers surrounded by screen cages at a pond in Missouri. One speaker per pair presented calls of constant duration, while the other speaker shifted between calls longer or shorter than the constant duration call. The period over which this change in call duration occurred differed for each of the four pairs of speakers. The numbers of females captured at the speaker array over the breeding season indicated that the most likely assessment time was close to 2 min. This estimate is similar, but not identical, to that obtained from additional laboratory tests.

4:45

2pAB11. Phonotaxis and chorus organization in African frogs. Phillip J. Bishop, Robert R. Capranica, and Neville I. Passmore (Dept. of Zoology, Univ. of the Witwatersrand, Johannesburg, S. Africa)

Detailed studies of the phonotactic behavior and chorus organization of several species of African amphibians were conducted from 1982–1992. A range of phonotactic experiments conducted in two- and three-dimensional arenas, using between one and four loudspeakers and a variety of different stimuli were carried out on the painted reed frog

(*Hyperolius marmoratus*). These experiments revealed the remarkable ability of very small anurans (<30 mm, SVL) to accurately localize a sound source, in both the horizontal and vertical planes. Furthermore, females were able to discriminate between two sound sources that differed in intensity by only 5 dB and their ability was significantly impaired by the introduction of further sound sources. The confounding effect of multiple sound sources on female choice may explain the presence of nonrandom mating with respect to size in this species. The chorus organization of five species of African anurans was investigated by using playback experiments. These experiments revealed four distinct categories of chorus organization which can be plotted on a continuum, from random calling to very precise triggered responses. The type of chorus organization was found to be directly related to the length of the call and the delay in response to the playback stimulus.

5:00–5:30

Questions and Comments

TUESDAY AFTERNOON, 3 DECEMBER 2002 CORAL GARDEN 1, 1:00 TO 5:45 P.M.

Session 2pAO

Acoustical Oceanography and Animal Bioacoustics: Using Ambient Sound in the Ocean

Jeffrey A. Nystuen, Chair
Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105

Chair's Introduction—1:00

Invited Papers

1:05

2pAO1. NOAA efforts in monitoring of low-frequency sound in the global ocean. Christopher G. Fox (NOAA/PMEL, 2115 S.E. OSU Dr., Newport, OR 97365, fox@pmel.noaa.gov), Robert P. Dziak, and Haruyoshi Matsumoto (Oregon State Univ., Newport, OR)

Since August 1991, NOAA/PMEL has collected continuous recordings from the U.S. Navy SOSUS arrays in the North Pacific. In May 1996, this effort was expanded through the use of PMEL-developed autonomous hydrophones deployed in the eastern equatorial Pacific, and later to the central North Atlantic between 15N and 35N (March 1999), the Gulf of Alaska (October 1999), and the North Atlantic between 40N and 50N (June 2002). Natural seismicity in the Pacific produces nearly 10 000 events per year with source levels exceeding 200 dB (*re*: 1 micro-Pa @ 1 m), with about 3500 events per year exceeding this level in the North Atlantic. Significant contributions from manmade sources are present throughout the data but have not been quantified. Recordings from North Atlantic arrays are dominated by noise from seismic airgun profilers working offshore Canada, Brazil, and west Africa. In September 2001, a cabled vertical hydrophone array was installed at Pioneer Seamount, offshore central California, which will provide continuous, unclassified acoustic data (in the range of 1–450 Hz) to the research community in real time. Future plans call for the expansion of the NOAA monitoring effort to other opportunities worldwide and making the raw data available to the community via the Internet.

1:25

2pAO2. Eight-year records of low-frequency ambient sound in the North Pacific. Rex K. Andrew, Charlotte Leigh, Bruce M. Howe, and James A. Mercer (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

Spectra of omnidirectional ambient sound have been collected since 1994 at 13 locations around the North Pacific. Data were acquired for 3 minutes every 6 minutes and spectra calculated from 0–500 Hz in 1 Hz bands. With a million spectra per site, this database allows investigation into the statistical character of low-frequency ambient sound at multiple scales. At the shortest scales, the spectral levels in the shipping bands have a fluctuation spectrum similar to a $1/f$ process, with decorrelation times less than 20 minutes. At intermediate scales, the seasonal baleen whale component becomes the most dominant and repeatable feature. At the longest scales (averaging over the entire record) the ambient levels (at the Pt. Sur site) seem to have increased by up to 10 dB since the 1960s. The distribution of the levels (in decibels) generally indicates a short tail for quieter levels but a long tail for loud events. The Pt. Sur data set has also been used to validate the new dynamic ambient noise model (DANM), which shows good agreement in one-third octave bands to within a couple of decibels for January 1998. These and further results will be discussed. [Work supported by ONR and SPAWAR.]

1:45

2pAO3. Creating a web-based library of underwater biological sound. Jack W. Bradbury and Carol A. Bloomgarden (Macaulay Library, Cornell Lab. of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850)

Establishing an archive of fish and other underwater biological sounds will meet many of the long-standing challenges faced by marine acousticians—the restoration and preservation of deteriorating recordings, the ability to catalog their sounds and data in a way that fosters the exchange and sharing of data, easy access to the sounds for analysis and identification, and the capacity to search through passive recordings for sounds of particular interest. The Macaulay Library of Natural Sounds, with a long history of working toward these goals in ornithology and animal behavior, recently launched into the realm of underwater sounds with the help of over 80 individual recordists and institutions worldwide. Researchers will be able to annotate their sounds through an online database

application, summarize search results in exportable tables and maps, and download copies of recordings for research, teaching, and conservation. MLNS is committed to dual goals of maintaining open access to allow other researchers to listen and help identify sounds, while protecting recordists copyrights and restricting access during the publication process. Detailed and extensive metadata are needed, however, to create the functionality such an archive requires.

2:05

2pAO4. Eavesdropping on marine mammals to monitor migration. Kathleen Stafford (Natl. Marine Mammal Lab., 7600 Sand Point Way NE, Seattle, WA 98115)

Many baleen whales produce low-frequency sounds that are sufficiently loud that they can be detected over long distances. Collectively these sounds contribute significantly to ambient noise levels in the ocean during certain times of the year [Curtis *et al.*, J. Acoust. Soc. Am. **106**, 3189–3200 (2000)]. By monitoring the occurrence of these sounds at different locations in the ocean, geographic and seasonal patterns begin to emerge that may be used to document the distribution and basin-scale movements of baleen whales. Each baleen whale species make distinctly different calls so distinguishing among them, and sometimes among populations within species, is possible. Although for many species the ecological function of sound production remains poorly understood, acoustic data can nevertheless be used to examine the seasonal occurrence and migratory behavior of large whales. One example of the utility of monitoring migrations by use of passive acoustics can be seen in the northeastern Pacific where sounds from blue whales are detected in the eastern tropical Pacific from November to May, off the west coast of the United States from July to January and in the Gulf of Alaska from September to December.

2:25

2pAO5. Toward a fisheries bioacoustics. David A. Mann (Univ. of South Florida, College of Marine Sci., 140 7th Ave. S., St. Petersburg, FL 33701, dmann@marine.usf.edu)

Fisheries bioacoustics is emerging as an application of passive acoustic detection of fish sound production. Many fishes produce sounds during reproductive activities that can be used to determine where and when they are spawning. Autonomous dataloggers and hard-wired hydrophones were used to record sound production by fishes in estuaries of southwest Florida. These data show that passive acoustics can be used to locate spawning sites and determine the timing of spawning by commercially important species. Ultimately fisheries bioacoustics should move the way of fisheries acoustics where the signal output is not the actual sound data, but the locations and intensity of fish spawning. A useful analogy is the development of SONAR systems for fish quantification. These systems do not deliver raw sound data to the researcher. They return processed data on fish location and abundance. One can envision the day when real-time fisheries bioacoustics systems will produce maps of the locations of sound-producing fishes that can provide managers with data on the temporal and spatial extent of spawning.

2:45

2pAO6. Expanding uses of ambient noise for imaging, detection, and communication. John R. Potter and Laurent Malod (ARL, TMSI, NUS, 12a Kent Ridge Rd., Singapore 119223)

The use of ambient noise to sense the marine environment has a human history of only two decades. Starting with incoherent processing inspired by optical analogies such as Acoustic Daylight (TM), the exploration of the potential for ambient noise to provide useful information about the environment has blossomed into many related areas and diverse algorithms with connections to multi-static active sonar, classic passive sonar, communications, matched field processing and others. This presentation will introduce some recent work in these areas and attempt to draw together how the use of ambient noise, both by mankind and marine animals, is beginning to form a more complete picture of the potential of this exciting area of research. [Work supported by the Defence Science and Technology Agency, Singapore.]

3:05

2pAO7. Estimating shallow water bottom geoacoustic parameters from ambient noise. Dajun Tang (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Knowing bottom geoacoustic parameters is of great importance for using sonar systems effectively in shallow waters. In this paper, ambient noise data recorded on a vertical hydrophone array taken in the frequency range of 1000 to 3000 Hz were used. Forward modeling and model/data comparison show that the energy ratio of down-looking and up-looking beams, after a proper average over time and frequency, is the energy reflection coefficient of the bottom. From the reflection coefficient, critical parameters of the sediments, the sound speed, density, and attenuation coefficient, are obtained. Core data taken at the experimental site support the inversion results.

3:25–3:40 Break

3:40

2pAO8. Monitoring air entrainment with breaking wave noise. Grant B. Deane and Dale M. Stokes (Code 0238, Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238)

It is now known that the dominant component of wind-driven oceanic noise comes from breaking waves. Bubbles ranging in size from tens of microns to centimeters are forced into the water column during the first second or so of whitecap formation. At the moment of creation, each bubble emits a pulse of sound at a center frequency inversely proportional to its radius. The ensemble of such events amounts to a burst of noise that continues throughout the active phase of bubble creation within the whitecap. As the noise spectrum is related to the bubble size distribution within the whitecap, it is natural to explore the possibility that underwater oceanic

ambient noise could be used as tool to remotely monitor air entrainment rates across the ocean surface. One of the problems with developing such a tool is understanding the relationship between bubble formation processes occurring within whitecaps and the concurrent noise emission. Here we will report recent progress in our understanding of bubble formation mechanisms in whitecaps, their role in ambient noise generation, and the implications for monitoring air entrainment rates.

Contributed Papers

4:00

2pAO9. Passive acoustic detection and measurement of rainfall at sea. Jeffrey A. Nystuen and Barry Ma (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, nystuen@apl.washington.edu)

It is well recognized that rainfall measurements are needed over the world's oceans. One method of providing these measurements is to passively listen for the underwater sound signal that is produced by rainfall striking the ocean surface. Since 1998, over 70 buoy-months of ambient sound data have been collected using Acoustic Rain Gauges (ARGs) deployed on deep ocean moorings that form part of the Tropical Atmosphere Ocean (TAO) array in the Pacific Ocean. These data demonstrate the acoustic measurement of oceanic wind and rain. Other ''noises'' are present in the ocean and need to be detected and rejected. This is accomplished by recognizing the unique spectral and temporal character of rain-generated sound. A quantitative relationship between absolute sound levels and rainfall rate is proposed. The probability of acoustic detection of rainfall events under different weather conditions will be discussed. A quantitative comparison of rainfall accumulation using the acoustic technique with co-located rainfall estimates from on-board R.M. Young rain gauges and from NASA TRMM satellite overpasses (Rainfall Product 3B42) shows promising agreement, but also points out problems associated with each measurement method. [Work supported by ONR—Ocean Acoustics, NSF—Physical Oceanography, and NOAA Office of Global Programs.]

4:15

2pAO10. Is anthropogenic ambient noise in the ocean increasing? Elena McCarthy (Dept. of Marine Affairs, Univ. of Rhode Island, Washburn Hall, Kingston, RI 02881, emcc7035@postoffice.uri.edu) and James H. Miller (Univ. of Rhode Island, Kingston, RI 02881)

It is commonly accepted that the ocean's ambient noise levels are rising due to increased human activities in coastal and offshore areas. It has been estimated that low-frequency noise levels increased more than 10 dB in many parts of the world between 1950 and 1975. [Ross, Acoustics Bulletin, Jan/Feb (1993)]. Several other sources cite an increase in man-made, or anthropogenic, noise over the past few decades. [D. A. Croll *et al.*, Animal Conservation **4**(1) (2001); Marine Mammal Commission Report to Congress (1999); C. W. Turl, NOSC Tech. Report 776 (1982)]. However, there are few historical records of ambient noise data to substantiate these claims. This paper examines several sectors of anthropogenic activities to determine their contributions to ambient noise. These activities include shipping, oil and gas exploration, military sonar development, and academic research. A series of indices for each of these industries is developed to predict ambient noise trends in the sea. It is found that the amount of noise generated by individual activities may have decreased overall due to new technologies and improved efficiency even if the intensity of such activities has increased.

4:30

2pAO11. Geoacoustic inversion of noise coherence in shallow water. David J. Thomson, Francine Desharnais (Defence Res. and Development Canada Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada), Matthew L. Drover (Dalhousie Univ., Halifax, NS B3H 4J1, Canada), and Chris A. Gillard (Defence Sci. and Technol. Organisation, Edinburgh, South Australia 5111, Australia)

It is known that the geoacoustic properties of a shallow-water sea-bed can be inferred from relatively simple measurements of the ambient noise coherence between a pair of vertically separated hydrophones [D. M. F. Chapman, ''Surface-generated noise in shallow water: A model,'' Proc. Inst. Acoust. **9**, 1–11 (1987)]. The design of an autonomous buoy package for acquiring geoacoustic information by this method is currently being considered by DRDC-Atlantic in support of matched-field localization efforts that are being developed for use with rapidly deployable arrays. Initially, vertical coherence estimates from a simple shallow water noise model were fit to measured coherences by adjusting geoacoustic parameters by a trial and error procedure. A more systematic approach involves combining noise coherence models with nonlinear global optimization methods based on matched-coherence processing concepts to search the space of possible sea-bed parameters more efficiently. In this paper, we report on recent efforts to use a hybrid simplex simulated annealing scheme [M. R. Fallat and S. E. Dosso, ''Geoacoustic inversion via local, global, and hybrid algorithms,'' J. Acoust. Soc. Am. **105**, 3219–3230 (1999)] to match an increasingly realistic suite of candidate geoacoustic parametrizations to some acoustic noise coherent data measured with modified sonobuoys deployed at several shallow water locations on the Scotian Shelf.

4:45

2pAO12. Model-based tracking of marine mammals near California using seismometers. Christopher O. Tiemann, Michael B. Porter (Sci. Appl. Intl. Corp., La Jolla, CA 92037), and John Hildebrand (Univ. of California, San Diego, La Jolla, CA 92037)

An algorithm originally developed for tracking humpback whales around a deep-water hydrophone array near Hawaii has been proven capable of localizing another species of marine mammal in the shallow waters off California. The new sources of interest are blue whales, animals with markedly different call characteristics from humpback whales, and the data under examination is from four bottom-mounted seismometers in a 3 km square array. The algorithm uses a range-dependent acoustic model to predict time differences of arrival (time-lag) of blue whale calls as measured between sensor pairs, while real pairwise time-lags are measured through a phase-only correlation process. Comparison between modeled and measured time-lags forms an ambiguity surface identifying the most probable whale location in a horizontal plane around the sparse array. The robustness of the model-based localization technique is illustrated in its application to a different scenario than for which it was developed, and it is also suitable for continuous, real-time, unattended alert and tracking applications. Examples of tracking whales along their migratory path will be provided.

5:00

2pAO13. Breaking waves and ambient sound in the ocean. W. Kendall Melville (Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0213, kmelville@ucsd.edu)

It is now well accepted that breaking waves at the ocean surface are the primary source of sea-surface sound. However, the relationship between the kinematics and dynamics of breaking waves, their acoustic source strength and their statistical description has not been explored in any comprehensive way. In this paper, the components of such a description are reviewed and discussed in the context of laboratory and field studies from the literature, theoretical models of breaking statistics, and recent field measurements of the kinematics and statistics of wave breaking. The implications of the results for the use of ambient sound to quantify processes of air–sea interaction are discussed. [Work supported by ONR and NSF.]

5:15

2pAO14. Hurricane classification using a full-field ocean surface noise model. Joshua D. Wilson and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Hurricanes generate noise in the ocean due to wind–wave interaction. The authors have previously discussed the possibility of using this noise to determine the size and strength of the hurricane with a modal model [Wilson and Makris, J. Acoust. Soc. Am. **108** (2000)]. Here the analysis is extended to include a full-field model for surface-generated ocean noise. Unlike previous surface noise models that contained far-field approximations, this full-field model can be used to calculate the acoustic field both inside and outside the hurricane. This full-field model is used to calculate the spatial covariance of the acoustic field generated by a hurricane. This spatial covariance is then used to determine the sound from a hurricane that would be detected by hydrophones and hydrophone arrays. Several examples are presented using single sensors and sensor arrays inside and outside the hurricane to determine the best method for classifying a hurricane. In addition shallow- and deep-water environments are compared to illustrate their effect on the propagation of surface-generated hurricane noise. Also simulations are shown for multiple frequencies to show the filtering effect of the waveguide on the propagation and to determine the optimal frequency for hurricane classification.

5:30

2pAO15. Laboratory measurements of noise generation by shoaling breakers. Steven L. Means and Paul J. Gendron (Naval Res. Lab., Code 7120, 4555 Overlook Ave. SW, Washington, DC 20375)

Simultaneous measurements of the surface gravity wave field, void fraction entrained during breaking, and the generated acoustic spectra were made in a sand beached wave tank at the Center for Applied Coastal Research at the University of Delaware during April 2002. The tank has the dimensions of 30 m (l)×2 m (w)×1.5 m (d) and a 0.2-mm sand beach. Four hydrophones were distributed along the sand beach beneath the region of active breaking. Wave gauges measured surface gravity wave properties along the beach cross-shore. Conductivity and temperature probes allowed for the measurement of the void fraction of the entrained bubble cloud. A video camera captured the evolution of the entrained bubble cloud through the tank's Plexiglas™ side panel. The paper presents generated acoustic power levels as a function of surface wave amplitude and period. Initial results in obtaining the relationships between void fractions of the entrained bubble clouds and the spectral components of the generated acoustic signal will also be discussed. [Work supported by ONR base funding at NRL.]

TUESDAY AFTERNOON, 3 DECEMBER 2002 CORAL KINGDOM 1, 1:00 TO 1:50 P.M.

Session 2pEA

Engineering Acoustics: Honoring Per Brüel's Contributions

Leo L. Beranek, Chair
975 Memorial Drive, Suite 804, Cambridge, Massachusetts 02138-7555

Chair's Introduction—1:00

Invited Paper

1:05

2pEA1. Working with Dr. Per V. Brüel. Svend Gade (Brüel & Kjær Univ., Brüel & Kjær, Sound & Vib. Measurement A/S, Skodsborgvej 307, DK-2850 Naerum, Denmark)

For more than a decade I have had the pleasure to work as an application specialist together with—and for—Dr. Brüel, one of the founders of the Brüel & Kjær Company, famous for sound and vibration measurement instrumentation, often nicknamed ''Green Boxes.'' It has been a great experience for me, and I recall this period in my life as one where I was much inspired by Dr. Brüel's methods, both as a private person and with his work as a director for the company and leader of both the sales and the innovation departments. In this presentation I will highlight some funny stories that are told about Dr. Brüel combined with the episodes that I have experienced myself. In short, the most simple way to characterize this rather complex person is maybe by repeating his vision statement for the company: ''We shall have fun and we shall make money. On the other hand we shall not have so much fun that we do not make any money, and we shall not make so much money that we do not have any fun!'' For Per Brüel, acoustics is one of his great hobbies. He has others such as cars, airplanes, motorbikes (he is the lucky owner of a Danish Nimbus) and wine.

Contributed Paper

1:35

2pEA2. Upcoming new international measurement standards in the field of building acoustics. Hans Goydke (Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, D-38116 Braunschweig, Germany, hans.goydke@ptb.de)

The extensively completed revision of most of the ISO measurement standards in building acoustics mainly initiated by the European Commissions demand for harmonized standards emphasized the insight that the main goal to avoid trade barriers between the countries can only be reached when the standards sufficiently and comprehensively cover the field when they are related to the actual state of the art and when they are sufficiently related to practice. In modern architecture one can observe the rapid change in the use of building materials, for instance regarding the use of glass. Lightweight constructions as well as heavyweight building elements with additional linings are increasingly in common use and unquestionably there are consequences to be considered regarding the ascertainment of sound insulation properties. Besides others, International Standardization is unsatisfactory regarding the assessment of noise in buildings from waste water installations, in the low frequency area and in general regarding the expression of uncertainty of measurements. Intensity measurements in building acoustics, rainfall noise assessment, estimation of sound insulation, impulse response measurement methods, assessment of sound scattering are examples of upcoming standards.

Session 2pMUa

Musical Acoustics: General Topics in Musical Acoustics

James P. Cottingham, Chair
Physics Department, Coe College, Cedar Rapids, Iowa 52402

Contributed Papers

1:00

2pMUa1. Observation of the laryngeal movements for throat singing. Ken-Ichi Sakakibara (NTT Commun. Sci. Labs., 3-1, Morinosato Wakamiya, Atsugi-shi, Kanagawa 243-0198, Japan), Tomoko Konishi, Emi Z. Murano, Hiroshi Imagawa (The Univ. of Tokyo, Tokyo, Japan), Masanobu Kumada (Natl. Rehabilitation Ctr. for the Disabled, Saitama, Japan), Kazumasa Kondo (Saitama, Japan), and Seiji Niimi (Intl. Univ. of Health and Welfare, Tochigi, Japan)

Throat singing is a traditional singing style of people who live around the Altai Mountains. Khöömei in Tyva and Khöömij in Mongolia are representative styles of throat singing. The laryngeal voices of throat singing is classified into (i) a *drone voice* which is the basic laryngeal voice in throat singing and used as drone and (ii) a *kargyraa voice* which is very low pitched with the range outside the modal register. In throat singing, the special features of the laryngeal movements are observed by using simultaneous recording of high-speed digital images, EGG, and sound wave forms. In the drone voice, the ventricular folds (VTFs) vibrate in the same frequency as the vocal folds (VFs) but in opposite phases. In the kargyraa voice, the VTFs can be assumed to close once for every two periods of closure of the VFs, and this closing blocks airflow and contributes to the generation of the subharmonic tone of kargyraa. Results show that in throat singing the VTFs vibrate and contribute to producing the laryngeal voice, which generates the special timbre and whistle-like overtone.

1:15

2pMUa2. A human vocal utterance corpus for perceptual and acoustic analysis of speech, singing, and intermediate vocalizations. David Gerhard (Dept. of Computing Sci., Simon Fraser Univ., 8888 University Dr., Burnaby, BC V5A 1S6, Canada)

In this paper we present the collection and annotation process of a corpus of human utterance vocalizations used for speech and song research. The corpus was collected to fill a void in current research tools, since no corpus currently exists which is useful for the classification of intermediate utterances between speech and monophonic singing. Much work has been done in the domain of speech versus music discrimination, and several corpora exist which can be used for this research. A specific example is the work done by Eric Scheirer and Malcom Slaney [IEEE ICASSP, 1997, pp. 1331–1334]. The collection of the corpus is described including questionnaire design and intended and actual response characteristics, as well as the collection and annotation of pre-existing samples. The annotation of the corpus consisted of a survey tool for a subset of the corpus samples, including ratings of the clips based on a speech–song continuum, and questions on the perceptual qualities of speech and song, both generally and corresponding to particular clips in the corpus.

1:30

2pMUa3. Computer-animated illustrations of vibrations and waves. Donald E. Hall (Phys. Dept., California State Univ., 6000 J St., Sacramento, CA 95819, hallde@dgweb.com)

Under this same title, Bruce Richards presented at the 143rd ASA meeting [J. Acoust. Soc. Am. **111**, 2394 (2002)] an admirable set of class demonstrations implemented on a Macintosh computer with C++ and the Mac Toolbox. This has inspired the preparation of a similar package of animations written in BASIC for the PC. They include visualizations of motion and corresponding graphs for plucked and bowed strings, bars and membranes, standing and traveling waves in pipes, and normal modes of cylindrical and conical pipes. The animations correspond closely to illustrations in standard textbooks such as the author's [*Musical Acoustics*, Brooks–Cole, 3rd ed., 2002].

1:45

2pMUa4. Pitch jnd and the tritone paradox: The linguistic nexus. Kourosh Safari (Music Cognition and Acoust. Lab., Univ. of California, Los Angeles, Box 951657, Los Angeles, CA 90095-1657)

Previous research has shown a connection between absolute pitch (the ability to name a specific pitch in the absence of any reference) and native competence in a tone language (Deutsch, 1990). In tone languages, tone is one of the features which determines the lexical meaning of a word. This study investigates the relationship between native competence in a tone language and the just noticeable difference of pitch. Furthermore, the tritone paradox studies have shown that subjects hear two tritones (with bell-shaped spectral envelopes) as either ascending or descending depending on their linguistic backgrounds (Deutsch, 1987). It is hypothesized that the native speakers of tone languages have a higher JND for pitch, and hear the two tones of the tritone paradox as ascending, whereas, native speakers of nontone languages hear them as descending. This study will indicate the importance of early musical training for the development of acute tone sensitivity. It will also underline the importance of language and culture in the way it shapes our musical understanding. The significance of this study will be in the areas of music education and pedagogy.

2:00

2pMUa5. The effects of timbre on melody recognition are mediated by familiarity. J. Devin McAuley and Chris Ayala (Dept. of Psych., Bowling Green State Univ., Bowling Green, OH 43403)

Two experiments examined the role of timbre in music recognition. In both experiments, participants rated the familiarity of a set of novel and well-known musical excerpts during a study phase and then were given a surprise old/new recognition test after a retention interval. The recognition test was comprised of the target melodies and an equal number of distractors; participants were instructed to respond yes to the targets and no to the distractors. In experiment 1, the timbre of the melodies was held constant throughout the study and then either stayed the same or switched to a different instrument sound during the test. In experiment 2, timbre varied randomly from trial to trial between the same two instruments used in experiment 1, yielding target melodies that were either mismatched or matched in their timbre. Switching timbre between study and test in experiment 1 was found to hurt the recognition of the novel melodies, but

not the familiar melodies. The mediating effect of familiarity was eliminated in experiment 2 when timbre varied randomly from trial to trial rather than remaining constant. Possible reasons for the difference between studies will be discussed.

2:15

2pMUa6. Recreating the real, realizing the imaginary—a composer's preoccupation with acoustic space. Rob Godman (Univ. of Coventry and Univ. of Hertfordshire, c/o 4 Mill Close Wotton-under-Edge, Glos GL12 7LP, UK)

For centuries composers have been concerned with spatialization of sound and with the use of acoustic spaces to create feeling, atmosphere, and musical structure. This paper will explore Rob Godman's own use of sound in space, including (1) his treatment of ancient Vitruvian principles and how they are combined with new technologies; (2) an exploration of virtual journeys through real and imaginary acoustic spaces; (3) how sounds might be perceived in air, liquid, and solids; and (4) how technology has allowed composers to realize ideas that previously had only existed in the imagination. While focusing on artistic concerns, the paper will provide information on research carried out by the composer into acoustic spaces that are able to transform in real time with the aid of digital technology (Max/MSP software with sensor technology) and how these have been used in installation and pre-recorded work. It will also explore digital reconstructions of Vitruvian theatres and how we perceive resonance and ambience in the real and virtual world.

2:30

2pMUa7. Music-therapy analyzed through conceptual mapping. Rodolfo Martinez (CIICIR, IPN, Oaxaca, Mexico, Rodolfo_cm@yahoo.com) and Rebeca de la Fuente (IMA, Mexico)

Conceptual maps have been employed lately as a learning tool, as a modern study technique, and as a new way to understand intelligence, which allows for the development of a strong theoretical reference, in order to prove the research hypothesis. This paper presents a music-therapy analysis based on this tool to produce a conceptual mapping network, which ranges from magic through the rigor of the hard sciences.

2:45

2pMUa8. Some recent examples of fractal music. Cesar Guerra-Torres, Moises Hinojosa-Rivera, Juan Angel Garza-Garza, and 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, cguerratorres@hotmail.com)

Since the first published studies of fractal music by Mandelbrot and Voss, the relationship between music, mathematics, and fractal geometry has been a very active field of research. It has been found that the music of classical composers can be characterized by fractal or self-affine parameters, which in turn serve as the basis for synthetic fractal music. In this work is presented a brief discussion of the state of the art as well as some recent examples of fractal music, including a live demonstration.

TUESDAY AFTERNOON, 3 DECEMBER 2002 CORAL SEA 1 AND 2, 3:30 TO 5:10 P.M.

Session 2pMUb

Musical Acoustics: Musical Instruments of the South American Dance Tradition

Paul A. Wheeler, Chair
Utah State University, Logan, Utah 84322

Invited Papers

3:30

2pMUb1. An overview of musical instruments used in South American dance traditions. Paul A. Wheeler (Utah State Univ., Logan, UT 84322)

Musical instruments used in South American dances combine elements from Amerindian, African, and European musical traditions. The Amerindian influence can be seen in Andean instruments, such as the end-blown flute, panpipe, and charango (modified from the European guitar). The berimau, a musical bow used in the Brazilian capoeira dance, is an example of African influence. The bandoneon is a square-ended German concertina most famous for its use in the tango from Argentina. This paper provides an overview of the musical instruments commonly used in South American dance traditions in relationship to their origins. The acoustics of some of these instruments, such as the guitar, has been studied in detail, whereas others, like the Brazilian cuica, provide opportunity for new studies.

3:50

2pMUb2. Musical instruments of Brazilian capoeira: Historical roots, symbolism, and use. Beatriz Ilari (Faculty of Music, McGill Univ., 555 Sherbrooke St. W., Montreal, QC H3A 1E3, Canada, bilari@po-box.mcgill.ca)

This paper describes the historical roots, symbolism, and uses of musical instruments in capoeira. A martial art form of Afro-Brazilian origin, capoeira is rhythmically performed to music in a roda (i.e., circle). Capoeira is at times defined as a martial art form disguised as dance because it is rooted in the struggles of African slaves. Elements of music, dance, fight, and ritual are part of this unique martial art form, which has two main styles: Angola and Regional. Capoeira styles are important as they determine rhythmic patterns, chant, movement, and musical instrumentation in a roda. The leading instrument in all capoeira styles is the berimbau. The instrument dictates the rhythm and movement of capoeira players in a roda (Ilari, 2001). Made out of a wooden stick, a wire, and a gourd and played with a stick and a coin, the berimbau is considered a sacred instrument due to its association with the cry of the slaves. Other instruments used in capoeira are pandeiros, agogo bells, reco-recos, and atabaques. A discussion regarding the use of these instruments within the context of capoeira will be presented at the conference. The incorporation of these instruments into contemporary Brazilian music will also be considered.

4:10

2pMUb3. An acoustic study of the Brazilian cuica. Paul A. Wheeler (Utah State Univ., Logan, UT 84322)

The cuica is a friction drum of African origin played in the batucada (an ensemble of instruments used for the samba) during the Brazilian carnival. It is played by rubbing a bamboo rod which is connected to the center of a drum head, giving a rhythmic grunting sound. Pitch is changed by applying pressure to the membrane. This paper discusses several acoustic aspects of a folk cuica (made of a gourd) including the waveforms, spectra, and time envelopes produced. Rubbing the bamboo rod gives a primitive saw-toothed excitation, similar to a bowed violin string. This is connected to the center of a membrane which modifies and radiates the sound. The body of the cuica contributes little to the sound.

4:30

2pMUb4. Samba and the other sambas: Instrumentation in different forms of Brazil's main musical genre. Pablo Majlis (Faculte de Musique, Universite de Montreal, 109 Chestnut Ave., Pointe-Claire, QC H9R 3B2, Canada, pablomajlis@yahoo.com) and Beatriz Ilari (McGill Univ., Montreal, QC H3A 1E3, Canada)

The aim of the present paper is to describe the instruments of the different forms of samba, their origin, and their uses, focusing on percussion instruments. Samba is a Brazilian popular genre that developed mainly during the 20th century, though being deeply rooted in the precedent centuries of colonization and metissage between the Portuguese colonizers and the Africans that were brought as slaves. From its origins to the present day, samba has branched into multiple forms and instrumentation. Perhaps the most famous samba form is the samba enredo. This type of samba accompanies the Carnival parade in Rio de Janeiro, and features hundreds of percussionists. Another possible samba group instrumentation can be as simple as a single voice and a matchbox played by the singer. Between these two extremes there are several possible formations for a samba group, depending on the social context and function in which it occurs. Different group formations sometimes imply different song forms. Examples include samba de roda (i.e., circle samba), samba de gafieira (i.e., ballroom samba), and samba-cancao (i.e., samba ballad), among others. Some instruments will be available for attendees to try during the conference.

4:50–5:10

An opportunity will be provided for those in attendance to try some of the instruments.

TUESDAY AFTERNOON, 3 DECEMBER 2002 CORAL GARDEN 2 AND 3, 1:00 TO 5:30 P.M.

Session 2pNS

Noise and Architectural Acoustics: Predicting Noise in Indoor Industrial Spaces

Murray R. Hodgson, Cochair

Occupational Hygiene Program, University of British Columbia, 2206 East Mall, Vancouver, British Columbia V6T 1Z3, Canada

Frank H. Brittain, Cochair

Bechtel Corporation, 50 Beale Street, San Francisco, California 94105

Chair's Introduction—1:00

Invited Papers

1:05

2pNS1. Overview of predicting noise levels in indoor industrial spaces. Frank Brittain (Bechtel Corp., 50 Beale St., San Francisco, CA 94105, fhbritta@bechtel.com)

Predicting indoor noise in industrial facilities is a vital part of designing industrial plants. The predicted levels are used in the design process to determine compliance with occupation noise exposure limits, and to estimate levels inside the walls as starting point for predicting community noise radiated by buildings. Once levels are predicted, the noise controls needed can be developed. Special methodologies are needed, because normal room acoustics found in architectural acoustics texts is valid only for nearly empty rooms with limited absorption and ranges of room dimensions. The fittings inside industrial spaces can profoundly affect the propagation of noise and the resulting noise levels. In an industrial space, such as a power plant, there is no such thing as a reverberant field, except in isolated areas. In industrial spaces, including factories, predicting noise levels by summing free and reverberant fields gives erroneous results that are usually overly conservative. This paper discusses normal empty room acoustics, and problems typically encountered when it is applied to industrial spaces, particularly those with a high density of fittings or very large spaces. Also, alternative methodologies for predicting indoor noise levels in industrial spaces, which are based on standards and software, are identified and discussed.

1:35

2pNS2. Acoustics of industrial buildings. Michael Vorlaender (Institut fuer Technische Akustik, Aachen Univ. (RWTH), 52056 Aachen, Germany)

Industrial halls have significantly different shapes as rooms for other purposes. Typically these enclosures are very large and/or flat, which means that the room height is usually very much smaller than width and length. Sound fields in these types of enclosures cannot be expected to have an approximate diffuse sound field. Instead, the decay curves are bent, while sound level versus distance (sound propagation curve) is declining. This contribution summarizes physical properties of sound fields in flat and long rooms, gives some examples of predictions based on image sources, on Kuttruffs integral equation, results of scale model experiments and ray-tracing algorithms with two different levels of complexity. It can be shown that in many cases predictions can yield sufficiently accurate results. However, problems will occur in cases where rooms are heavily filled with scattering objects (fittings). On the other hand, the exact spatial distribution of sound in these nondiffuse cases has shown to have less influence on the perceived annoyance than expected. Accordingly, predictions of sound levels with reasonable accuracy should be sufficient to describe the subjective effect of listening, unpleasantness and annoyance in industrial halls.

2:05

2pNS3. Predicting noise in industrial workrooms using empirical models. Murray Hodgson (School of Occupational and Environ. Hygiene, Univ. of British Columbia, 3rd Fl., 2206 East Mall, Vancouver, BC V6T 1Z3, Canada)

Sound fields in complex industrial workrooms can be predicted well using numerical procedures such as the method of images and ray tracing. However, this requires acoustical expertise, as well as computational resources and times which result in prediction methods only being used in special cases. This paper discusses alternative empirical prediction methods which have the potential to be sufficiently accurate in ''typical'' cases, and more readily accessible to practitioners, making them more likely to be used in practice. The first method discussed is a hybrid approach, whereby characteristic workroom sound-propagation curves are predicted using ray tracing. These are then input into an empirical model which sums the energy contribution of all sources at a receiver position based on those curves and the applicable source/receiver distances. Next, the development of empirical models for predicting frequency-varying sound-propagation curves and reverberation times using regression techniques is discussed. These were developed from data measured in actual workrooms when empty or fitted, without and with sound-absorptive treatment. Empirical methods for estimating workroom fitting densities and multisource noise levels, and the integration of the empirical models into the PlantNoise prediction system, are also discussed.

2:35

2pNS4. A web-based noise control prediction model for rooms using the method of images. Stephen Dance (School of Eng., South Bank Univ., Borough Rd., London, UK, dances@sbu.ac.uk)

Previous simple models could only predict sound levels in untreated rooms. Now, using the method of images, it has become possible to accurately predict the sound level in fitted industrial rooms from any computer on the Internet. Thus, a powerful tool in an acoustician's armory is available to all, while requiring only the minimal amount of input data to construct the model. This is only achievable if the scope of the model is reduced to one or two acoustic parameters. Now, two common noise control techniques have been implemented into the image source model: acoustic barriers and absorptive patches. Predictions using the model with and without noise control techniques will be demonstrated, so the advantages can be clearly seen in typical industrial rooms. The models are now available on the web, running directly inside Netscape or Internet Explorer.

3:05–3:15 Break

3:15

2pNS5. Application of indoor noise prediction in the real world. David N. Lewis (SEAC, Unilever R&D, Colworth House, Sharnbrook, Bedfordshire MK44 1LQ, UK)

Predicting indoor noise in industrial workrooms is an important part of the process of designing industrial plants. Predicted levels are used in the design process to determine compliance with occupational-noise regulations, and to estimate levels inside the walls in order to predict community noise radiated from the building. Once predicted levels are known, noise-control strategies can be developed. In this paper an overview of over 20 years of experience is given with the use of various prediction approaches to manage noise in Unilever plants. This work has applied empirical and ray-tracing approaches separately, and in combination, to design various packaging and production plants and other facilities. The advantages of prediction methods in general, and of the various approaches in particular, will be discussed. A case-study application of prediction methods to the optimization of noise-control measures in a food-packaging plant will be presented. Plans to acquire a simplified prediction model for use as a company noise-screening tool will be discussed.

3:45

2pNS6. Validation–comparison of predicted and measured levels in industrial spaces. Wolfgang Probst (ACCON GmbH, Grfelfinger Str. 133A, 81375 Munich, Germany)

Financed by the German Federal Agency for Labor and Social Affairs, several methods of calculating the noise propagation in industrial halls were used for about 150 halls and compared with measurements. With all noise sources, such as machines and equipment stopped, a dodecahedron loudspeaker emitting broadband noise was used, and the octave-band levels were measured on different propagation paths. The room geometry and equipment in the room were entered into uniform datasets, the calculation methods were applied to each dataset, and the results in terms of deviations between calculated and measured values were evaluated

statistically. The prediction method with the smallest deviations was chosen for further evaluation. This method, which uses mirror images with approximations and takes into account diffraction with a method first developed by Kuttruff and extended by Jovicic, uses mean values for fittings and absorption at walls and ceilings. This method has been incorporated into VDI 3760, and will be extended in the future to take into account screening by single objects and the real distribution of absorptive materials on surfaces. Representative experimental results, calculation techniques, predicted levels, and deviations between measured and predicted levels are presented.

4:15–5:30

Panel Discussion

TUESDAY AFTERNOON, 3 DECEMBER 2002 CORAL ISLAND 1 AND 2, 1:00 TO 4:45 P.M.

Session 2pPA

Physical Acoustics: Bubbles, Drops and Foams II

Joachim Holzfuss, Chair
Nonlinear Physics Group, Institute of Applied Physics, Technical University of Darmstadt, Darmstadt D-62489, Germany

Invited Papers

1:00

2pPA1. Encapsulated bubble dynamics. John Allen III (510 Arthur St., Apt. # 116, Davis, CA 95616, johnallen312@hotmail.com)

The study of a bubble encapsulated by a fluid or an elastic shell is a subject of interest for a wide variety of applications. In particular, ultrasound contrast agents are encapsulated bubbles 1–5 microns in radius developed for diagnostic imaging and, more recently, therapeutic purposes involving drug delivery. The previous formulations of the equations for encapsulated bubbles have originated from a generalized Rayleigh–Plesset equation, which follows from the Eulerian form of the fluid dynamics equations. Dynamic formulations of empty spherical shells in an inviscid medium have been also developed independently of these gas bubble studies. Little effort has gone into unifying the two approaches. The equations for a gas-filled incompressible, isotropic elastic spherical shell are derived from the Lagrangian frame using the first Piola–Kirchoff stress tensor. Some previous results are obtained and compared in different limiting cases. Instabilities of the shell are determined for sufficiently flexible shell materials or high internal gas pressures. Also highlighted are the nonspherical instabilities associated with acoustically driven bubbles with fluid shells of different density and viscosity than that of the surrounding fluid.

1:25

2pPA2. Single bubble sonoluminescence: Unstable diffusion and the kinetics of chemical reactions. Joachim Holzfuss (Nonlinear Phys. Group, Inst. of Appl. Phys., Tech. Univ. of Darmstadt, Germany)

Sonoluminescence of a single bubble in water driven by ultrasound is accompanied by long-term stability of the radial bubble wall oscillation. However, in a certain parameter range instabilities of the phenomenon are observed. They are characterized by growing of the ambient (and maximum) radius of the bubble and a sudden microbubble split-off with spatial dislocation (recoil). Experimental results using images of shock waves shed into the surrounding water at bubble collapse are shown to visualize the effect. Numerical calculations show that the observed nonlinear dynamical effects can be interpreted by the influence of diffusional, chemical, and spatial instabilities of the bubble.

1:50

2pPA3. Acoustically-driven spherical implosions and the possibility of thermonuclear reactions. D. Felipe Gaitan, Ross A. Tessien (Impulse Devices, Inc., 12731-A, Grass Valley, CA 95945, gaitan@impulsedevices.com), and William C. Mead (Adaptive Network Solutions Res., Inc., Los Alamos, NM)

Acoustically driven, gas-filled cavities in liquids have been known to collapse violently, generating short flashes of light of ~100 psec duration. This phenomenon has been known as Sonoluminescence (SL) and was first observed by Marinesco *et al.* in 1933. Ten years ago the author pioneered a technique for observing the oscillations of a single, gas-filled cavity (termed Single-Bubble Sonoluminescence) which has provided new insights into the phenomenon of sonoluminescence. More recently, the possibility of generating fusion reactions using acoustics has been considered. Results of computer simulations and preliminary experimental data will be presented. Back-of-the-envelope calculations in terms of the acoustical and thermodynamic parameters necessary to achieve Thermonuclear reactions will be presented in an effort to evaluate the feasibility of Sonofusion as an energy source.

2pPA4. Evidence for nuclear emissions during neutron seeded acoustic bubble cavitation. R. P. Taleyarkhan, C. D. West, J. S. Cho (Oak Ridge Natl. Lab., P.O. Box 2009, Bldg. 9204-1, Oak Ridge, TN 37831), R. T. Lahey, Jr., R. C. Block (Rensselaer Polytechnic Inst., Troy, NY), and R. Nigmatulin (Russian Acad. of Sci., Ufa, Russia)

In cavitation experiments with deuterated acetone, statistically significant tritium decay activity above background levels was detected. In addition, evidence for statistically significant neutron emissions near 2.5 MeV was also observed, as would be expected for deuterium–deuterium fusion. Control experiments with normal acetone did not result in tritium activity or neutron emissions. Hydrodynamic shock code simulations supported the observed data and indicated compressed, hot (10^6–10^7 K) bubble implosion conditions, as required for thermonuclear fusion reactions. Separate experiments with additional fluids are under way and results appear to support those observed with acetone. Scalability potential to higher yields, as well as evidence for neutron–tritium branching ratios are presented.

2:40–3:00 Break

Contributed Papers

3:00

2pPA5. Dispersion relation measurements of acoustic waves in bubbly water. Gregory J. Orris and Michael Nicholas (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375)

Recent theoretical work on the propagation of acoustic waves in bubbly media has highlighted the need for more precise and modern measurements of the relationship between the phase speed and attenuation in bubbly media. During the engineering tests of the new Salt-Water Tank Facility at the Naval Research Laboratory measurements of the dispersion of acoustic waves in fresh water were performed over a broad range of environmental conditions under semi-free field conditions. Large aquaculture aeration tubes were used to create bubble clouds completely filling the facility with bubbles whose radii ranged from a few tens of microns to 1 cm with total void fractions that reached to a few percent. We discuss these experimental results within the context of current theories and their implications on ocean acoustic experiments. [Work supported by ONR.]

3:15

2pPA6. Hydroacoustical interaction of bubble clouds. Stefan Luther and Detlef Lohse (Phys. of Fluids, Faculty of Appl. Phys., Univ. of Twente, The Netherlands)

Acoustically driven cavitation bubble fields consist of typically 10^4 micron-sized bubbles. Due to their nonlinear hydroacoustical interaction, these extended multiscale systems exhibit the phenomenon of spatiotemporal structure formation. Apart from its significance for the theory of self-organization, it plays a major role in design and control of many industrial and medical applications. Prominent examples are ultrasound cleaning, sono-chemistry and medical diagnostics. From a fundamental point of view the key question to ask is ''How does the fast dynamics on small length scales determine the global slow dynamics of the bubble field?'' To clarify the complex interplay of acoustical and hydrodynamical forces acting on the bubbles, we employ high-speed particle tracking velocimetry. This technique allows the three-dimensional reconstruction of the bubbles' trajectories on small and fast scales as well as the measurement of the bubble density on large and slow scales. A theoretical model is derived that describes the nonlinear radial and translational dynamics of the individual bubbles and their interaction. The numerical solution of this N body problem is presented.

3:30

2pPA7. Phase velocity measurements in a bubble swarm using a fiber optic sensor near the bubble resonant frequency. Stanley A. Cheyne (Dept. of Phys. and Astron., Hampden-Sydney College, Hampden-Sydney, VA 23943)

Acoustic phase velocity measurements of a bubble swarm in a cylindrical tube have been made with a fiber optic sensor. The fundamental design of this system is similar to one used in a previous experiment [S. A. Cheyne *et al.*, ''Phase velocity measurements in bubbly liquids using a fiber optic laser interferometer,'' J. Acoust. Soc. Am. **97**, 1621 (1995)]. This new system is more robust and is more easily constructed than the previous system. Results will be presented that will show conclusively that data have been obtained just after the bubble resonance in the regime where the attenuation of sound is very high. Other results will be presented at different air-to-water ratios (void fraction).

3:45

2pPA8. Statistical characteristics of cavitation noise. Karel Vokurka (Phys. Dept., Tech. Univ. of Liberec, Halkova 6, CZ-461 17 Liberec, Czech Republic, karel.vokurka@vslib.cz)

Cavitation noise originates as a superposition of pressure waves emitted during oscillations of individual cavitation bubbles. These pressure waves contain useful information on bubbles generating them and efforts are done to extract it. Unfortunately the pressure waves emitted by different bubbles usually overlap heavily and thus in experiments it makes sense to measure statistical characteristics only. Typical statistical characteristics determined experimentally encompass autospectral densities and instantaneous autospectra. To be able to extract information concerning the oscillating bubbles, suitable models of both cavitation bubbles and cavitation noise are necessary. It has been found out recently that a reasonable insight into the cavitation noise structure may be obtained by simulating cavitation noise on a computer and comparing statistical characteristics of simulated cavitation noise with those determined experimentally. By varying different parameters in theoretical models used to simulate the noise, a good agreement between the simulated and measured cavitation noise statistical characteristics can be obtained. The models parameters thus found may be then analyzed from a physical point of view and conclusions on behavior of cavitation bubbles can be drawn. [Work supported by the Ministry of Education of the Czech Republic as the research Project No. MSM 245100304.]

4:00

2pPA9. Time-scales for quenching single-bubble sonoluminescence in the presence of alcohols. Jingfeng Guan and Thomas Matula (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

A small amount of alcohol added to water dramatically decreases the light intensity from single-bubble sonoluminescence [Weninger *et al.*, J. Phys. Chem. **99**, 14195–14197 (1995)]. From an excess accumulation at the bubble surface [Ashokkumar *et al.*, J. Phys. Chem. **104**, 8462–8465 (2000)], the molecules evaporate into the bubble interior, reducing the effective adiabatic exponent of the gas, and decreasing the bubble temperature and light output [Toegel *et al.*, Phys. Rev. Lett. **84**, 2509–2512 (2000)]. There is a debate as to the rate at which alcohol is injected into the bubble interior. One camp favors the notion that molecules must be repetitively injected over many acoustic cycles. Another camp favors the notion that most quenching occurs during a single collapse. An experiment has been conducted in order to resolve the debate. Quenching rates were measured by recording the instantaneous bubble response and corresponding light emission during a sudden increase in pressure. It was found that

complete quenching in the presence of methanol requires over 8000 acoustic cycles, while quenching with butanol occurs in about 20 acoustic cycles. These observations are consistent with the view that quenching requires the repetitive injection of alcohol molecules over repetitive acoustic cycles.

4:15

2pPA10. A KC-135 experiment for studying single bubble sonoluminescence: Design, fabrication, and results. Ronald A. Roy and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

A detailed description of the design and fabrication of an apparatus to study SBSL on the NASA's KC-135 parabolic flight aircraft will be presented. The apparatus was used during two recent flights aboard the KC-135 (March and July 2002); data from these flights will be presented. Parameters measured during the flights include the acoustic pressure, ambient pressure, acceleration, bubble size, bubble location, water temperature, and light intensity. All measurements were made in a simultaneous fashion, and time tagged using the time code from a VCR which was used to record video images of the bubble. A review of and comparison to data from other KC-135 SL experiments will also be given. [Work supported by NASA.]

4:30

2pPA11. High frequency acoustic scattering by a gas bubble. Nail A. Gumerov (UMIACS, Univ. of Maryland, 115 A. V. Williams Bldg., College Park, MD 20742)

A problem of acoustic scattering by a gas bubble when the length of the incident acoustic waves can be comparable with or smaller than the bubble size is considered. For such waves pressure and temperature distributions inside and outside the bubble are not spherically symmetrical even though the bubble is spherical in the absence of the acoustic field and the amplitude of the acoustic field is small. This case is considered in the present study. Genearal three-dimensional pressure and temperature distributions together with capillary effects are taken into account by expansion of solutions for coupled thermal and acoustic problems in series of spherical multipoles. The acoustic response of the bubble including volume and surface oscillations is analyzed. Limits for modeling of the bubble as a sound-soft sphere in acoustic scattering problems are discussed.

TUESDAY AFTERNOON, 3 DECEMBER 2002 GRAND CORAL 3, 1:30 TO 5:00 P.M.

Session 2pPP

Psychological and Physiological Acoustics: Psychological and Physiological Acoustics Potpourri (Poster Session)

Walt Jesteadt, Cochair
Boys Town National Research Hospital, 555 North 30th Street, Omaha, Nebraska 68131

Joel Flores, Cochair
3A Privada de Mimosas No. 617, Villa de las Flores, Coacalco, Edo. de Mexico CP 55710, Mexico

Contributed Papers

All posters will be on display from 1:30 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:30 p.m. to 2:45 p.m. and contributors of even-numbered papers will be at their posters from 2:45 p.m. to 5:00 p.m.

2pPP1. Computer-based method for the diagnostics and rehabilitation of tinnitus patients. Bozena Kostek, Andrzej Czyzewski, and Henryk Skarzynski (Inst. of Physiol. and Pathol. of Hearing, Pstrowskiego 1, Warsaw, Poland)

The proposed method is an electronic diagnostic and rehabilitation system for people suffering from internal ear noise, namely tinnitus patients and people with an abnormally high hearing sensitivity—hyperacusis patients. Thanks to the method employing multimedia personal computers to the fitting masking sounds to the patient's needs, tinnitus patients can be rehabilitated using the masking or habituation method. The subject is asked to answer detailed questions in the electronic questionnaire and next to identify those sounds which most strongly resemble those they perceive as ear noise. Following an algorithm analysis of the results and analysis of the selection of sounds made by the patient, the computer diagnoses the patient as free from ear noise and hyperacusis or classifies them into a risk group. Next, the patient is informed about the result and can then read about the causes of the ailment and the recommended treatment and also can download compressed sounds applicable in therapy to the rehabilitation tool coupled with the personal computer which is a programmable ear noise masker. In the paper the algorithm of analysis of patients answers and acoustic characteristics of tinnitus maskers are discussed. Hitherto obtained results of this method application are demonstrated.

2pPP2. Principles and acoustical foundations of the computer-based hearing screening method. Henryk Skarzynski, Andrzej Czyzewski, and Bozena Kostek (Inst. of Physiol. and Pathol. of Hearing, Pstrowskiego 1, Warsaw, Poland)

The hearing impairment is one of the fastest growing diseases in modern societies. Therefore it is very important to organize screening tests allowing us to find people suffering from this kind of impairment. The computer-based system was designed to conduct hearing screening,

mainly in children and youth. The test uses automatic questionnaire analysis, audiometric tone test procedure, and testing speech intelligibility in noise. The starting point of the test is an automatic interview with the individual to be tested. Based on the interview, the electronic questionnaire is filled out. After the questionnaire has been filled out and the specially conceived three tone audiometric test is completed, the mode of the speech-in-noise based test might be selected as appropriate for the specific age. When all the testing is completed, the system ''I CAN HEAR . . . '' automatically analyzes the results for every person examined. Based on the number of wrong answers those who may have hearing problems are referred to cooperating medical consulting centers. In the paper foundations and principles of the hearing tests are discussed and results of testing of more than 200 000 children with this method are demonstrated.

2pPP3. The effects of high intensity pure tone on visual field, eye fixation, pupil size, and visual false positive and negative errors. Hashir Aazh, Ali Nouraeinejad, Ali Asghar Peyvandi, and Latif Gachkar (P.O. Box 17445-177, Tehran, Iran, hash_aazh@hotmail.com)

Our aim in this study is to evaluate the effects of pure tone on visual field, pupil size, eye fixation, and visual false positive and negative errors. Thirty-two young adult subjects with normal hearing and normal visual acuity were tested in this study. Measurements were performed over two test sessions. In one session visual factors were measured in quiet, and another measurement was performed during presenting a pure tone (1000 Hz, 100-dB HL) binaurally via a headphone. The statistical program SPSS10 was used for all analyses. Fixation loss was significantly lower (better) in quiet condition compared with continuous pure tone condition, and other visual factors had no significant differences in the two conditions. As a conclusion it can be stated that changes in attentional focus which result from altered levels of arousal or autonomic system activity during presenting a high intensity pure tone affect on fixation of the eye. However, it is suggested that these effects are possibly related to stress caused by sound masking the hearing of speech and other wanted environmental sounds, and not from some direct autonomic system arousal by sound.

2pPP4. On the lateralization of the Huggins pitch. Peter Xinya Zhang and William M. Hartmann (Phys.-Astron., Michigan State Univ., East Lansing, MI 48824 and Biomed. Eng., Boston Univ., Boston, MA 02215)

The central activity pattern (CAP) model of Raatgever and Bilsen [J. Acoust. Soc. Am. **80**, 429–441 (1986)] correctly predicts that Huggins pitch (HP^+) is lateralized in the center whereas HP^- is lateralized to the left or the right. Experiments show that some listeners (left-eared listeners) always hear the pitch sensation on the left and others always hear it on the right. Still others can hear it on one side or the other. The CAP model also predicts that the laterality of HP^- should follow a hyperbolic function of the boundary frequency. To test this prediction, laterality was measured in careful laterality estimation experiments, wherein HP^- was combined with a set of interaural time differences (ITDs). Although laterality estimates followed predictions for finite (ITDs), on those trials where the ITD was zero the hyperbolic law was violated for five out of five listeners. Instead, the laterality of HP^- was very insensitive to the boundary frequency over the range tested, 200 to 1000 Hz. A search for a satisfactory variation on the CAP model continues. [Work supported by the NIDCD under Grants Nos. DC00181 and DC00100.]

2pPP5. Development of the positional presumption system of sound source which combines sound source information and picture image. Yamashita Yasuhiro (4-17-1 Wakasato, Nagano City, Japan)

The development of the positional presumption system of sound sources which combines sound information and picture images has been done by using five microphones. The calculation of the positional presumption of the sound source is obtained by calculating the time difference from the cross spectrum of the output of several microphones. There was a tendency that the direction presumption varies by enlarging the elevation of the position of the sound source and the receiving sound point. It is thought that this cause is influenced by the ground level reflection. We tried to separate the reflected sound and the removed method was shown. In addition, the methods of improving the calculation accuracy in consideration of the acoustic impedance of the ground surface were shown, and compared.

2pPP6. Functional segregation of segmental features, pitch-accents, and nondistinctive suprasegmental features in working memory. Akihiro Tanaka, Koichi Mori (Res. Inst. of Natl. Rehabilitation Ctr. for the Disabled, 4-1, Namiki, Tokorozawa-shi, Saitama 359-8555, Japan, tanaka@rehab.go.jp), and Yohtaro Takano (Univ. of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan)

Japanese has pitch-accent which contribute to lexical distinctiveness just as segmental features. A dual-task experiment was conducted to examine the independence among segmental features, pitch-accents, and nondistinctive suprasegmental features in working memory. Japanese speakers of Tokyo dialect participated in the experiment. The primary task was a working memory task which required the retention of both segmental and suprasegmental features of nonsense words. The suprasegmental features were either regular pitch-accent of Tokyo dialect or pseudopitch changes that do not exist in a Japanese accent pattern. The secondary task was silent mouthing of irrelevant verbal material either without putative pitch change, with pitch accent, or with pseudopitch change, that was performed during the retention period of the primary task. The results revealed selective interference according to distinctivity of pitch patterns in native speakers of Tokyo dialect, suggesting functional segregation between distinctive and nondistinctive suprasegmental features. In contrast, functional segregation was not observed in ''non-native'' speakers of Tokyo dialect, suggesting that the prosody of the non-native dialect acts as ''second language.'' Additionally, in both groups, suprasegmental processing did not interfere with the retention of segmental features. The results suggest that there are at least three substores for segmental features, pitch-accents, and nondistinctive suprasegmental features.

2pPP7. Electroacoustic verification of FM benefits in advanced hearing aid circuitry. Erin C. Schafer and Linda M. Thibodeau (Adv. Hearing Res. Ctr., Univ. of Texas, Dallas, 1966 Inwood Rd., Dallas, TX 75235, eschafer@utdallas.edu)

Miniature FM receivers attached to ear-level hearing aids can provide significant improvements in speech recognition in noisy environments. These receivers are designed to couple to the hearing aid via a boot connection with limited or no adjustment of hearing aid settings. Ideally, the circuitry allows the FM signal to be approximately 10 dB more intense than the typical signal from the hearing aid microphone, allowing for an FM advantage. Furthermore, for intense input levels, the output limiting should not differ for the hearing aid compared to the hearing aid and FM receiver combined, i.e., FM transparency. Following procedures recommended by the American Speech, Language, and Hearing Association, the electroacoustic responses of digital and conventional aids were measured with and without coupling to FM systems at typical conversational input levels and maximum input levels. The rms difference between the electroacoustic responses of the hearing aid alone and coupled to the FM was used to quantify the FM advantage at typical input levels and the FM

transparency at high input levels. The finding of great variability in both FM advantage and transparency supports the need for additional fitting controls or design modifications to obtain the maximum FM benefit.

2pPP8. A comparison of middle ear acoustic admittance in adults and 3-week-old infants based on multifrequency tympanometry. Linda Polka (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, linda.polka@McGill.ca), Navid Shahnaz (Univ. of British Columbia, BC, Canada), and Anthony Zeitouni (McGill Univ., Montreal, QC, Canada)

The assessment of newborn hearing requires information on middle ear status yet the interpretation of tympanometry in newborns is unclear. This study aims to further our understanding of acoustic admittance in the newborn middle ear. Multifrequency tympanograms were recorded from sixteen 3-week-old infants (30 ears) and sixteen young normal-hearing adults (30 ears). Tympanometry was conducted using the Virtual 310 middle ear analyzer using 9 probe tone frequencies between 226 and 1000 Hz at roughly 100 Hz intervals. All infants passed a hearing screening using automated ABR (Algo II) shortly after birth (within 24 h) and again at 3-weeks of age. At 226 Hz, admittance tympanograms had a single peak in all adult ears while 60% of infant ears had multiple peaks or irregular patterns. At 1000 Hz admittance tympanograms had a single peak for 74% of infant ears while 78% of adult ears showed multiple peak or irregular patterns. Analyses of tympanometric shape (using the Vanhuyse classification scheme), as well as static admittance, static susceptance, and static conductance also reveal differences in adult and infant middle ear function. Implications for the clinical application of tympanometry in the first month of life will be discussed.

2pPP9. Estimates of the strength of repetition pitch in infants. Marsha G. Clarkson (Dept. of Psych., Georgia State Univ., University Plaza, Atlanta, GA 30303-3083, mclarkson@gsu.edu), Cynthia M. Zettler, Michelle J. Follmer, and Michael J. Takagi (Georgia State Univ., Atlanta, GA 30303)

To measure the strength of the pitch of iterated rippled noise (IRN), 24 7- to 8-month-olds and 24 adults were tested in an operant conditioning procedure. To generate IRN, a 500-ms Gaussian noise was delayed by 5 or 6 ms (pitches of 200 and 166 Hz) and added to the original noise for 16 iterations. IRN stimuli having one delay were presented repeatedly, and on signal trials the delay changed for 6 s. Overall stimulus level roved from 63–67 dBA. Infants learned to turn their heads toward the sound, and adults learned to press a button when the delay of the stimulus changed. Testing started with IRN stimuli having 0 dB attenuation (i.e., maximal pitch strength). Then, stimuli having weaker pitches (i.e., progressively greater attenuation applied to the delayed noise) were presented. Strength of pitch can be quantified as the maximum attenuation for which pitch can be discerned. For each subject, threshold attenuation for pitch strength was extrapolated as the 71% point on a psychometric function depicting percent-correct performance as a function of attenuation. Mean thresholds revealed that the pitch percept was significantly weaker for infants (6.9 dB) than for adults (19.1 dB).

2pPP10. Noise addiction. 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)

Progress in understanding the brain processes involved in addiction to chemicals has resulted in the recognition of other forms of addiction related to energy and information, i.e., TV and Internet. In that context noise becomes a cause of addiction and, therefore, it is necessary to develop more adequate approaches even to conceptualize noise addiction. In this presentation some of the main aspects to be considered to approach an operational definition of noise addiction are discussed.

2pPP11. Effect of noise on social relationships. Karla Panuszka (Acoust. Consulting, 2929 Beverly Ln., Lafayette, IN 47904) and Ryszard Panuszka (Staszic Univ., AGH Krakow, Poland)

Sociology is the scientific study of human society and social behavior. The effect of noise and vibrations under unwanted vibroacoustical fields shows important changes in social interactions. A new social question arises from these studies: What effect does unwanted noise pollution play on human environments and how are relationships changed? It has been observed in workers that noise and vibrations play an important factor in machinery safety. Music shows a positive effect on relationships by providing a person's mental attitude with a happy euphoria. Natural sounds of waterfalls, streams, and rivers also show an improved mood that enhances relationships in a positive way. Staszic University AGH has documented these interactions under the influence of acoustical fields. The human ear cannot hear sound waves below 20 Hz, but these waves have been shown to have effects on people's interactions. The main focus of sociology is interactions of large populations; thus, noise interactions need to be studied. Three areas of investigations could arise from these studies: family relationship responses, occupational safety, and domestic and business violence. Broad investigation results will be displayed.

2pPP12. Validation of audiometric method for measuring temporary threshold shifts. Rodrigo Ordoñez, Ioanna-Rigina Karariga, Brian L. Karlsen, Karen Reuter (Dept. of Acoust., Aalborg Univ., Denmark, rop@acoustics.auc.dk), and Dorte Hammershøi (Aalborg Univ., Denmark)

The proposed experiment aims at testing the audiometric method developed earlier [Ordoñez *et al.*, Acta Acustica (Beijing) **88**, 450–452 (2002)] by evaluating if it is able to reproduce some of the known aspects of temporary threshold shifts (TTS), such as: Immediate sensitization, the 2 min maximum of the recovery and the half-octave shift of the peak of TTS. The subjects will be exposed to pure tones of 500 Hz and 2 kHz for 2 to 5 min and the levels will cover the range between 40- to 100-dB SPL. The aim is to induce a maximum of 15 dB of TTS. For showing the half-octave shift, the threshold will be determined from one octave below the exposure frequency to one octave above in 1/2 octave steps. For the 2 min maximum, the threshold will be measured at the most affected frequency and it can be determined continuously for 4 min after the exposure. Sensitization at probe frequencies lower than the exposure frequencies will be tested by measuring the threshold at 1/2 and 1 octave below the exposure frequency.

2pPP13. Scientific discovery of the function of ear in the light of material property. Hari S. Paul and M. Kumaresan (Int. Res. Inst. for the Deaf—A component of Acoust. Foundation, 94/159 Avvai Sanmugam Salai, Chennai-600014, India, mirc@md3.vsnl.net.in)

Piezoelectricity is conversion of mechanical energy to electric energy and vice versa. This property exhibits in noncentrocymetric materials. Bone is well-known piezoelectric material in living body. Eardrum is connected with bones (malleus, incus, stapes, and bony cochlea). Cochlea is snail-shaped and filled with fluid and hair cells (human electrodes). Fluid is centrocymetric; hence, it is nonpiezoelectric material. Acoustic pressure on eardrum sets mechanical energy to these bones. Bones convert mechanical energy to electric polarization, which is direct piezoelectric property. Electric charges generated in bones are transmitted through fluid as $\pm$ions (like a car battery charger) and picked up by $-/+$ hair cells and

auditory nerves for transmission to brain. Transmissions of ± charges through fluid generate movement of fluid. Cochlea replacement is required when hair cells in cochlea lose their power to transmit charges. Vocal cord surrounded by cricoid and arytenoid cartilages is known as vocal box/larynx. Larynx transforms pulse energy to sound energy (converse piezoelectric property). Vocal cord narrowing and opening its air passage from lungs can also produce sound. Present concept of ear function contradicts majority of doctor's view, that is, fluid in cochlea transforms sound (mechanical) energy to electric energy, which is untrue.

2pPP14. Tempo discrimination of isochronous tone sequences: The multiple-look model revisited. Nathaniel Miller and J. Devin McAuley (Dept. of Psych., Bowling Green State Univ., Bowling Green, OH 43403)

Previous research has shown that increasing the number of intervals in an isochronous tone sequence reduces tempo discrimination thresholds [C. Drake and M. Botte, Percept. Psychophys. **54**, 277–286 (1993)]. One question that arises is whether increased tempo sensitivity in this instance is attributed to multiple looks at the standard interval, comparison interval, or both. The present study addressed this question by examining tempo discrimination using isochronous tone sequences that contained variable numbers of standard and comparison intervals. In all cases, participants judged the tempo of the comparison sequence relative to a standard sequence (responding faster or slower). Preliminary results suggest that in some cases increases in tempo sensitivity are more due to repetitions of the comparison interval than to repetitions of the standard. The implications of these findings for theories of auditory tempo discrimination will be discussed.

2pPP15. Listening strategies used by normal-hearing adults during loudness estimation. Lori J. Leibold and Lynne A. Werner (Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98105-6246)

This study describes individual differences in the weight normal-hearing adults give to the frequency components of a complex sound when performing loudness estimation. In addition, the relationship between listening strategy and loudness was examined. A multi-tone complex (1000, 2000, and 4000 Hz) was presented for 500 ms to the left ear of seven normal-hearing adults. The level of each tone was selected independently and randomly from a rectangular distribution ranging in 10 dB steps from 40 to 80 dB SPL on each of 500 presentations. In the nonselective condition, subjects provided a numerical estimate of the loudness of the complex. In selective conditions, listening strategy was controlled by instructing subjects to attend to a single target frequency. Relative weights for each subject were estimated by normalizing the raw correlations calculated from the level of each component and the subjects magnitude estimate for each trial. Results for the nonselective condition revealed individual differences in subjects' weighting functions. In contrast, similar weighting functions were observed in the selective conditions, with subjects giving the greatest weight to the target frequency. The loudness growth data can best be described by weighting the intensity of each frequency component prior to determining the loudness growth.

2pPP16. Melody recognition with temporal and spectral cues in normal-hearing and cochlear-implant listeners. Ying-Yee Kong (Dept. of Cognit. Sci., Univ. of California, Irvine, CA 92697) and Fan-Gang Zeng (Univ. of California, Irvine, CA 92697)

Cochlear-implant users can achieve a high level of speech recognition, but their ability to appreciate music is severely limited. This study investigates the relative contribution of the temporal and spectral cues to melody recognition. Two sets of 12 familiar songs were generated: one set contained both rhythm and melody information (the rhythm–melody condition), whereas the other set contained only melody information (the no-rhythm condition, where all notes had the same duration). Normal-hearing listeners achieved 95%–100% accuracy in both conditions. While cochlear-implant listeners achieved 42%–94% accuracy in the rhythm–melody condition, they performed essentially at chance level in the no-rhythm condition. To further identify the cues used in melody recognition, temporal envelopes from the original broadband signal were extracted and modulated with white noise. When the rhythm cue was available, the normal-hearing listeners achieved 60%–100% recognition, which was similar to the performance achieved by the cochlear-implant listeners. When the rhythm cue was not available, both normal-hearing and cochlear-implant listeners performed essentially at chance level. These results suggest that the present cochlear-implant listeners relied solely on temporal information to recognize familiar melodies. The fine structure information is certainly needed to allow true appreciation of music for cochlear-implant users.

2pPP17. The effects of aging on spatial hearing in listeners with normal hearing. Janet Koehnke, Joan Besing, Ianthe Dunn-Murad, and Caryn Neuvirth (School of Grad. Medical Education, Seton Hall Univ., 400 S. Orange Ave., South Orange, NJ 07079)

This study was designed to examine the effects of aging on spatial hearing in young, middle-aged, and older adults with normal hearing. In order to determine the onset and progression of age-related changes in spatial hearing, performance was measured for three groups of listeners: 18 to 30 years, 38 to 50 years, and over 60 years. All listeners had hearing thresholds of 25 dB HL or better at octave frequencies from 250 through 8000 Hz. A series of virtual tests were administered including spatial localization in quiet, spatial localization in noise, spatial detection in noise, and speech intelligibility gain. For all of these tests, the source locations and listening environments were simulated using head-related-transfer-functions. Results indicate no clear or consistent differences between the young and middle-aged groups on any of the spatial tasks. In contrast, localization errors, spatial detection thresholds, and speech intelligibility thresholds are poorer for the older listeners than for the young and middle-aged listeners. It is noteworthy that all three groups obtain comparable gain in spatial detection thresholds and speech intelligibility thresholds in some listening conditions when the signal and noise sources are spatially separated. [Work supported by NIH/NIDCD Grant No. DC004402.]

2pPP18. Binaural phase masking experiments in stereo audio. Alexander I. Iliev and Michael S. Scordilis (Dept. of Elec. and Computer. Eng., Univ. of Miami, 1251 Memorial Dr., Coral Gables, FL 33124-0640)

Researchers have established that in binaural hearing the smallest detectable angular separation between two sources, commonly referred to as the minimal audible angle (MAA), for a pair of sources on the horizontal plane depends on the frequency of the emitted pure tone and the azimuth angular separation between the sources. One interesting way is to view the sources' angular perturbation within the MAA limits as noise in the phase domain, and the listener inability to detect this perturbation as the result of a masking process. The present discussion focuses on experimental procedures for examining the perception of MAA and the corresponding interaural phase difference (IPD) when complex sound sources are located in the most sensitive region, which is directly in front of the observer (both azimuth and elevation angles at 0° degrees). Sound stimuli were viewed as the linear combination of pure tones, as provided by Fourier analysis. Results indicate that masking is achieved when the IPD is disturbed within

a threshold limit corresponding to the MAA for pure tone sources. Listening tests using stereo audio further validated our observations. [Work supported by Watermark Technologies.]

2pPP19. Distinguishing feature misperception from illusory conjunctions in spatially distributed musical tones. Michael D. Hall and Kimberly Wieberg (Psych. Dept., Univ. of Nevada, Las Vegas, 4505 Maryland Pkwy., Box 455030, Las Vegas, NV 89154, hallm@unlv.edu)

Recent questions have been raised in the visual search literature concerning whether illusory conjunctions of correctly registered features occur (indicating a feature integration process) or are an artifact of feature misperception. The current investigation raised and addressed similar questions for findings of auditory illusory conjunctions using simultaneous, spatially distributed musical tones. Two experiments were conducted where musically trained listeners identified pairs of tones that reflected possible combinations of two instrument timbres with two pitches. Experiments differed in the spatial separation between simultaneous tones to potentially evaluate the effects of distance on feature perception/integration. In Experiment 1 tones were presented to opposing ears. In Experiment 2 tones were only slightly lateralized by a manipulation of interaural time disparities. Conjunction responses, reflecting the incorrect combination of features, frequently occurred, and were more common for slightly lateralized tones. To evaluate the perceptual event(s) responsible for conjunction responses, data were submitted to multinomial models that differed with respect to whether or not they allowed for illusory conjunctions, the misperception of features, or both errors. Across experiments, data fitting by models was improved by feature misperception, but was not further improved by illusory conjunctions. Implications for models of search performance and feature binding are discussed.

2pPP20. Dichotic pitch and the missing fundamental. Joseph Hall III, Emily Buss, and John Grose (Dept. of Otolaryngol., Univ. of North Carolina at Chapel Hill, 610 Burnett-Womack Bldg., CB 7070, Chapel Hill, NC 27599-7070)

A dichotic pitch can be heard by introducing interaural phase shifts in one or more low-frequency spectral regions of an otherwise diotic noise. The present study examined the perceived pitch associated with phase disparities in spectral regions corresponding to harmonics of a low-frequency fundamental. Of particular interest was whether the pitch of the complex was related to the frequency region of the harmonics or whether it was associated with the missing fundamental. Discrimination thresholds for this dichotic virtual pitch were estimated in a 3AFC paradigm, where listeners were asked to identify the interval with the lower virtual pitch. After modest training, listeners were able to perform this task well even when a randomized subset of harmonics above $F0$ were presented on each trial, a manipulation designed to make spectral region an unreliable cue. In another manipulation, frequency regions of the phase disparity were chosen such that although the frequency increased, the missing fundamental decreased (or vice versa). In these cases listeners reported that the perceived pitch shifted according to the change in $F0$, not according to the spectral region of the harmonics. Results were consistent with the interpretation that the perceived pitch correspond to the missing fundamental.

2pPP21. Auditory performance in an open field. Kim Fluitt, Tomasz Letowski, and Timothy Mermagen (U.S. Army Res. Lab., Auditory Res. Team, Bldg. 520, Aberdeen Proving Ground, MD 21005, kfluitt@arl.army.mil)

Detection and recognition of acoustic sources in an open field are important elements of situational awareness on the battlefield. They are affected by many technical and environmental conditions such as the type of sound, the distance to a sound source, terrain configuration, meteorological conditions, hearing capabilities of the listener, the level of background noise, and the listeners' familiarity with the sound source. A limited body of knowledge about auditory perception of sources located over long distances makes it difficult to develop models predicting auditory behavior on the battlefield. Our purpose in the present study was to determine the listeners' abilities to detect, recognize, localize, and estimate distances to sound sources from 25 to 800 meters from the listening position. Data were also collected for meteorological conditions (wind direction and strength, temperature, atmospheric pressure, humidity) and background noise level for each experimental trial. Forty subjects (men and women, ages 18–25) participated in the study. Nine types of sounds were presented from six loudspeakers in random order; each series was presented four times. Partial results indicate that both detection and recognition declined at distances greater than approximately 200 m and distance estimation was grossly underestimated by listeners. Specific results will be presented.

2pPP22. Comparing some convolution-based methods for creation of surround sound. Andrzej Czyzewski and Bozena Kostek (Sound & Vision Eng. Dept., Gdansk Univ. of Technol., Narutowicza 11/12, 80-952 Gdansk, Poland)

Spatialization of the sound using the multichannel techniques is now getting widespread. One can derive many rules for surround sound recording and reproduction. However, there exists only few methods suitable for recording sound in large auditoria ensuring its proper subsequent reproduction in small reproduction rooms, preserving spatial properties of sound acquired in the original recording location. Some experiments presented in the paper were devoted to the simulation of acoustics of the recording hall using the convolution of a monophonic audio signal with the multichannel impulse response of the hall. A special microphone setup was created for this task and an original method of recording multichannel impulse response of auditory halls was conceived and implemented. In this method the acoustical signal recorded quasianechoically was convolved with five impulse responses of the simulated room measured in the room corners and on the stage. Another examined method which is more standard employed convolution of monophonic signals with long-term averaged HRTFs (Head-Related Transfer Funtions). Surround recordings made with both mentioned convolution techniques were then compared on the basis of subjective testing results. The details of the examined surround recording methods and results of their assessments will be discussed in the paper. [Work supported by KBN, Grant No. 8 T11D 00218.]

2pPP23. Auscultation in noise: A program to develop a stethoscope capable of functioning during aeromedical transport. Paul A. Cain, William A. Ahroon (US Army Aeromedical Res. Lab., Bldg. 6901, Ft. Rucker, AL 36362), John M. Sewell, and William N. Bernhard (Active Signal Technologies, Linthicum Heights, MD)

U.S. Army helicopters used to provide aeromedical transport are extremely noisy (105 dB). This prevents auscultation with current stethoscopes. Ten subjects used four different stethoscopes in noise at 70–100 dB in order to determine the detection threshold of a body sound. The stethoscopes differed in performance ($p=0.001$); the best being the acoustic, followed by the three electronic stethoscopes. The threshold of noise for the detection of heart and breath sounds was 80 dB and 70–75 dB, respectively. A comparison of a standard stethoscope and one modified with Communications Ear Plugs revealed a difference ($p=0.05$) at 70 and 80 dB that disappeared at higher noise levels, implying that it was not simple masking, but that noise was amplified after entering the sensing head. Despite the poorer performance of the electronic stethoscopes, the need for a gain of 30–35 dB indicated that this was the preferred route for development. A sensor with an acoustic impedance match close to body tissue, but poorly matched to air-coupled noise, was developed, and trials indicate that heart sounds can be heard at 105 dB. Development continues with the aim of reducing or eliminating the need for signal processing.

Session 2pSAa

Structural Acoustics and Vibration: Vibration of Floors of Buildings

Eric E. Ungar, Chair
Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138-1118

Chair's Introduction—1:00

Invited Papers

1:05

2pSAa1. The art of building floor vibration evaluation. Thomas M. Murray (Dept. of Civil and Environ. Eng., Virginia Tech, Blacksburg, VA 24061)

Annoying floor vibration caused by building occupant activity is an increasingly common occurrence. Optimized floor systems and the advent of the electronic office are the main causes. Because humans are very sensitive to vertical movement, especially in quiet environments, careful floor design is required. When this is not done or when the occupancy does not meet the design criteria, complaints are commonly received. The complex nature of human excitations (walking, jumping, running, exercising) makes the evaluation of problem floors as much art as science. Typical floor motions include accelerations caused by frequencies above the human threshold limit, which must be considered in any evaluation. Further, most floor systems exhibit several closely spaced modes, which make remedial measures difficult to implement. Structural modifications, passive control in the form of tuned mass dampers, and active control are all potential remedies but a choice is highly dependent on the particulars of the problem floor and the occupancy. This paper discusses techniques for evaluating human induced floor motion and gives examples of successful and unsuccessful retrofits.

1:35

2pSAa2. Diagnosing a case of occupant-induced whole building vibration. Linda M. Hanagan (Dept. of Architectural Eng., Penn State Univ., 104 Eng. Unit A, University Park, PA 16802)

People in the tenth floor office suite of a 10-story building were complaining of annoying floor vibrations. These vibrations were worse on some days than on others and seemed to be emanating from a dance studio on the floor below. The building owners wanted the problem fixed; however, the exact mode of transmission to the tenth floor was, as yet, unknown. Understanding how the vibration was being transmitted was essential to developing a repair solution. Among the possibilities for transmission were a full height partition on the ninth floor, the curtain wall, column flexure, and column shortening. Through vibration testing, it was determined that a whole building mode, with an estimated equivalent mass of almost 1 000 000 kg, was excited to cause disturbing levels of vibration at the tenth floor with as few as six people jumping on the floor below. Details of the vibration testing are provided.

2:05

2pSAa3. Estimation of vibrations due to walking on floors that support sensitive equipment. Eric E. Ungar and Jeffrey A. Zapfe (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138-1118)

The development of the extensively used simple method for predicting footfall-induced vibrations of floors of buildings (American Institute of Steel Construction Design Guide 11) is reviewed. The empirical basis and underlying analytical assumptions of the method are delineated and critiqued. Its limitations are discussed and suggestions for its extension are presented. Results of some recent related measurements are summarized.

2:35

2pSAa4. Vibrations of raised access floors. Hal Amick, Michael Gendreau, and Colin G. Gordon (Colin Gordon and Assoc., 883 Sneath Ln., Ste. 150, San Bruno, CA 94066)

Raised access floors play a critical role in modern cleanroom design. They have unique mechanical properties that make them respond to dynamic loading in a manner quite different from conventional floors. For example, an unbraced floor is much more flexible horizontally than in the vertical direction. Horizontal vibration amplitudes with walker excitation may exceed 100 μm/s in an unbraced floor, exceeding the sensitivity of 1000$\times$ inspection microscopes by as much as an order of magnitude. Issues such as these become important when moderately vibration-sensitive instruments, such as optical microscopes, are supported on access floors, typically the case in cleanrooms. This paper presents results of experimental studies involving a 3 m$\times$3 m segment of floor and a large floor installed in a cleanroom, both of which were subjected to dynamic loads using a shaker. Both drive-point and propagation properties were examined. In addition, data are presented for variations in bracing and bolting using the 3 m$\times$3 m segment.

3:05

2pSAa5. Vibration sensitivity of a laboratory bench microscope. Hal Amick (Colin Gordon and Assoc., 883 Sneath Ln., Ste. 150, San Bruno, CA 94066) and Matthew Stead (Bassett Acoustics, Kent Town SA 5067, Australia)

Bench-mounted optical microscopes have a wide variety of applications in science and technology. The vibration sensitivity is a function of both magnification and support conditions. In this paper we present the results of experimental studies addressing vibration sensitivity as well as the amplification and attenuation provided by typical laboratory casework. The benchtop vibration amplitudes at which the effects of motion first become perceptible were found for magnifications of 40×, 100×, 400×, and 1000× using sinusoidal excitation. Frequency response functions were determined for benchtop motion with respect to floor motion, using both conventional casework and a popular pneumatic isolation bench. Floor vibration criteria were developed for microscopes with the two types of support.

TUESDAY AFTERNOON, 3 DECEMBER 2002 CORAL GALLERY FOYER, 3:50 TO 4:50 P.M.

Session 2pSAb

Structural Acoustics and Vibration: Vibration Abatement

Jeffrey S. Vipperman, Chair
Department of Mechanical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261

Contributed Papers

3:50

2pSAb1. Vibration analysis and building vibration isolation design for an automated people mover system in airport terminal structures. James E. Phillips (Wilson, Ihrig & Assoc., Inc., 5776 Broadway, Oakland, CA 94618)

Detailed Finite Element Analysis (FEA) models were developed for a proposed airport terminal expansion project. An Automated People Mover (APM) system is incorporated into the airport structures for shuttling passengers quickly between terminals. The dynamic forces imparted onto the structures by the moving APM vehicles and the analysis approach were based upon established techniques developed for addressing ground-borne and structure-borne vibrations from rail systems. Measurements were conducted at two other major airports with existing rubber tire APM systems on aerial structures. These measurements provided baseline vibration levels for the analysis as well as the forces imparted to the structure by the APM vehicles. The results of the analyses were utilized in the design of a vibration isolation system included in the structural design of a guideway within a terminal building in order to mitigate structure-borne vibration from APM operations.

4:05

2pSAb2. Helmholtz design for noise transmission attenuation on a chamber core composite cylinder. Deyu Li and Jeffrey S. Vipperman (Dept. of Mech. Eng., Univ. of Pittsburgh, 531 Benedum Hall, Pittsburgh, PA 15261)

This work explores the feasibility of using Helmholtz resonators to attenuate a subscale ChamberCore cylinder noise transmission. The ChamberCore cylindrical composite is an innovative new sandwich-type structure. It consists of an outer skin, an inner skin, and linking ribs. There are wedge-cross-section chambers along the axis direction between the outer and inner skins. These chambers provide a potential for the acoustic Helmholtz resonator design in order to reduce the noise transmission, which is dominated by the internal acoustic cavity. In this experimental work, the sound transmission behavior of the ChamberCore fairing is investigated and divided into four interesting frequency regions: the stiffness-controlled zone, cavity resonance-controlled zone, coincidence-controlled zone, and mass-controlled zone. It is found that the noise transmission in the low-frequency band is controlled by the structural stiffness and cavity resonances, where the acoustic Helmholtz design method has the potential to improve the noise transmission.

4:20

2pSAb3. Evidence of the existence of phononic band gaps: A practical example of a tunable sound insulation by a periodic device of rods. Cecile Goffaux, Philippe Lambin, Jean-Pol Vigneron (LPS, Facultés Universitaires Notre-Dame de la Paix, 61 rue de Bruxelles, B-5000 Namur, Belgium, cecile.goffaux@scf.fundp.ac.be), and Fabrizio Maseri (Ctr. de Conception Solutions Acier pour la Construction, B-4000 Liege, Belgium)

Based on the principle of phononic band gap materials, the control of acoustic frequency gaps by altering the geometry of the system is analyzed in the particular case of a set of parallel solid square-section columns distributed in air on a square lattice. This system is shown to be sensitive enough to the rotation of the columns to be considered for a practical sonic band gap width engineering. For different geometric configurations, experimental and theoretical results are presented and a discussion about the application of such structures as sound insulators is discussed. We acknowledge the use of the Namur Scientific Computing Facility (Namur-SCF), a common project between FNRS, IBM Belgium, and the Facultés Universitaires Notre-Dame de la Paix (FUNDP). [C. Goffaux acknowledges the financial support of the FIRST program of the Walloon Region Government and the Research and Development Centre of Cockerill-Sambre (ARCELOR Group).]

4:35

2pSAb4. Physics of a pneumatic vibration isolator revisited. Vyacheslav M. Ryaboy (Newport Corp., 1791 Deere Ave., Irvine, CA 92606, vryaboy@newport.com)

This paper gives detailed consideration to two phenomena that affect the performance of a pneumatic isolator but have not received adequate attention in the past. The first one is the thermal conductivity of gas. It is usually assumed that the compression of gas in the pneumatic chamber is an adiabatic process. Thermal conductivity can, nevertheless, affect the performance, especially in the low-frequency domain and for small-size isolators. A simple explicit expression for the acoustic compliance is derived from exact solution of conductive gas equations in the cylindrical domain. Another factor affecting the isolator response is the stiffness introduced by the diaphragm. A formula for the stiffness is derived based on the mathematical model of the diaphragm as an elastic membrane. Comparison to experimental data shows that adequate representation of these two factors results in accurate predictions of the isolator performance as a function of the supported load.

Session 2pSP

Signal Processing in Acoustics and Speech Communication: Feature Extraction and Models for Speech

Jose A. Diaz, Cochair
Universidad de Carobobo, Valencia, Venezuela

Shrikanth Narayanan, Cochair
Signal and Image Processing Institute, University of Southern California, 3740 McClintock Avenue, Los Angeles, California 90089-2564

Chair's Introduction—1:00

Invited Papers

1:05

2pSP1. Feature extraction and models for speech: An overview. Manfred Schroeder (Univ. of Goettingen, Buergerstrasse 42–44, 37073 Goettingen, Germany)

Modeling of speech has a long history, beginning with Count von Kempelens 1770 mechanical speaking machine. Even then human vowel production was seen as resulting from a source (the vocal chords) driving a physically separate resonator (the vocal tract). Homer Dudley's 1928 frequency-channel vocoder and many of its descendants are based on the same successful source-filter paradigm. For linguistic studies as well as practical applications in speech recognition, compression, and synthesis (see M. R. Schroeder, Computer Speech), the extant models require the (often difficult) extraction of numerous parameters such as the fundamental and formant frequencies and various linguistic distinctive features. Some of these difficulties were obviated by the introduction of linear predictive coding (LPC) in 1967 in which the filter part is an all-pole filter, reflecting the fact that for non-nasalized vowels the vocal tract is well approximated by an all-pole transfer function. In the now ubiquitous code-excited linear prediction (CELP), the source-part is replaced by a code book which (together with a perceptual error criterion) permits speech compression to very low bit rates at high speech quality for the Internet and cell phones.

1:30

2pSP2. New pitch detection system for concurrent speech on a parallel processing system. Yoshifumi Chisaki (Dept. of Computer Sci., Faculty of Eng., Kumamoto Univ., 2-39-1 Kurokami, Kumamoto 860-8555, Japan), Akira Nakagawa (Kumamoto Univ., Kumamoto, Japan), Hidetoshi Nakashima (Kumamoto Natl. College of Technol.), Tsuyoshi Usagawa (Kumamoto Univ., Kumamoto, Japan), and Masanao Ebata (Kumamoto Natl. College of Technol., Japan)

Pitch detection system based on the harmonic wavelet transform (HWT) algorithm for concurrent speech [Y. Chisaki, 6C.15, Proc. ICA2001] has been proposed. The HWT algorithm has two major advantages. One is to detect a pitch from a frame signal, which is a short duration, such as 20 ms. The other one is robustness against noise; namely, all pitches can be detected from concurrent speech in a short delay, simultaneously. However, processing time for the modified correlation method, as a conventional method is 8.6 times faster than that for the HWT based system. Due to the development of computers, processing time can be reduced easily by using PCs. This paper proposes a new pitch detection system for concurrent speech with a new parallel algorithm. The system is based on the HWT method, and implemented on a parallel processing system. Simulations are performed with some node arrangements. As a result, a processing speed for a parallelized block of the system is 6.3 times faster as that for the original one without being accompanied by degradation of pitch detection accuracy. [Part of this work is supported by The Sagawa Foundation for Promotion of Frontier Science, the Cooperative Research Project Program of the RIEC, Tohoku Univ., and the Ono Acoustics Research Fund (2001).]

1:55

2pSP3. Methodology for analyzing spectral differences. Howard Rothman (Dept. of Commun. Sci. and Disord., Univ. of Florida, P.O. Box 117420, Gainesville, FL 32611) and Jose Diaz (Universidad de Carabobo, Valencia, Venezuela)

The human voice can be an exceptional instrument. The larynx contributes significantly to making up the exceptional voice: it provides the speaking/singing fundamental frequency (SFF) and the resultant harmonic structure. The vocal tract modulates energy and enhances the contributions of the larynx. There is evidence that the singers/speakers formant provides brilliance and increased amplitude to the voice. Several papers have examined the effects of the spectrum on the perception of the aging voice, the charismatic voice and the professional voice. Spectral differences can be seen in voices that have been identified as good/bad, young/old, brilliant/ordinary. One of the authors presented data differentiating between good and bad vibrato samples that aid in identifying singers that are experiencing vocal difficulty. However, there are singers, identified by vocal pedagogues, critics and other singers as

experiencing vocal difficulty, whose vibrato remained unchanged over time. While changes in vibrato are most probably laryngeal in nature, there are no data available as to the changes in the spectra of these singers that provide the perceptual cue for their vocal problems. This paper will present methods detailing the spectral differences and their contributions to perceptual judgments.

2:20

2pSP4. Novel features for robust speech recognition. Alexandros Potamianos (Bell Labs, Lucent Technologies, 600 Mountain Ave., Murray Hill, NJ 07974)

Recently there has been much research in the area of robust front-ends and new features for automatic speech recongition (ASR). These efforts have had limited success for certain databases and recording conditions. In this work, we review some recent work on features for ASR: the articulatory front-end of Li *et al.* (2000), nonlinear modulation features [Quatieri (2002); Dimitriadis and Maragos (2002)], chaotic features [Pitsikalis and Maragos (2002)], short-time spectral moments [Paliwal *et al.* (2000)], etc. We extend the work of Potamianos and Maragos (2001) to show how some of these features relate to the standard front-end of short-time smooth spectral envelope. We also analyze some of these new features using classification and regression trees to show the relevance of the features for phone-classification task.

2:45–3:00 Break

3:00

2pSP5. Convolutive mixture separation in time–frequency domain for robust automatic speech recognition. Shubha L. Kadambe (HRL Labs., LLC, 3011 Malibu Canyon Rd., Malibu, CA 90265)

In a mobile environment, automatic speech recognition (ASR) systems are being used for information retrieval. Due to the presence of multiple speakers, noise sources and reverberations in such environments, the ASR performance degrades. Here, the problem of improvement of ASR performance by separating the convolutively mixed speech signals that predominantly exist in mobile environments is addressed. For the separation, an extension of the algorithm published in A. Ossadtchi and S. Kadambe [''Over-complete blind source separation by applying sparse decomposition and information theoretic based probabilistic approach,'' ICASSP, 2001] is applied in the time–frequency domain. In the extended algorithm, the dual update algorithm that minimizes L1 and L2 norms simultaneously is applied in every frequency band. The problem of channel swapping is also addressed. The experimental results of separation of convolutively mixed signals indicate about 6-dB SNR improvement. The enhanced speech signals are then used in GMM based continuous speech recognizer. The recognition experiments are performed using the real speech data collected inside a vehicle. During the presentation, complete ASR performance improvement results will be provided.

3:25

2pSP6. Linguistically informed automatic speech recognition. Carol Y. Espy-Wilson (Dept. of Elec. Eng., Univ. of Maryland, A. V. Williams Bldg., Cambridge, MA 20742)

The development of an automatic event-based speech recognition system (EBS) that relies heavily on acoustic phonetics (to guide the recognition process and to extract relevant information) and combines a phonetic-feature hierarchy with a uniform statistical framework (at present, Support Vector Machines) to provide adaptability and flexibility is currently under way. This recognition framework allows for easy assessment and distinction of the performances of the acoustic parameters versus that of the pattern recognizer. The overall structure of EBS involves (1) landmark detection based on acoustic parameters that are related to the source and manner-of-articulation phonetic features and (2) use of the landmarks in the extraction of other acoustic parameters related to the place-of-articulation phonetic features. This talk will focus on the development of the acoustic parameters and the need for relative parameters for speaker independence, multi-time-scale processing to capture the dynamics of phonetic segments and extensive evaluation of the parameters to hone in on direct measures of the relevant acoustic properties. [Work supported by NSF and NIH.]

3:50

2pSP7. Graphical models and automatic speech recognition. Jeff A. Bilmes (Univ. of Washington, Seattle, WA 98195-2500)

Graphical models (GMs) are a flexible statistical abstraction that has been successfully used to describe problems in a variety of different domains. Commonly used for ASR, hidden Markov models are only one example of the large space of models constituting GMs. Therefore, GMs are useful to understand existing ASR approaches and also offer a promising path towards novel techniques. In this work, several such ways are described including (1) using both directed and undirected GMs to represent sparse Gaussian and conditional Gaussian distributions, (2) GMs for representing information fusion and classifier combination, (3) GMs for representing hidden articulatory information in a speech signal, (4) structural discriminability where the graph structure itself is discriminative, and

the difficulties that arise when learning discriminative structure (5) switching graph structures, where the graph may change dynamically, and (6) language modeling. The graphical model toolkit (GMTK), a software system for general graphical-model based speech recognition and time series analysis, will also be described, including a number of GMTK's features that are specifically geared to ASR.

Contributed Papers

4:15

2pSP8. Fundamental frequency estimation using signal embedding in state space. Dmitry Terez (SoundMath Technologies LLC, 6 N. 9th St., Millville, NJ 08332, terez@soundmathtech.com)

A new robust nonlinear method for determination of fundamental frequency ($F0$) was recently proposed with application to speech pitch detection [Terez, Proc. ICASSP **1**, 345–348 (2002)]. The method uses state-space embedding technique originally introduced for analyzing chaotic signals. The new method has been generalized and tested on different types of speech signals, as well as on a variety of other acoustic signals. In addition, some artificially generated nonstationary and complex wave forms have been used to test the limits of the method in comparison with other known (short-term) $F0$-estimation techniques (e.g., correlation, spectrum or cepstrum-based methods). Evaluation results demonstrate a unique combination of properties distinguishing the new method from conventional techniques. In particular, reliable and accurate $F0$ estimates can be obtained for clean periodic signals using signal segments slightly longer than one complete fundamental period. Other properties include immunity to speech formants and robust performance on noisy and band-limited speech signals. Some improvements are introduced to reduce the number of required computations and to achieve higher (subsample) accuracy. The method waused to implement a robust pitch-tracking algorithm for speech processing applications. Further information and demo software can be found at http://www.soundmathtech.com/pitch.

4:30

2pSP9. Human factor cepstral coefficients. Mark D. Skowronski and John G. Harris (Computational Neuro-Eng. Lab., Univ. of Florida, Gainesville, FL 32611, markskow@cnel.ufl.edu)

Automatic speech recognition (ASR) is an emerging field with the goal of creating a more natural man/machine interface. The single largest obstacle to widespread use of ASR technology is robustness to noise. Since human speech recognition greatly outperforms current ASR systems in noisy environments, ASR systems seek to improve noise robustness by drawing on biological inspiration. Most ASR front ends employ mel frequency cepstral coefficients (mfcc) which is a filter bank-based algorithm whose filters are spaced on a linear-log frequency scale. Although center frequency is based on a perceptually motivated frequency scale, *filter bandwidth is set by filter spacing and not through biological motivation*. The coupling of filter bandwidth to other filter bank parameters (frequency range, number of filters) has led to variations of the original algorithm with different filter bandwidths. In this work, a novel extension to mfcc is introduced which decouples filter bandwidth from the rest of the filter bank parameters by employing the relationship between filter center frequency and critical bandwidth of the human auditory system. The new algorithm, called human factor cepstral coefficients (hfcc), is shown to outperform the original mfcc and two popular variations in several ASR experiments and noise sources.

4:45

2pSP10. Vowel landmark detection. Andrew W. Howitt (Res. Lab. of Electron., MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Landmark based speech processing is a component of Lexical Access From Features (LAFF), a novel paradigm for feature based speech recognition. Detection and classification of landmarks is a crucial first step in a LAFF system. Vowel landmarks are detected using an existing syllabic segmentation algorithm with several novel extensions that incorporate durational information, absolute energy level, and $F1$ track information. The detector is scored against the TIMIT database, using a novel algorithm to convert the segmental transcriptions to a landmark representation for scoring. Previous experiments have validated the predictions of acoustic theory, specifically the presence of $F1$ peak in vowels, and demonstrated that amplitude peak in a fixed frequency band is practically as good as a formant tracker. Substantial improvements in performance are achieved by optimizing the fixed frequency band to the $F1$ range, and use of a trainable neural network to combine the multiple acoustic cues. The neural network can also be used to generate confidence scores for detected landmarks, which provide vital information for later stages of processing.

5:00

2pSP11. A Gaussian-selection-based preclassifier for speaker identification. Marie A. Roch (Dept. of Computer Sci., San Diego State Univ., San Diego, CA 92182-7720)

For practical reasons driven by the need to periodically adapt speaker models or enroll new members of the population, most speaker identification systems train individual models for each speaker. When classifying a speech token, the token is scored against each model and the maximum *a posteriori* decision rule is used to decide the classification label. Consequently, the cost of classification grows linearly for each token as the population size grows. When considering that the number of tokens to classify is also likely to grow linearly with the population, the total work load increases exponentially. In this work, a new system is presented which builds upon the so-called ''Gaussian selection'' techniques. The system uses the speaker-specific models as source data and constructs N-best hypotheses of speaker identity. The N-best hypothesis set is then evaluated using individual speaker models. This process results in an overall reduction of workload. The cost of the model generation is low enough to permit enrollment and adaptation, and the accuracy of the preclassifier is such that there is minimal impact on the recognition rate.

5:15

2pSP12. Quality of service (QoS) on public telephonic networks for multimedia transmission systems. Salvador Alvarez-Ballesteros and Miguel Alvarez-Rangel (IPN–ESIME, Zacatenco Edif. #1 Col., Lindavista, Mexico)

The object of this paper is to determine which QoS maximum can be obtained by applying commercial telecommunications technology in the transmission of audio and video on public telephone lines and the Internet, procuring to achieve the necessary quality to reach the conditions required on H323. It is certain that the top technology that allows to assure that the QoS is most demanding, are a reality. However, these technologies alone pay off practically under conditions that are not very feasible for formalizing would be of the ambient control of the laboratory. In this work we analyzed the technical problems of the telecommunications that affect the yield of the links for applications of multimedia on the Internet.

5:30

2pSP13. The challenges of archiving networked-based multimedia performances (Performance cryogenics). Elizabeth Cohen (Cohen Acoustical, Inc., 132 S. Lucerne Blvd., Los Angeles, CA 90004-3725 and CCRMA/Dept. of Elec. Eng., Information Systems Lab., Stanford Univ., CA, akustik@attbi.com), Jeremy Cooperstock (McGill Univ., Montreal, Canada), and Chris Kyriakakis (Univ. of Southern California)

Music archives and libraries have cultural preservation at the core of their charters. New forms of art often race ahead of the preservation infrastructure. The ability to stream multiple synchronized ultra-low latency streams of audio and video across a continent for a distributed interactive performance such as music and dance with high-definition video and multichannel audio raises a series of challenges for the architects of digital libraries and those responsible for cultural preservation. The archiving of such performances presents numerous challenges that go beyond simply recording each stream. Case studies of storage and subsequent retrieval issues for Internet2 collaborative performances are discussed. The development of shared reality and immersive environments generate issues about, What constitutes an archived performance that occurs across a network (in multiple spaces over time)? What are the families of necessary metadata to reconstruct this virtual world in another venue or era? For example, if the network exhibited changes in latency the performers most likely adapted. In a future recreation, the latency will most likely be completely different. We discuss the parameters of immersive environment acquisition and rendering, network architectures, software architecture, musical/choreographic scores, and environmental acoustics that must be considered to address this problem.

TUESDAY AFTERNOON, 3 DECEMBER 2002 GRAND CORAL 1, 1:00 TO 5:35 P.M.

Session 2pUW

Underwater Acoustics: Geoclutter and Boundary Characterization II

Purnima Ratilal, Cochair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 5-435, Cambridge, Massachusetts 02139

Nicholas C. Makris, Cochair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 5-204, Cambridge, Massachusetts 02139

Chair's Introduction—1:00

Contributed Papers

1:05

2pUW1. The Geoclutter Experiment 2001: Remote acoustic imaging of sub-bottom and seafloor geomorphology in continental shelf waters. Nicholas C. Makris, Purnima Ratilal, Yisan Lai, Deanelle T. Symonds (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, makris@mit.edu), Lilimar A. Ruhlmann, and Edward K. Scheer (MIT, Cambridge, MA 02139)

In the Geoclutter experiment of April–May 2001, an active sonar system was used to remotely and rapidly image geomorphology over wide areas in continental shelf waters by long-range echo sounding. The bistatic system, deployed in the strataform area south of Long Island, imaged extensive networks of buried river channels and inclined subseafloor strata over tens of kilometers in near real time. Bathymetric relief in the strataform area is extremely benign. The vast majority of features imaged apparently correspond to sub-bottom geomorphology that sound waves reach after tunneling as well as propagating through the overlying sediment. Returns from buried river channels were often found to be as discrete and strong as those from calibrated targets placed in the water column. Since buried river channels are expected to be ubiquitous in continental shelf environments, sub-seafloor geomorphology will play a major role in producing ''false alarms'' or clutter in long-range sonar systems that search for submerged objects such as underwater vehicles or marine mammals. Wave guide scattering and propagation are inherent to this new remote sensing technology because source signals are transmitted over hundreds of water-column depths in range to image sub-seafloor and seafloor geomorphology.

1:20

2pUW2. Coherent versus diffuse surface and volume Reverberation in an ocean wave guide: Reverberation rings, modal decoupling, and possible fish scattering in Geoclutter 2001. Purnima Ratilal and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, makris@mit.edu)

Conditions necessary for the field scattered from volume inhomogeneities and surface roughness in a wave guide to become diffuse in nonforward azimuths are derived from Green's theorem. Diffuse scattering leads to a decorrelation of the wave guide modes that simplifies the expression for the intensity of the total scattered field. When diffuse scattering conditions are not satisfied, seafloor scattering becomes coherent and leads to the formation of ''reverberation rings'' sometimes observed in high-resolution sonar systems. The formation of these rings will be investigated both analytically and with simulations. The possibility that some prominent clutter events in the Geoclutter 2001 experiment are due to diffuse scattering from fish schools is investigated theoretically and experimentally.

1:35

2pUW3. Fish schools as potential clutter and false targets: Observations on the New Jersey shelf. Redwood W. Nero, Charles H. Thompson (Naval Res. Lab., Stennis Space Center, MS 39529-5004), and Richard H. Love (BayouAcoust., Pass Christian, MS 39571-2111)

Fish schools can appear as clutter or false targets on search sonars and can confuse the interpretation of scattering from the sea floor. During Boundary Characterization 2001, a high-frequency echosounder was used

to quantify fish schools in an effort to provide estimates of false targets and clutter at low frequency. Schools were quantified from the echosounder data using an image processing algorithm designed to provide estimates of school size, acoustic intensity, and a variety of diagnostic features. The number of schools that had the potential to be low-frequency false targets was estimated using information on fish species obtained from fisheries research trawls and an NRL swimbladder scattering model. A few pelagic fish schools of intermediate size and high intensity, with low-frequency target strengths estimated at +12 dB, occurred near the sea surface at night in the northern corner of the study site, at a density of about one per km^2. These schools were most likely to appear as false targets. Demersal fish, those near the sea floor, although abundant along the 80-m contour, were not likely to be strong false targets at low frequency. [Work supported by ONR.]

1:50

2pUW4. Model for coherent scattering from a network of buried river channels in a stratified ocean waveguide. Sunwoong Lee, Purnima Ratilal, and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, makris@mit.edu)

A theoretical model of three-dimensional bistatic scattering from a network of buried river channels in a continental shelf waveguide is presented. The scattered field on the river channel walls is estimated to be locally specular using the Kirchhoff approximation. This field is then propagated out to a distant receiving array using Green's theorem and the waveguide Green function, so that it is also valid for scattering in the near field of the river channel network. This model is applied to calculate the scattered and beamformed field from a network of buried river channels in the East Coast strataform area, whose geomorphology was mapped out by numerous geophysical surveys. The modeled scattered field is found to be dependent upon the bistatic orientation and projected area of the river channel walls relative to the source and receiver. To verify the modeling, the result of the simulation is compared with data from the Geoclutter 2001 acoustic experiment where images of buried river channel networks were acquired from long-range using a bistatic sonar system.

2:05

2pUW5. Evidence for single-mode propagation and its implications for range-dependent geoacoustic inversion in the SWAT experiment. George V. Frisk (Appl. Ocean Phys. & Eng. Dept., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, gfrisk@whoi.edu) and Kyle M. Becker (Penn State Univ., State College, PA 16804)

During the Shallow Water Acoustic Technology (SWAT) experiments, which were conducted in October 2000 on the New Jersey Shelf, a 20 Hz pure tone was transmitted from a source in 73 m of water to two drifting receivers. The behavior of the pressure magnitude and phase out to ranges of about 8 km suggest that the field is dominated by a single normal mode. In particular, an adiabatic modal interpretation of the phase variation with range is consistent with an independent autoregressive spectral estimate of the wave number content of the field. It is then shown, within the context of an isovelocity bottom model with varying water depth, that the measured range-varying bathymetry does not account for the range dependence in the single modal eigenvalue. On the other hand, an alternative interpretation of the data in terms of a wave guide with constant depth and range-varying sound speed in the bottom yields reasonable estimates of the lateral seabed variability. [Work supported by ONR.]

2:20

2pUW6. Effects of measured bathymetry and subbottom variability on low-frequency shallow-water reverberation. Altan Turgut and Roger Gauss (Naval Res. Lab., Acoust. Div., Washington, DC 20375)

A pseudo-spectral numerical method is used to study the effects of bathymetry and subbottom variability on shallow-water reverberation at low frequencies (less than 500 Hz). Bathymetry and subbottom variability at New Jersey Shelf Geoclutter and Boundary Characterization 2001 experimental sites were measured by using chirp sonar and sediment core data at different wavelength scales. A stochastic bottom simulator was incorporated with the measured deterministic bottom/subbottom features, including dipping layers, erosional channels, and buried river channels. Three-dimensional simulations were also performed to study the effects of anisotropic bottom/subbottom variability on the shallow-water reverberation. It has been shown that, for the Geoclutter and Boundary Characterization 2001 experimental sites, predicted anisotropy in the acoustic scattering field is mainly due to the deterministic features rather than due to anistropic small-scale variabilities in the sediment. [Work supported by ONR.]

2:35

2pUW7. Analysis of the effects of sea floor and sub-sea floor geology on underwater sound propagation estimation and performance prediction. Richard Katz and Michael Sundvik (NUWC Newport, 1176 Howell St., Newport, RI 02841)

In this study, we investigate shallow water sound propagation measurements and interactions with the seabed as a function of geologic composition and waveform characterization. We have analyzed sound transmission loss measurements in the shallow acoustic channel from the geoclutter experiments conducted in 2001 and 2002 and found them to be in good agreement with standard modeling results. The sound propagation through the acoustic channel and interactions with the boundaries lead to energy spreading losses, which make acoustic performance prediction and modeling uncertain. We quantify these losses to the extent possible based on the signals transmitted during the experiments. The contributions to and effects on sound propagation variability from sound speed profile (degree of downward refraction) are discussed with emphasis placed on interactions with the bottom and sub-bottom geology. Analysis is conducted for signals in the 400 Hz to 4 kHz range.

2:50

2pUW8. Range-dependent modal eigenvalue estimation in the SWAT experiment. Luiz L. Souza (MIT/WHOI Joint Prog. in Oceanogr./Appl. Ocean Sci. and Eng., 77 Massachusetts Ave., Rm. 5-435, Cambridge, MA 02139), George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Kyle M. Becker (Penn State Univ., State College, PA 16804)

Three modal mapping experiments (MOMAX) have been conducted in shallow water environments: two in the East coast STRATAFORM site off the New Jersey coast; and one in the Gulf of Mexico off the Florida coast. A low-frequency source emits a set of CW signals in the range 20–500 Hz, and the field is measured by drifting buoys. The ultimate goal of these experiments is to invert the acoustic field data in water for the sound speed profile in the seabed. The analysis has focused on measuring the modal characteristic wavenumbers (eigenvalues) as a function of position. In general, one has to resort to range-dependent spectral analysis. Techniques based on both the short-term Fourier transform and high-resolution parametric spectrum estimation have been used for measuring eigenvalues. New results from the year 2000 SWAT (MOMAX III) experiment at the STRATAFORM site will be discussed. These include a comparison of data obtained from tracks along and across the New Jersey shelf. [Work supported by ONR and WHOI Academic Programs Office.]

3:05–3:20 Break

3:20

2pUW9. Geoacoustic measurements on the New Jersey shelf. Charles Holland (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

Two key acoustic parameters for predicting shallow water reverberation are bottom reflection and bottom scattering strength. Historically, seabed reflection and scattering have been treated in a disconnected fashion both in measurement and analysis. A coupled approach to reflection and scattering was conducted during the GeoClutter and Boundary2001 Experiments. By coupled, it is meant that the reflection and scattering measurements are not only colocated, but sample commensurate spatial scales of the seabed. It is also meant that the processing and analyses employ commensurate assumptions and modeling. The coupled nature of the measurements provides the opportunity for determining the physical mechanisms responsible for the reflection and scattering. Deterministic and stochastic geoacoustic properties obtained from an analysis of these mechanisms show very high spatial variability across the measurement area. [Work supported by the NATO SACLANT Undersea Research Center and ONR.]

3:35

2pUW10. Statistical characterization of geologic clutter observed on the STRATAFORM. John R. Preston and Douglas A. Abraham (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804)

Together with SACLANTCEN, the authors recently participated in the 2001 Geoclutter Experiment to study shallow water bottom reverberation in the STRATAFORM off New Jersey. Sources were bistatic and monostatic coherent pulses near 400 Hz. The receivers were horizontal arrays. The STRATAFORM is known to have benign surface morphology but contains many buried river channels and other subsurface horizons. Some highlights of the reverberant returns are discussed that include the correlation of returns with the buried river channels and some other subbottom features. The main objective of this study is the statistical characterization of the geologic clutter. The K distribution has been shown to be useful in describing non-Rayleigh behavior. The clutter from STRATAFORM is described by its K-distribution parameters as a function of location and sediment type. [Work supported by ONR Code 32, Grant N00014-97-1-1034.]

3:50

2pUW11. Geoacoustic inversion of ambient noise. Chris H. Harrison (SACLANT Undersea Res. Ctr., Viale San Bartolomeo, 400, 19138 La Spezia, Italy, harrison@saclantc.nato.int)

The vertical directionality of ambient noise is strongly influenced by seabed reflections. Therefore, geoacoustic parameters can be inferred by inversion of the noise. In this approach, using vertical array measurements (16 m aperture), the reflection loss is found directly by comparing the upward with the downward going noise. Theory suggests that this simple ratio is, in fact, the power reflection coefficient (a function of angle and frequency). A layer model and a search is required to find geoacoustic parameters, but no such model or search is required for reflection loss alone. Experimental data have been gathered at 11 sites, 10 in the Mediterranean, and 1 on the New Jersey Shelf during BOUNDARY2001. Site to site variations are discussed, and comparisons are made with simple layer models. Usually the vertical array is bottom-moored, but at two of the Mediterranean sites the array was allowed to drift over a few miles so that bottom properties could be surveyed. Bottom variations could indeed be seen. Simulations using the parabolic equation to mimic spatial variations of bottom properties confirm that the bottom loss inferred from this type of measurement is indeed a local one rather than a regional average.

4:05

2pUW12. Geoacoustic inversion using genetic algorithm. Itaru Morishita, Hirotaka Murakami (Oki Electric Industry Co., Ltd., 4-10-3, Shibaura, Minato-ku, Tokyo 108-8551, Japan, morishita555@oki.com), Kazuhiko Ohta, Kouki Okabe, and Masamichi Oikawa (TRDI, Yokosuka 239-0826, Japan)

Geoacoustic parameters of bottom sediment have a great influence on acoustic propagation in a water column. In order to estimate the parameters, the genetic algorithm (GA) was applied to the low frequency acoustic data obtained at the SWAT (Shallow Water Acoustic Technology) experiments which were conducted at the East China Sea and the New Jersey Shelf. In the experiments, complex acoustic pressure were measured while distance between the source, which transmitted the cw signal, and the receiver was changed. Horizontal wave number spectrum were obtained from the measured data by synthetic aperture process and the Hankel transform. For estimating geoacoustic parameters, an error function was defined as the difference between the measured spectrum and a calculated one which is obtained by applying the Hankel transform to calculated complex pressure, and GA searched geoacoustic parameters which minimize the error. To evaluate the estimated parameters, transmission loss was calculated using the parameters and was compared with measured data.

4:20

2pUW13. Geoacoustic inversion of broadband acoustic recordings of a surface ship on an horizontal line array in the Gulf of Mexico. Robert A. Koch and David P. Knobles (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

As part of a comprehensive experiment to support research on geoacoustic inversion techniques, acoustic signals from a surface ship were recorded on a bottom mounted, horizontal line array at a soft sediment, shallow-water site in the Gulf of Mexico off Port Aransas, TX. Along with the surface ship recordings, water sound-speed profiles, cw TL data at 53, 103, 153, 503, and 953 Hz, and recordings of light bulb implosions were obtained. In this paper a method for performing inversions from broadband beam data is demonstrated and compared with corresponding inversions from broadband element data and with previously reported inversions of the light bulb implosions. The performance in a simulated annealing algorithm of cost functions that involve coherent or incoherent sums over frequency and multiple time segments is presented, and the uncertainties in the geoacoustic parameter values are examined. Comparisons of measured TL with predicted TL for the bottom descriptions resulting from the inversions are used to judge validity of the inverted profiles and the efficacy of the various cost functions. The applications of this work to ocean environments for which the bottom parameter values are unknown or uncertain will be discussed. [Work supported by ONR.]

4:35

2pUW14. Finite-element technique for multi-fluid 3-D two-way propagation in ocean waveguides. Mario Zampolli, Finn B. Jensen, David S. Burnett, and Carlo M. Ferla (SACLANT Undersea Res. Ctr., Viale S. Bartolomeo 400, 19138 La Spezia, Italy)

A finite-element technique is presented which makes it possible to treat 3-D forward- and backward scattering in underwater waveguides with inhomogeneous layers of fluid and sediment of arbitrary geometry. The

finite-element multi-fluid scattering model was derived from the linear wave equation via the Galerkin residual formulation. From the numerical model, software was developed by customizing an *hp*-adaptive finite-element package developed by COMCO/Altair Engineering in Austin, Texas. Based on automatic error estimation, it is possible to adapt the user-supplied inhomogeneous computational mesh by making the element size smaller, so-called *h*-refinement, and/or by increasing the order of the approximating polynomials, so-called *p*-enrichment, in selected regions of the domain. This process is repeated until a desired level of accuracy in the numerical solution is eventually achieved. The code is validated against a coupled normal-mode model for two-dimensional geometries, and results for shallow-water two-way propagation and scattering with strongly three-dimensional features are presented. The finite-element software is particularly useful for the prediction of scattering from geometries with high local complexity, and as a benchmarking tool for possible future three-dimensional underwater-propagation models.

4:50

2pUW15. Seafloor characterization from inversion of high-frequency backscatter. Frank W. Bentrem, John Sample, and Will Avera (Marine Geosciences Div., Naval Res. Lab., Stennis Space Center, MS 39529, frank.bentrem@nrlssc.navy.mil)

Geoacoustic inversion of high-frequency backscatter data is presented for characterization of the seafloor. The APL-UW (Applied Physics Laboratory at the University of Washington) backscattering model is used to model the grazing-angle dependence of the backscattering strength for a number of seafloor parameters. Backscattering strength versus grazing angle data sets (with frequencies 12–35 kHz) are provided for three sites along with sediment ground truth. Roughness measurements are also available for two of the sites. Inversion of these data sets is performed via simulated annealing, with some of the parameters constrained by empirical relationships with mean grain size. From the inversion, estimates are obtained for mean grain size, roughness, and volume interation. Inversion of data from a smooth, silty site in Arafura Sea yields estimates in good agreement with ground truth. Results also compare well with ground truth in a rough, sandy site (Quinault) and in Onslow Bay (sand). [Work supported by SPAWAR.]

5:05

2pUW16. Modal inversion results for geoacoustic properties in the SWAT experiments. Kazuhiko Ohta, Kouki Okabe, Masanichi Oikawa (5th Res. Ctr., TRDI, JDA 3-13-1 Nagase, Yokosuka 239-0826, Japan), Itaru Morishita, Hirotaka Murakami (Oki Electric Industry Co., Ltd., Tokyo 108-8551, Japan), and George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Bottom sediment properties at sites on the New Jersey Shelf and in the East China Sea were studied in the Shallow Water Acoustic Technology (SWAT) experiments. In these experiments, a source towed at constant depth transmitted low-frequency cw signals, which were measured on a bottom-moored vertical line array. The Hankel transform was applied to the acoustic field measured on the resulting synthetic aperture horizontal array created at each receiver depth. The horizontal wave number spectra, with peak positions corresponding to the modal eigenvalues, were observed to be slightly different among the different receiver depths, partially due to noise and range dependency. Thus stochastic mode inversion was exploited by using all of the observed peak positions for estimation of the geoacoustic properties. The sound field simulated using the inversion results agrees well with the measured one for each receiver depth.

5:20

2pUW17. Source tow geoacoustic inversions from the 2001 ASIAEX East China Sea experiment. Chen-Fen Huang and William S. Hodgkiss (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238, chenfen@mpl.ucsd.edu)

During the 2001 ASIAEX East China Sea experiment, source tow data were collected by a 16-element, 75-m aperture, autonomously recording vertical line array in 105-m-deep water. Transmissions from two similar 6-km-long tracks have been analyzed. In the first, CW tonals at 95, 195, 295, 395, 805, 850, and 905 Hz were transmitted from a J-15 transducer at a nominal depth of 46 m. In the second, CW tonals at 1.6, 2.4, 3.5, and 4.4 kHz were transmitted from an ITC-2015 transducer at a nominal depth of 49 m. Inversion results for seafloor geoacoustic parameters from these transmissions will be presented. [Work supported by ONR.]

Session 3aAA

Architectural Acoustics, Musical Acoustics and Physical Acoustics: Ancient Acoustics I Megaliths and Pyramids

David Lubman, Cochair
David Lubman & Associates, 14301 Middletown Lane, Westminster, California 92683

Fernando J. Elizondo, Cochair
Acoustics Laboratory, FIME, Univ. A de Nuevo Leon, P.O. Box 28 "F," Cd. Universitaria, San Nicolas 66450, N.L. Mexico

Chair's Introduction—8:30

Invited Papers

8:35

3aAA1. Psychoacoustic influences of the echoing environments of prehistoric art. Steven J. Waller (Rock Art Acoust., 5381 Wellesley St., La Mesa, CA 91942, wallersj@yahoo.com)

Cave paintings and ancient petroglyphs around the world are typically found in echo rich locations such as caves, canyons, and rocky cliff faces. Analysis of field data shows that echo decibel levels at a large number of prehistoric art sites are higher than those at nondecorated locations. The selection of these echoing environments by the artists appears not to be a mere coincidence. This paper considers the perception of an echoed sound as a psychoacoustic event that would have been inexplicable to ancient humans. A variety of ancient legends from cultures on several continents attribute the phenomenon of echoes to supernatural beings. These legends, together with the quantitative data, strongly implicate echoing as relevant to the artists of the past. The notion that the echoes were caused by spirits within the rock would explain not only the unusual locations of prehistoric art, but also the perplexing subject matter. For example, the common theme of hoofed animal imagery could have been inspired by echoes of percussion noises perceived as hoof beats. Further systematic acoustical studies of prehistoric art sites is warranted. Conservation of the natural acoustic properties of rock art environments—a previously unrecognized need—is urged.

8:55

3aAA2. MegaSound: Sound in Irish megalithic buildings. Victor Reijs (Geniet, 15 Shenick Grove, Skerries, Dublin, Ireland, geniet@iol.ie)

Stimulated by the studies done by Paul Deveraux and Robert Jahn, research has been conducted on the sound properties of two megalithic chambers is Ireland: Dowth South and Fourknocks I. As reference measurements two normal rooms (bed- and bathroom) have been studied. The following aspects will be covered in the presentation: some theoretical background on acoustical modes (within a passage, a chamber, and a combination of them: Helmholtz resonator); tips for doing sound experiments inside megalithic chambers (like: equipment, measurement software, power provisioning and calibrating); frequency response measurements (between 20 and 200 Hz) for the surveyed chambers/rooms; comparison of the results with other researchers' results; background on the pitch of the human (male, female, and child) voices in neolithic times and recommendations for future research. The presentation also provides insight in the aeralization (simulation) of sound in a megalithic chamber, covering: software that can do these simulations; issues in finding the basic information, e.g., acoustic absorption coefficients and provide examples of the results. I would like to thank all the people who have provided constructive feedback on my work (http://www.iol.ie/~geniet/eng/megasound.htm).

9:15

3aAA3. The accidental (acoustical) tourist. Wayne Van Kirk (7368 Brace St., Houston, TX 77061)

The acoustical phenomenon observed at an ancient temple in the Great Ball Court at Chichen Itza was described as ''little short of amazing—an ancient whispering gallery'' by Silvanus G. Morley, leader of the Carnegie Institute's archaeological team that excavated and restored these structures in the 1920s. Since then, many others have experienced the extraordinary acoustics at Chichen Itza and other Maya sites. Despite these reports, archaeologists and acousticians have until recently shown little interest in understanding these phenomena. After experiencing Chichen Itza's remarkable acoustics as a tourist in 1994, the author commenced collecting and disseminating information about acoustical phenomena there and at other Mayan sites, hoping to stimulate interest among archaeologists and acousticians. Were these designs accidental or intentional? If intentional, how was the knowledge obtained? How were acoustical features used? This paper highlights the author's collection of anecdotal reports of mysterious Mayan acoustics (http://www.ianlawton.com/pal.htm), recommended reading for scientists and engineers who wish to pursue this fascinating study. Also recounted are some of the reactions of archaeologists—ranging from curious, helpful, and insightful to humorous and appalling—to outsiders' efforts to bring serious scientific attenation to the new field of acoustical archaeology.

9:35

3aAA4. Tonal response on the stairway of the main pyramid at La Ciudadela, Teotihuacan archaeological site. Sergio Beristain, Cecilia Coss, Gabriela Aquino, Jose Negrete (P.O. Box 75805, Mexico, D. F. C.P. 97300, Mexico, sberista@hotmail.com), and Pablo Lizana (ESIME, IPN, Mexico)

This paper presents new research on the very interesting audible effects produced by the stairways of many archaeological sites in Mexico. This investigation was made at the main stairway of the pyramid at La Ciudadela, Teotihuacan archaeological site. The effect previously studied was a chirped echo reflected from the stairway at normal incidence, which resembles the singing of the Quetzal. Now it is presented with the impulsive sound source and the listeners located at different angles, where apart from the characteristic chirped sound, several musical notes could be obtained and identified, covering a range of at least one half an octave. This evaluation was made at the site, where the effect is clearly audible, and it is supported with simple mathematics.

9:55–10:05 Break

10:05

3aAA5. Acoustical features of two Mayan monuments at Chichen Itza: Accident or design? David Lubman (David Lubman & Assoc., 14301 Middletown Ln., Westminster, CA 92683-4514, dlubman@ix.netcom.com)

Chichen Itza dominated the early postclassic Maya world, ca. 900–1200 C.E. Two of its colossal monuments, the Great Ball Court and the temple of Kukulkan, reflect the sophisticated, hybrid culture of a Mexicanized Maya civilization. The architecture seems intended for ceremony and ritual drama. Deducing ritual practices will advance the understanding of a lost civilization, but what took place there is largely unknown. Perhaps acoustical science can add value. Unexpected and unusual acoustical features can be interpreted as intriguing clues or irrelevant accidents. Acoustical advocates believe that, when combined with an understanding of the Maya worldview, acoustical features can provide unique insights into how the Maya designed and used theater spaces. At Chichen Itza's monuments, sound reinforcement features improve rulers and priests ability to address large crowds, and Ball Court whispering galleries permit speech communication over unexpectedly large distances. Handclaps at Kukulkan stimulate chirps that mimic a revered bird (''Kukul''), thus reinforcing cultic beliefs. A ball striking playing field wall stimulates flutter echoes at the Great Ball Court; their strength and duration arguably had dramatic, mythic, and practical significance. Interpretations of the possible mythic, magic, and political significance of sound phenomena at these Maya monuments strongly suggests intentional design.

Contributed Papers

10:25

3aAA6. *Quetzal* or not *Quetzal*, that is the question . . . . On the stairs of the *Castillo* monument in *Chichen Itza*. 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)

Some speculation will be presented about the hypothesis that states that the reflected sounds in the stairs of the *Castillo* building in *Chichen Itza*, Mexico, imitates the song of the *quetzal* bird. Some aspects of construction, both technical and social, are discussed as well as issues related with myth and fantasies.

10:40

3aAA7. Theoretical interpretation of a case study: Acoustic resonance in an archaeological site. Jorge Carrera (Eng. Faculty, UNAM, Mexico, jorgec00@yahoo.com) and Sergio Beristain (ESIME, IPN, Mexico)

It is well-known that the stairways of some Mexican archaeological sites, like Chichen-Itza or Teotihuacan, present an interesting sound reflection and resonance phenomenon which causes a special audible effect. In this paper, mathematical modeling of this situation is presented, and the practical phenomenon is discussed from a theoretical standpoint. More than an end in itself, the idea is, once this validated model is obtained, to use the results for the analysis of a more extensive architectural environment in order to establish whether this kind of phenomenon would have been purposely introduced in the design of the site. This will be presented in future publications.

10:55

3aAA8. Navajo oral history of a pre-Columbian amphitheater in Chaco Canyon, New Mexico. Taft Blackhorse, Jr. (Navajo Nation Historic Preservation Dept., P.O. Box 4950, Window Rock, AZ 86515) and Jay S. Williams (Univ. of New Mexico, Albuquerque, NM 87108)

A large performance space attached to a natural amphitheater in the cliff face has recently been identified at the center of the pre-Columbian Chaco Complex in Chaco Canyon, New Mexico. This location is known in the ceremonial history of the Navajo people (Din) as Tsbiinaholtsa Yalti (Concavity in Bedrock that Speaks). Tsbiinaholtsa Yalti is a portal to the dimension of the deities and it is opened by way of tonally induced acousma. The physical manifestation of the amphitheater and its acoustical properties invoke the concepts of Tal (Chants) and Taal (Ceremonial pathway). Navajo ceremonies are called Haataal and religious practitioners Hataalii (Chanters). The origin of the tones which give power to contemporary Navajo chants may be traced to the Tsbiinaholtsa Yalti. These tones are produced vocally and are accompanied by the shell trumpet, eagle bone whistle, and reed flute (jadzoosh).

11:10

3aAA9. Computer analysis of sound recordings from two Anasazi sites in northwestern New Mexico. Richard Loose (MEVATEC Corp., P.O. Box 1419, Las Cruces, NM 88004)

Sound recordings were made at a natural outdoor amphitheater in Chaco Canyon and in a reconstructed great kiva at Aztec Ruins. Recordings included computer-generated tones and swept sine waves, classical concert flute, Native American flute, conch shell trumpet, and prerecorded music. Recording equipment included analog tape deck, digital minidisk recorder, and direct digital recording to a laptop computer disk. Microphones and geophones were used as transducers. The natural amphitheater lies between the ruins of Pueblo Bonito and Chetro Ketl. It is a semicircular arc in a sandstone cliff measuring 500 ft. wide and 75 ft. high. The radius of the arc was verified with aerial photography, and an acoustic ray trace was generated using CAD software. The arc is in an overhanging cliff face and brings distant sounds to a line focus. Along this line, there are unusual acoustic effects at conjugate foci. Time history analysis of recordings from both sites showed that a 60-dB reverb decay lasted from 1.8 to 2.0 s, nearly ideal for public performances of music. Echoes from the amphitheater were perceived to be upshifted in pitch, but this was not seen in FFT analysis. Geophones placed on the floor of the great kiva showed a resonance at 95 Hz.

Session 3aAB

Animal Bioacoustics and Psychological and Physiological Acoustics: Electrophysiological Investigations of Animals I

Eduardo Mercado III, Chair
Center for Molecular and Behavioral Neuroscience, Rutgers University, 197 University Avenue, Newark, New Jersey 07102

Invited Papers

8:00

3aAB1. Noise changes receptive fields of auditory neurons in A1. David T. Blake and Michael M. Merzenich (513 Parnassus Ave. S-877, San Fransisco, CA 94143-0732)

Primates engage in auditory behaviors under a broad range of signal to noise conditions. In this study, optimal linear receptive fields were measured in alert primate A1 in response to stimuli that vary in spectrotemporal density. As the random tone pip density increased, selective A1 excitatory receptive fields systematically changed. Receptive field sensitivity, expressed as the expected change in firing rate after a tone pip onset, decreased by an order of magnitude. Spectral selectivity more than doubled. Inhibitory subfields, which were rarely recorded at low sound densities, emerged at higher sound densities. The ratio of excitatory to inhibitory population strength changed from 14.4:1 to 1.4:1. At low sound densities, the sound associated with the evocation of an action potential from an A1 neuron was broad in spectrum and time. At high sound densities, a spike-evoking sound was more likely to be a spectral or temporal edge, and was narrower in both time and frequency range. Prediction experiments were performed to validate the assumption that linear receptive fields were representative of neural responses at high noise densities. The auditory context alters A1 responses across multiple parameter spaces; this presents a challenge for reconstructing neural codes.

8:20

3aAB2. Psychophysical and physiological measures of spectro-temporal processing in humans. David A. Eddins and Ann C. Eddins (Ctr. for Hearing and Deafness, Dept. of Communicative Disord. and Sci., Univ. at Buffalo, Buffalo, NY 14214)

The normal perception of natural acoustic stimuli requires the brain to extract, encode, and interpret multidimensional patterns derived from the acoustic input as relayed by the auditory periphery. Of particular importance in audition are patterns of intensity variations across time and frequency. To better understand the relation between perceptual abilities and underlying physiological mechanisms, one may study the relation between measures of psychophysical performance and physiological coding in humans on a range of auditory tasks. Furthermore, similar physiological measures are easily obtained in a variety of animal species, providing a critical link between human perception and relevant animal models. The first series of experiments to be reviewed involves behavioral and physiological measures of auditory temporal processing including the perception of sinusoidal amplitude modulation and nonsimultaneous masking. The second series involves the perception of spectral envelope features using behavorial and physiological estimates of the spectral modulation transfer function. The ability to combine or compare temporal patterns across different spectral regions will be examined via the comodulation masking release (CMR) paradigm. Finally, the processing of local and global spectral features will be examined and hemispheric dominance for each will be discussed.

8:40

3aAB3. Physiological measures of auditory temporal integration in chinchillas and humans. Ann Clock Eddins and David A. Eddins (Dept. of Communicative Disord. & Sci., Ctr. for Hearing and Deafness, Univ. at Buffalo, 215 Parker Hall, Buffalo, NY 14214)

Temporal integration is a phenomenon in which signal detection improves with increasing signal duration. The magnitude of threshold improvement has been shown to vary as a function of signal frequency and is often diminished in the presence of hearing loss. In studies of chinchilla auditory nerve fibers and cochlear nucleus (CN) neurons, we have shown that single neuron thresholds improve with increasing duration in a manner similar to data from psychophysical studies. Moreover, it appears that a critical feature of duration coding in the CN is the pattern of spike activity rather than the total number of spikes over time. Thresholds measured in evoked-potential studies with chinchillas at the level of the inferior colliculus and with humans at the level of the cortex also reflect physiological evidence of temporal integration, with comparable frequency effects and degradation with hearing loss like that observed psychophysically. Although frequency effects that show greater threshold improvement for low versus high frequencies have been reported in the literature for decades, the underlying mechanism remains unclear and is under further investigation psychophysically and physiologically in our laboratories. The results of these investigations will be discussed along with potential underlying mechanisms.

9:00

3aAB4. The effects of selective inner hair cell loss on auditory evoked potentials recorded from multiple levels of the auditory system of the unanesthetized chinchilla. Robert Burkard (Ctr. for Hearing & Deafness, Univ. at Buffalo, 215 Parker Hall, Buffalo, NY 14214, RFB@acsu.buffalo.edu)

Most animal models of sensorineural hearing loss show either selective outer hair cell (OHC) loss, or combined OHC and inner hair cell (IHC) loss. Several years ago, Robert Harrison and colleagues discovered that carboplatin could produce a selective IHC loss in the chinchilla. Carboplatin is a second-generation, platinum-based anti-neoplastic agent. In chinchillas, the effect of carboplatin administered systemically is dose dependent, and for the appropriate dosage, produces selective, patchy IHC loss that extends the length of the cochlea. As the vast majority of auditory-nerve afferents innervate the IHCs, this patchy IHC loss due to carboplatin leads to a partial deafferentation of the cochlea. This invited presentation will review a series of studies investigating near-field auditory evoked potentials in the chinchilla that have been performed at the University at Buffalo by the author and his colleagues. Chronic electrode implantation combined with passive animal restraint allow us to record from unanesthetized animals, thus avoiding the confounding effects of anesthesia. Responses from multiple levels of the nervous system (e.g., round window, auditory nerve, inferior colliculus, auditory cortex) will be reported, comparing responses before and after inducing partial IHC loss. [Work supported by NIH NICDC DC03600.]

9:20

3aAB5. Neural response characteristics in auditory cortex of the awake ferret. Didier A. Depireux, Bing-Zhong Chen, Peter Marvit, and Yaan Li (Univ. of Maryland, Baltimore, 685 S. Baltimore St., HSF 222, Baltimore, MD 21201)

We are developing an awake preparation for chronic physiological recording in the ferret. In this talk, we describe neural responses in auditory cortex in the restrained preparation to a variety of acoustic stimuli including tones, noise, auditory gratings, and their combinations. We characterize population responses, show how the cells properties change as a function of the stimulus (e.g., how the receptive field of a cell changes, depending on whether the carrier under the gratings spectral envelope is a harmonic complex or noise) and correlated the activity of single units from neighboring cells in a cortical column in response to these stimuli. [Work supported in part by NIH Grant No. R01 DC05019-01A1 and a training grant to the Program in Neuroscience at UMAB.]

9:40

3aAB6. Models of birdsong learning. Daniel Margoliash (Dept. of Organismal Biol. & Anatomy, Univ. of Chicago, Chicago, IL 60637, dan@bigbird.uchicago.edu)

The study of birdsong learning and the passerine song system represents an excellent opportunity to understand how complex acoustic perceptions are represented in the brain, and how these interact with motor systems to produce learned behavior. Recent advances suggesting specific functional roles for different nuclei within the song system help to constrain proposed models of the sensorimotor phase of learning. In particular, a descending motor pathway organized in a temporal hierarchy is involved in motor programming whereas an anterior forebrain basal ganglialike pathway may provide auditory and motor feedback-mediated error signals to adjust the motor pathway. A separate pathway outside of the song system may be devoted to perception of conspecific songs. Central to the problem of vocal learning is the time delay between motor command and auditory feedback. Both real time online models and offline models have been proposed as solutions to the temporal credit assignment problem. The models differ in the types of synaptic mechanisms envisioned that involve forward predictions, and the relative importance of state dependency of song system activity and song learning. Recent data will be given that describe the insight into neural solutions to these problems.

10:00–10:10 Break

10:10

3aAB7. Audiomotor integration for active sensing in the echolocating bat, *Eptesicus fuscus*. Cynthia Moss and Shiva Sinha (Dept. of Psych., Inst. for Systems Res., Univ. of Maryland, College Park, MD 20742)

Echolocation depends upon the dynamic interplay between auditory information processing and adaptive motor control. We are conducting experiments aimed at understanding the neural mechanisms supporting audiomotor integration for echolocation in the big brown bat, *Eptesicus fuscus*. In this work we focus on the superior colliculus (SC), a midbrain structure implicated in species-specific orienting behaviors. We have characterized the spatial response profiles of auditory neurons in the bat SC, and have discovered a population of cells that show echo-delay tuning, a response property believed to encode target distance in the bats sonar receiver. These data reveal a 3-D spatial coordinate system that may be used to guide appropriate orienting responses. Our studies also demonstrate that the bat SC plays a functional role in the execution of motor commands used for acoustic orientation by sonar. In particular, SC microstimulation elicits head and pinna movements, along with the production of sonar vocalizations. Multi-unit recordings from the SC of tethered, vocalizing bats reveal bursts of neural activity preceding the production of each sonar cry. Collectively, these results suggest that the bat SC plays a functional role in both auditory information processing and orienting behaviors that operate together in echolocation. [Work supported by the NSF, the NIMH, and the Whitehall Foundation.]

10:30

3aAB8. Neuroethology of audition and bat evasion in praying mantises. David D. Yager (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, dy5@umail.umd.edu)

Many praying mantises have a unique auditory system comprising a single ear located in the ventral midline of the thorax. The hearing is nondirectional, and best frequencies range from 25 kHz to over 100 kHz depending on the species. Bat-like ultrasound triggers complex evasive maneuvers in flying mantises that requires CNS processing in the head. A mirror-image pair of interneurons

(501) rapidly ($<$20 ms) carries auditory information from the auditory nerve to the brain and these neurons are the best candidates as the primary input to the evasive response. The 501 recordings from an electrode chronically implanted in a mantis while it is being attacked by a flying, echolocating bat show that 501 responds strongly to bat cries and faithfully reports the temporal pattern of the echolocation cries during the early stages of the attack. However, during the last 200–250 ms before capture (beginning in mid-Buzz I), 501 falls silent. Although consistent with several physiological characteristics of 501, the result is surprising behaviorally. However, based on behavioral latencies, the latest that a successful response could be triggered is 200–250 ms before capture, and silencing 501 after that may prevent potentially dangerous habituation of the interneuron.

10:50

3aAB9. Time and frequency information processing in the midbrain: Evidence found in the gerbil inferior colliculus. Hiroshi Riquimaroux and Katuhiro Maki (Dept. of Knowledge Eng. and Computer Sci., Doshisha Univ., Japan)

We have investigated the firing pattern of neurons in the central nucleus of the inferior colliculus (ICc) in time-frequency coordinate. The ICc is the major nucleus in the auditory midbrain. Single-unit recordings were made from ICc contralateral to the monaurally stimulated ear in anesthetized gerbils. We could classify firing patterns into four groups, except the simple ON neurons. Three out of four groups of neurons (53.8%) demonstrated that firing distribution changed depending on time and frequency, which was not shown previously. One group (37.6%) exhibited a frequency response range that changed little with time. The remainder (8.6%) belonged to none of the above. The time-frequency sensitive neuron in ICc may be a good candidate to code communication sounds, which comprise complex temporal change in frequency. [Research supported by Special Coordination Funds from the Science and Technology Agency and a grant to RCAST at Doshisha University from the Ministry of Education and Science of Japan.]

11:10

3aAB10. FM signals produce a robust paradoxical latency shift that is important for coding of target range in the bat inferior colliculus. Albert S. Feng and Alexander Galazyuk (Dept. of Molec. & Integ. Physiol. and Beckman Inst., Univ. of Illinois, 405 N. Mathews, Urbana, IL 61801)

Echolocating bats utilize the time delay between an outgoing ultrasonic pulse and its echo to determine target range. Many neurons at the inferior colliculus (IC) and above are tuned to time delays between these sound pulses of unequal amplitudes. Sullivan previously proposed that paradoxical latency shift (PLS), characterized by a quantal increase in firing latency to loud sounds, is important for this attribute because PLS permits coincidence detection that is important for the creation of delay-tuned responses. In the IC of little brown bats, Galazyuk and Feng recently reported that, in response to tone pulses, the proportion of neurons showing PLS was low ($<$20%). This study was undertaken to determine whether PLS is a function of the acoustic stimulus. For this, the temporal discharge patterns of single IC neurons were investigated over a broad range of sound levels, using tone pulses at CF as well as FM sound pulses that mimicked bats' ultrasonic cry as stimuli. For many IC neurons, tone pulses did not elicit PLS but FM sound pulses produced robust PLS. This result showed that PLS is stimulus dependent and that the bat's auditory system is optimal for processing FM sounds employed during echolocation. [Work supported by NIH R01DC04998.]

11:30

3aAB11. Phase sensitivity of auditory brain-stem responses in echolocating big brown bats. Michael J. Ferragamo (Dept. of Biol., Gustavus Adolphus College, St. Peter, MN), Mark I. Sanderson, and James A. Simmons (Brown Univ., Providence, RI 02912)

Multiple behavioral experiments show that echolocating big brown bats perceive 180° phase shifts of ultrasonic (20–100 kHz) FM echoes as delay changes of ± 15 μm. These bats represent FM sweeps as coherent auditory spectrograms in which low-pass smoothing of half-wave-rectified hair-cell excitation is the critical limiting parameter. Computational modeling of auditory spectrograms combined with Monte Carlo simulation of echo delay psychophysics reveals that coherence is preserved when the auditory low-pass smoothing cutoff is as low as 7–10 kHz, which is not nearly as high as the 20- to 50-kHz ultrasonic frequencies that seem necessary intuitively. Local-field-potential recordings from the bat's auditory brain-stem are sensitive to the starting phase of tone-bursts at frequencies up to 14 kHz, manifested as a change in LFP wave-form shape with phase. For most sites this phase sensitivity exhibits a strong dependence on stimulus amplitude. Typically, tone-bursts of 65–75 dB SPL evoke significant changes in LFP responses across different phase conditions. These results appear to confirm that low-pass smoothing in the bat's auditory transduction process occurs at higher frequencies than for most mammals, and at frequencies that have the ability to confer coherence on the resulting auditory representation of biosonar echoes.

11:50

3aAB12. Genesis of a space map in the owl's brain. Jose L. Pena (Dept. of Biol., 216-76, Caltech, Pasadena, CA 91125, jose@etho.caltech.edu)

In order to localize a sound source, the brain uses monaural and binaural cues that convey spatial information. In the owl, space-specific neurons of the inferior colliculus respond only to sounds coming from a particular direction and represent the results of this computation. The parallel pathways that process interaural time difference (ITD) and interaural level difference (ILD) merge in the external nucleus of the inferior colliculus where the space-specific neurons are selective to combinations of ITD and ILD. How is the combination selectivity to ITD–ILD pairs achieved? A multiplication of postsynaptic potentials tuned to ITD and ILD can account for the response of these neurons to ITD–ILD pairs. There are very few examples of multiplication by neurons or neural circuits; however, some computational models assume the existence of this basic arithmetic in sound localization. The owl's auditory system uses such operation in creating a two-dimensional map of auditory space.

Session 3aBB

Biomedical Ultrasound/Bioresponse to Vibration: Lithotripsy I

Robin O. Cleveland, Cochair
Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Achim M. Loske, Cochair
Centro de Física Aplicada y Technología Avanzada, UNAM, A.P. 1-1010, Querétaro, Qro. Mexico

Chair's Introduction—8:25

Invited Papers

8:30

3aBB1. A new fragmentation mechanism in extracorporeal shock wave lithotripsy and a first clinical study in China. Wolfgang Eisenmenger (Physikalisches Institut, Universitt Stuttgart, Pfaffenwaldring 57, D-70550 Stuttgart, Germany)

With ESWL-focus diameters of the order of the stone dimension or larger, fragmentation in planes perpendicular and parallel to the wave plane is observed. This is explained by circumferential pressure or ''squeezing'' of the stone by the wave propagating at the outside of the stone in the liquid or tissue. Since the pressure zone propagates with the sound velocity in the liquid which is below the propagation velocity in the stone, it causes an evanescent pressure zone in the stone resulting in tensile stress in planes perpendicular and parallel to the wave plane. A quantitative model predicting the ratio of pulses needed to fragment the stone to 2-mm particle size in relation to the number of pressure pulses needed for the first fragmentation is well in accord with experiments, supporting the ''squeezing mechanism with binary fragmentation.'' On the basis of these results it now appears possible to optimize the pressure pulse parameters measured using the fiber-optic probe hydrophone (FOPH). With correspondingly optimized shock wave generator systems, a clinical study of the concept ''wide focus and low pressure'' ESWL was performed in scientific cooperation between the Physical Institute of the University of Stuttgart and the Xixin Medical Instruments Co., Ltd. in Suzhou, China. [W. Eisenmenger, ''The mechanism of stone fragmentation in ESWL,'' Ultrasound Med. Biol. **27**, 683–693 (2001); Eisenmenger, Du, Tang *et al.*, ''The first clinical results of wide focus and low pressure ESWL,'' *ibid.* **8**, 769–774 (2002) (in press).]

9:00

3aBB2. The relative contribution of stress waves and cavitation to the overall success of stone comminution in shock wave lithotripsy. Pei Zhong, Songlin Zhu, and Yufeng Zhou (Dept. of Mech. Eng. and Mat. Sci., Duke Univ., Box 90300, Durham, NC 27708)

The disintegration of kidney stones in shock wave lithotripsy (SWL) is caused primarily by the stress waves propagating inside the stone and by the collapse of cavitation bubbles in the surrounding fluid near the stone surface. Understanding the relative contribution of these different mechanical forces to the overall success of stone comminution is critical for improving the treatment efficiency of SWL. In this talk, we will provide an overview of the experimental studies under way at Duke University to delineate the contribution of stress waves and cavitation to the overall success of stone comminution in SWL. Limitations of individual contributing factors and their synergistic interaction will also be discussed. Finally, strategies to improve stone comminution efficiency while reducing tissue injury in SWL will be presented. [Work supported by NIH DK52985 and DK58266.]

9:20

3aBB3. Compact self-focusing piezoelectric shock wave generator using electrically prestressed transducer. Dominique Cathignol, Alain Birer, and Mohammad Ghohestani (INSERM Unite 556, 151 Cours Albert Thomas, 69424 Lyon Cedex 03, France)

Compact piezoelectric shock heads need to increase the pressure at the surface of the shell. However, the pressure is limited by the possible mechanical breakdown. So, we have proposed to electrically pre-stress the piezoelectric material by applying an electric field in the opposite direction of the polarization. Using this method, it was possible to reach, on a plane piston, a 5 MPa pressure value. According to this idea, a very compact shock wave generator made of a 1–3 piezocomposite material having a diameter and a focal length of 120 mm was developed. The maximum pressure and the width of the compressive wave at the focus were, respectively, 60 MPa and 1.5 microseconds. The focal zone is an ellipsoid 6 mm in the propagating axis and 3 mm in the perpendicular direction. The efficacy of this generator was measured as the number of shocks necessary to disintegrate plaster balls 15 mm in diameter. At full power, the number of shocks was only 150 which is rather the same number as the one obtained using electrohydraulic machine generally considered as the gold standard. These results show that piezoelectric material may be advantageously used for the manufacturing of shock wave generators.

9:40

3aBB4. Dual-pulse lithotripter accelerates stone comminution and reduces cell injury *in vitro*. Dahlia L. Sokolov, Michael R. Bailey, and Lawrence A. Crum (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105)

Peak acoustic pressures and cavitation generated in shock wave lithotripsy (SWL) appear to contribute to both desired stone comminution and undesired injury to surrounding renal tissue. Our dual pulse system, comprised of two opposing, confocal lithotripters and generating simultaneous, converging shock pulses, localizes and intensifies the peak pressures and cavitation. Comparison of cavitation damage to aluminum foil shows an 8-cm stripe of pits produced by a single-pulse lithotripter and a 1-cm stripe of deep pits produced by the dual-pulse lithotripter. 100 dual pulses generated at 15 kV comminuted gypsum stones placed at the geometric focus $F2$ into 8 times as many fragments and significantly reduced hemolysis in dilute blood 2 and 4 cm off $F2$ when compared to 200 single pulses generated at 18 kV. Thus the dual-pulse lithotripter enhanced comminution and reduced injury while cutting treatment time in half. Additionally, when cavitation was suppressed by placing the stones in glycerol, the improvement in comminution was reduced to only a twofold increase. This result indicates that the localized and intensified cavitation is the dominant mechanism in the accelerated comminution produced by the dual-pulse lithotripter. [Work supported by NIH Grants Nos. P01-DK43881 and R01-DK55674.]

10:00–10:20 Break

10:20

3aBB5. Lithotripter shock wave interaction with bacteria. Achim M. Loske (Centro de Física Aplicada y Tecnología Avanzada, UNAM A.P. 1-1010, Querétaro, Qro., México, loske@fisica.unam.mx

The worldwide success of extracorporeal lithotripsy and shock wave treatment of some orthopedic diseases is not doubted. In the future, shock waves could also be used as a method for *in vivo* drug delivery and gene transfer into target cells; however, basic research is still necessary to understand shock wave interaction with biological tissue, bacteria, and cells. The destructive effects of ultrasonic waves on micro-organisms and the reduction in renal infections after extracorporeal shock wave lithotripsy led to the idea of studying the effects of lithotripter shock waves on bacteria. Independent of possible applications to medicine, this could lead to a new, nonthermal food preservation method. The bactericidal effect of underwater shock waves on *Escherichia coli* ATCC 10536 and O157:H7, *Listeria monocytogenes* L8, and *Salmonella enterica servovar typhimurium* suspensions in isotonic saline solution were studied. Shock waves of about 55 MPa were generated using an electrohydraulic generator. Results indicate that pressure variations, shock wave-created cavitation, and the radiation resulting from the high-voltage spark significantly reduce the viability of these micro-organisms. An analysis of variance revealed that the bactericidal action of shock waves seems to depend on multiple-factor interactions, which vary depending on the type of bacteria.

Contributed Papers

10:40

3aBB6. Numerical simulation of shock and bubble dynamics in shockwave lithotripsy. Tim Colonius and Michel Tanguay (California Inst. of Technol., Pasadena, CA 91125, colonius@caltech.edu)

Theoretical evaluation of the efficacy of stone comminution (and potential for tissue damage) during shockwave lithotripsy requires knowledge of the complex stress fields associated with both the incident focussing shock and the dynamics of cavitation bubbles that it induces. While simple models from geometrical acoustics and subsequent modeling of spherical bubbles in isolation (Gilmore equation) can provide estimates, high-speed photography *in vitro* reveals a far more complex flow with bubble number densities that are sufficiently high such that collective effects associated with a cloud of bubbles are important. This talk will describe a modeling effort aimed at estimating stresses from these complex lithotripter generated flow fields. We compute the time-dependent, compressible, ensemble-averaged two-phase flow equations with a finite-difference scheme. Detailed modeling of the dynamics of bubbles (on the microscale) and high-order weighted essentially nonoscillatory shock-capturing schemes are employed. The model is compared to hydrophone and passive cavitation detection measurements, as well as qualitative comparison with high-speed photography. Finally, we explore collective bubble mechanisms ranging from defocusing and shielding of the stone (for high bubble densities in the focal region) to enhanced stresses due to concerted cloud collapse in a dual-pulse lithotripsy configuration. [Work supported by NIH P01 DK-43881 and NSF under grant CTS-9979258.]

10:55

3aBB7. Dynamical response of a bubble submitted to two following shock waves. Marie-Caroline Jullien (Faculty of Appl. Phys., Univ. of Twente, WB Bldg., P.O. Box 217, 7500 AE Enschede, The Netherlands)

A numerical study of the dynamical response of a bubble submitted to two following shock waves is reported. After the passage of a shock wave, a micron-size bubble expands enormously, reaching millimeter size, and then inertially collapses; this is the so-called cavitation phenomenon. The influence of the passage of a second shock wave on the bubble inertial collapse control is investigated. This control may have a considerable impact for drug delivery or gene transfer, by the use of extracorporeal shock wave lithotripsy, which is a technique already used in medical treatment for kidney stones. It is shown that the dynamical response of the bubble qualitatively depends strongly on the forcing shape. Furthermore, for a given forcing shape, the dynamical response depends on both the delay and forcing amplitude ratio between the two applied shock waves. [This work has been done in collaboration with Ruediger Toegel, Claus-Dieter Ohl, and Detlef Lohse (Twente University).]

11:10

3aBB8. Modeling the pressure pulse shape of piezoelectric lithotripters. Thomas Dreyer and Rainer Riedlinger (Universitaet Karlsruhe, IHE-Akustik, Kaiserstr. 12, D-76128 Karlsruhe, Germany, Thomas.Dreyer@ihe.uka.de)

Piezoelectric focusing transducers are widely used in extracorporeal lithotripsy. To optimize the therapeutically relevant focal pressure pulse it is necessary to affect the generated pulse shape at the transducer surface. Therefore a modeling approach is required containing the acousto-mechanical properties of the transducer structure as well as the influence of the electrical drive. The procedure presented here uses three dimensional transient finite element simulations to calculate an electro-acoustical impulse response of the transducer structure and linear systems theory to model the influence of the driving circuit on the emitted acoustical signal. Applying a short electrical pulse an acoustical impulse response can be simulated under plane wave conditions, which is valid at sufficiently large distances from the transducer. Focal pressures are estimated rapidly by linear calculations or accurately by a nonlinear propagation model. The influence of electrical drive conditions on the emitted acoustical signal is investigated very efficiently by a convolution with the desired electrical input, avoiding FEM simulations for each case. Reverting this process the

required driving voltage course for a given pressure signal is determined. Alterations of the pressure signal in terms of pulse width and tensile components are demonstrated theoretically, varying the design parameters of the transducer.

11:25

3aBB9. Comparison of the cavitation fields of an electromagnetic and electrohydraulic lithotripter. Parag Chitnis and Robin Cleveland (Aerosp. and Mech. Eng. Dept., Boston Univ., 110 Cummington St., Boston, MA 02215)

We contrast the cavitation of two lithotripters of different shock wave (SW) generation: electrohydraulic (EHL) and electromagnetic (EML). The cavitation field induced by the SWs was measured using a dual passive cavitation detector consisting of two 1 MHz focused transducers. The characteristic time (t_c) was obtained from the time difference of the arrival of the SW and the inertial collapse of the cavitating bubble. The t_c increased significantly with energy level for both lithotripters. At the highest energy level, $t_c = 332\ \mu s \pm 65\ \mu s$ for the EML and $t_c = 277\ \mu s \pm 43\ \mu s$ for the EHL. Increasing the pulse rate frequency from 1 Hz to 2 Hz led to a significant increase in t_c for the EHL but no significant change for the EML. The region of strong cavitation activity corresponded with the focal region of the lithotripters. Based on a 6 dB decrease in the amplitude of the inertial collapse signal, the region of strong cavitation was 7 mm wide for the EHL and 2 mm wide for the EML. Measurements of the acoustic pressure indicate that the -6 dB width of the focal spots were 12 mm and 4 mm, respectively. [Work supported by the NIH P01-DK43881.]

WEDNESDAY MORNING, 4 DECEMBER 2002 CORAL SEA 1 AND 2, 8:30 TO 11:20 A.M.

Session 3aMU

Musical Acoustics: Wind Instrument Measurement Techniques

Peter L. Hoekje, Cochair
Department of Physics and Astronomy, Baldwin-Wallace College, 275 Eastland Road, Berea, Ohio 44017

Leonardo Fuks, Cochair
Escola de Musica, Universidade do Brazil-UFRJ, Rua do Passeio 98, Rio de Janeiro 20021-290, Brazil

Chair's Introduction—8:30

Invited Papers

8:35

3aMU1. From air to rubber: New techniques for measuring and replicating mouthpieces, bocals, and bores. Leonardo Fuks (Escola de Musica, Universidade do Brasil-UFRJ, Rua do Passeio 98, Rio de Janeiro 20021-290, Brazil)

The history of musical instruments comprises a long genealogy of models and prototypes that results from a combination of copying existing specimens with the change in constructive parameters, and the addition of new devices. In making wind instruments, several techniques have been traditionally employed for extracting the external and internal dimensions of toneholes, air columns, bells, and mouthpieces. In the twentieth century, methods such as pulse reflectometry, x-ray, magnetic resonance, and ultrasound imaging have been made available for bore measurement. Advantages and drawbacks of the existing methods are discussed and a new method is presented that makes use of the injection and coating of silicon rubber, for accurate molding of the instrument. This technique is harmless to all traditional materials, being indicated also for measurements of historical instruments. The paper presents dimensional data obtained from clarinet and saxophone mouthpieces. A set of replicas of top quality clarinet and saxophone mouthpieces, trombone bocals, and flute headjoints is shown, with comparative acoustical and performance analyses. The application of such techniques for historical and modern instrument analysis, restoration, and manufacturing is proposed.

9:05

3aMU2. Measurement of the acoustic response of a wind instrument with application to bore reconstruction. Maarten van Walstijn and Murray Campbell (Dept. of Physics and Astronomy, Univ. of Edinburgh, Mayfield Rd., Edinburgh EH9 3JZ, UK)

Reconstruction of a bore from measured acoustic response data has been shown to be very useful in studying wind instruments. Such data may be obtained in different ways; directly measuring the frequency–domain response of an acoustic bore has some distinct advantages over directly measuring time–domain data (for example, by pulse reflectometry), but so far has been unsuitable for producing input data for deterministic bore reconstruction algorithms, due to the limited accuracy at high frequencies. In this paper a method is presented for large-bandwidth measurement of the input impedance of a wind instrument using a cylindrical measurement head with multiple wall-mounted microphones. The influence of the number of microphones and the types of calibration impedance on the accuracy will be discussed, and bore reconstructions derived using this technique will be compared with reconstructions obtained using pulse reflectometry. [Work supported by EPSRC.]

9:35

3aMU3. Impedance comparison of high-quality alto and soprano saxophones. Vincent Gibiat (LAMI Univ. Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France) and Jerome Selmer (Henri Selmer, Paris, France)

Measuring the input impedance of high-quality wind instruments requires both a system able to provide the physical quantity in less than a few seconds and the possibility to do the measurement on a wide number of instruments. If we have such a system measuring system giving us the possibility to obtain impedance curves for the whole musical range of the instrument the second condition cannot be achieved without the strong collaboration of an instrument maker. The present study has been realized on Selmer saxophones and on instrmuents from two other factories. They include vintage instruments (as Mark VI) and the most recent commercialized ones. It allows one to study the evolution of the instruments. We will show that some alto saxophone models have reached an incredible harmonicity of their resonances while less common models present other characteristics. We will discuss these results and their relations with the audible quality of the instruments.

10:05

3aMU4. Measuring acoustic impedances using a semi-infinite waveguide reference: Applications to wind instruments and vocal tracts. Joe Wolfe, John Smith, John Tann, and Ryan France (School of Phys., Univ. of New South Wales, Sydney 2052, Australia)

Acoustic pressures may generally be measured with much greater sensitivity, dynamic range, and frequency response than acoustic currents. Consequently, most measurements of acoustic impedance consist of comparison with standard impedances. The method reported here uses a semi-infinite waveguide as the reference because its impedance is purely resistive, frequency independent and accurately known, independent of theories of the boundary layer. Waveguides are effectively infinite for pulses shorter than the echo return time, or if the attenuation due to wall losses (typically 80 dB) exceeds the dynamic range of the experiment. The measurement signal from a high output impedance source is calibrated to have Fourier components proportional to fn, where n may be 1 for convenience or chosen to improve the signal:noise ratio. The method has been used on diverse systems over the range 50 Hz to 13 kHz. When applied to systems with simple geometries, the technique yields results with a little higher wall losses than those expected from the calculations of Rayleigh and Benade. Discontinuities introduce further losses as well as the expected departures from simple one-dimensional models. Measurements on musical wind instruments and on the human vocal tract are reported. [Work supported by the Australian Research Council.]

Contributed Papers

10:35

3aMU5. Acoustic input impedance measurements on brass instruments. Robert W. Pyle, Jr. (11 Holworthy Pl., Cambridge, MA 02138, rpyle@post.harvard.edu)

Measurement of the acoustic input impedance of a brass instrument can reveal something about the instrument's intonation, its reasonable playing range, its tone color, and perhaps whether the mouthpiece used for the impedance measurement is appropriate for the instrument. Such measurements are made at sound-presssure levels much lower than those encountered under playing conditions. Thus, impedance measurements may offer the only feasible way to infer something about the playing characteristics of instruments, typically museum specimens, that are too rare or too fragile to be played. In this paper the effects of some of the available choices of sound source and stimulus signal on measurement accuracy will be explored. Driver–transducer nonlinearity, source impedance, signal-to-noise ratio, and any necessary signal processing will be discussed.

10:50

3aMU6. Pulse reflectometry as an acoustical inverse problem: Regularization of the bore reconstruction. Barbara J. Forbes, David B. Sharp (Dept. of Environ. and Mech. Eng., The Open Univ., Walton Hall, Milton Keynes MK7 6AA, UK, b.forbes@open.ac.uk), and Jonathan A. Kemp (Univ. of Edinburgh, Edinburgh EH9 3JZ, UK)

The theoretical basis of acoustic pulse reflectometry, a noninvasive method for the reconstruction of an acoustical duct from the reflections measured in response to an input pulse, is reviewed in terms of the inversion of the central Fredholm equation. It is known that this is an ill-posed problem in the context of finite-bandwidth experimental signals. Recent work by the authors has proposed the truncated singular value decomposition (TSVD) in the regularization of the transient input impulse response, a non-measurable quantity from which the spatial bore reconstruction is derived. In the present paper we further emphasize the relevance of the singular system framework to reflectometry applications, examining for the first time the transient bases of the system. In particular, by varying the truncation point for increasing condition numbers of the system matrix, it is found that the effects of out-of-bandwidth singular functions on the bore reconstruction can be systematically studied.

11:05

3aMU7. Body vibrational spectra of metal flute models. Clare M. Hurtgen (Jaffe Holden Acoustics, Inc., 114A Washington St., Norwalk, CT 06854, churtgen@jhacoustics.com) and Dewey T. Lawson (Duke Univ., Durham, NC 27708)

For years, flutists have argued over the tonal advantages of using different precious metals for their instruments. Occasionally, scientists have entered the fray and attempted to offer an objective point of view based on experimental measurements. However, their research often involved actual instruments and performers, ignoring variations in wall thickness, craftsmanship, and human consistency. These experiments have been conducted using a variety of methods; all have concluded that wall material has no effect on tone. This paper approaches the question using simple tubular models, excited by a wind source through a fipple mouthpiece. The amplitude and phase of the harmonic components of the body vibrational signal were measured with a stereo cartridge. Results demonstrated the existence of complex patterns of wall vibrations in the vicinity of a tone hole lattice, at frequencies that match significant harmonics of the air column. Additionally, the tube wall was found to expand in a nonuniform or ''elliptical'' manner due to the asymmetry of the tone holes. While this method is somewhat removed from direct musical applications, it can provide an objective, quantitative basis for assessing the source of differences among flutes. [Work financed by two Undergraduate Research Support grants from Duke University.]

Session 3aNS

Noise: Hearing Protection I

Elliott H. Berger, Cochair
EAR/Aearo Company, 7911 Zionsville Road, Indianapolis, Indiana 46268-1657

Daniel P. Salomon, Cochair
Comaudi, Patriotismo 706, Colonia Mixcoac 03730, D.F. Mexico

Chair's Introduction—8:30

Invited Papers

8:35

3aNS1. International hearing protector standardization. Torben Poulsen (Acoust. Technol., Oersted-DTU, Bldg. 352, DTU, DK-2800 Lyngby, Denmark, tp@oersted.dtu.dk)

Hearing protectors shall fulfill some minimum requirements to their performance. As hearing protector manufacturers sell the products all over the world, the testing and certification of hearing protectors has become an international issue. The ISO working group WG17 under the headlines Acoustics, Noise, produce hearing protector standards to be used at an international level. The presentation will cover the ongoing work in WG17, including the revision of existing standards (ISO 4869-1, ISO 4869-3), upcoming new standards (ISO 4869-7) and the plans and status for future standards (performance in impulse noise, protectors with active noise reduction). Furthermore, an overview of the present European standards (CEN) and the relation to American and Australian/New Zealand standards will be discussed.

8:55

3aNS2. Development of Mexican standard on hearing protection devices. Daniel P. Salomon and Jose N. Razo (Comaudi, Patriotismo 706, Colonia Mixcoac, 03730, Mexico D. F., gerencia@comaudi.com)

A working group was established by the Secretaria del Trabajo (similar to OSHA) to work on a standard whose objective was to protect workers from noise by trying to predict the performance of HPDs as personal protection equipment in the workplace. Various standards were analyzed including ISO4869, AS-NZ 1270, and ANSI. All evaluate attenuation as a performance descriptor. Experimenter-Fit methods tend to overrate field performance due to very different field versus laboratory conditions. Subject-Fit methods result in lower attenuation values which are more likely to be achieved by groups of workers, even if inadequately motivated and trained. Subject-Fit methods encountered large resistance, including laboratories' difficulties in finding naive subjects, commercial interests of manufacturers, risk of overprotection, etc. Having settled for an Experimenter-Fit method, the intent is to do more to help protect workers' hearing through other means available. A thorough, yet fairly simple, HPD Selection and Usage Guide was prepared as an Informative Appendix. Relevant information should be provided to the user with the Packaging. NRR wars should be avoided. The standard includes a list of contents, instructions, and warnings that must be included. A sample label containing key information is detailed in another Informative Appendix.

9:15

3aNS3. The subjective testing of hearing protectors and issues of interlaboratory variability—The Australian/New Zealand experience. Warwick Williams (Natl. Acoust. Labs., 126 Greville St., Chatswood, NSW Australia, warwick.williams@hearing.com.au)

Over the past few years the proposal to use ''inexperienced'' subjects to test the attenuation of hearing protectors has been controversial, particularly in the USA and in connection with the development of *draft ISO 4869–7 Acoustics–hearing protectors–Part 7: Subjective method for measurement of sound attenuation–Subject-fit method*. Much of the discussion centers around the belief that by using ''inexperienced,'' ''unassisted'' test subjects attenuation results will be unrepresentative of the ''true'' figure. Both Australia and New Zeland have been using subjective testing of hearing protectors with inexperienced, unassisted test subjects for many years. The results have been positive and the conjectured difficulties have not eventuated. Any ''unusual'' results have been readily explained. The Australian acoustic community has no difficulty with the test methodology. Further it is the opinion of Occupational Health and Safety workers that ''inexperienced,'' subjective testing most closely represents attenuation values typical in the well informed workplace.

9:35

3aNS4. Hearing protection: Limits to attenuation. Elliott H. Berger, Ronald W. Kieper (E-A-R/Aearo Co., 7911 Zionsville Rd., Indianapolis, IN 46268-1657), and Dan Gauger (Bose Corp., The Mountain, Framingham, MA 01701-9168)

With louder and louder weapon systems being developed and military personnel being exposed to steady noise levels approaching and sometimes exceeding 150 dB, the interest in greater amounts of protection is evident. When the need for communications is included in the equation, the situation is even more extreme. New initiatives are underway to design improved hearing protection, including active noise reduction (ANR) earplugs and perhaps even cancellation of head-borne vibration. With that in mind it may be useful to explore the limits to attenuation, and whether they can be approached with existing technology. Data on the noise reduction achievable with high-attenuation foam earplugs, as a function of insertion depth, will be reported. Previous studies will be reviewed that provide indications of the bone-conduction (BC) limits to attenuation that, in terms of mean values, range from 40–60 dB across the frequencies from 125 Hz to 8 kHz. Additionally, new research on the effects of a flight helmet on the BC limits, as well as the potential attenuation from deeply inserted passive foam earplugs or communication earplugs, worn with passive earmuffs, or with active-noise reduction (ANR) earmuffs, will be also examined.

9:55

3aNS5. Do sound restoration hearing protectors provide adequate attenuation for gunfire noise? John R. Franks and William J. Murphy (NIOSH Hearing Loss Prevention Section, 4676 Columbia Pkwy., M.S. C-27, Cincinnati, OH 45226-1998, jrf3@cdc.gov)

Sound restoration earmuffs have an electronics package installed in a passive earmuff shell or earplug body. The electronics package consists of a microphone placed on the outside of the ear cup or earplug and a limiting amplifier driving a loudspeaker placed inside the ear cup or the earplug. The devices are designed to provide unity gain or better to low-level signals and to shut off when sound exceeds a given level. The effects of several sound restoration earmuffs and two nonlinear orifice earplugs on gunfire for several different weapon types were analyzed. In general, the nature of the electronics packages was such that the studied devices provided the same attenuation when turned on as when turned off. However, as with passive devices, a single protector, be it by earplug or earmuff, appears to be inadequate for gunfire when more than just a few shots are fired. Sound restoration earmuffs are best used with a well-fitted earplug, since the electronics can compensate for both the insertion loss of the earmuff and the earplug. However, nonlinear orifice earplugs may not provide sufficient protection for extended sessions of target practice.

10:15–10:30 Break

10:30

3aNS6. Industrial warning signal detection under augmented hearing protection devices. John G. Casali and Gary S. Robinson (Dept. of Industrial & Systems Eng., Auditory Systems Lab., Virginia Tech, 260 Durham Hall, Blacksburg, VA 24061, jcasali@vt.edu)

The results from several experiments concerning the effects of passive, Amplitude-Sensitive Sound Transmission (ASST), and Active Noise Reduction (ANR) HPDs on vehicle backup alarm detection by normal hearers are reviewed. Masked thresholds under a Peltor electronic ASST earmuff and an AEARO orifice-based ASST earmuff were found not to differ significantly from thresholds under conventional earmuffs of similar volume to the ASST devices, in pink noise of 75, 85, and 95 dBA. A separate analysis using a KEMAR manikin indicated that subjects' chosen settings of the Peltor ASST earmuff gain control did not significantly increase the noise exposure dose over the earmuff's amplifier-off setting. In low frequency-biased ''red'' noise conditions (averaged across 85 and 100 dBA), masked thresholds under a Bose ANR earmuff and an AEARO foam earplug were significantly lower (better) than those under a Peltor passive earmuff having weaker low frequency attenuation, but there were no differences among the 3 devices in pink noise, nor at 85 dBA in either noise. Detection thresholds in 85 dBA noise were 3.2–4.5 dB lower in the 3 protected conditions than in the unprotected condition, attesting to the fact that HPD use can facilitate signal detection under certain conditions for normal hearers. [Work supported by U.S. Bureau of Mines (now NIOSH), Bose Corp., Aearo Corp.]

10:50

3aNS7. Single number to represent hearing protector attenuation. Samir N. Y. Gerges, Eric N. Gaste (Mech. Eng. Dept., Federal Univ. of Santa Catarina, Cx.P. 476, Florianopolis, SC, Brazil, CEP: 88040-900), and Vera M. Steffen (Federal Univ. of Rio rand de sul, Porto Alere, RS, Brazil)

The existing single numbers for hearing protector noise attenuation are NRR, SNR, HML, and NRRsf. All of these numbers are calculated using a standard ambient noise called pink noise with a total level of 100 dBC. This pink noise does not represent real industrial noise. This pink noise has its peak values in dBA at 2 kHz, which is not usually the case for most real industrial environments. Also, this single number is not associated with statistical distribution and standard deviation. In this paper a trial is presented to calculate the NRRsf using measured industrial noise spectrum. The calculation is carried out for a typical foam-hearing protector. The REAT naive subject results obtained from laboratory measurements of 20 subjects using ANSI 12.6-1997 (B) gives 20 individual attenuations as function of frequency. For each one of the real noise spectra measured in industry, using one subject attenuation, an NRRsf individual number is obtained. The average NRRsf and standard deviation of these attenuation values are presented along with their statistical distribution. These results are discussed and compared to the single classical numbers.

11:10

3aNS8. Issues regarding hearing protection device use in manufacturing and mining. Thais Morata (Hearing Loss Prevention Section, NIOSH C27/4676, Columbia Pkwy., Cincinnati, OH 45226)

In this study reasons offered by workers for not consistently using hearing protectors are examined. Study groups were comprised of: (1) 124 workers exposed to various levels of noise in a manufacturing company, located in Sao Paulo, Brazil. Data on work history, psychosocial aspects of the workers job, medical history, present health, stress, occupational and nonoccupational exposures to noise or chemicals, and lifestyle factors were collected through interviews. The participants had their hearing and noise exposure assessed. Sixty-four percent of the workers indicated that they wore hearing protectors, but only 20 percent of that subgroup stated that they wore the devices all the time when noise-exposed. (2) Forty-four workers and supervisors from the manufacturing and mining sectors in Cincinnati and Pittsburgh. Participants were interviewed or took part in focus groups discussions on a variety of hearing conservation topics, including problems with the use of HPDs. The variables significantly associated with a workers' decision not to wear hearing protectors consistently included interference with communication, interference with job performance, comfort issues, self-perception of hearing condition, and inability to detect warning signals when wearing HPDs. These are issues to be addressed in order to promote a more effective use of HPDs.

Contributed Papers

11:30

3aNS9. Software development for NIOSH hearing protector testing. William J. Murphy, John R. Franks (NIOSH Hearing Loss Prevention Section, 4676 Columbia Pkwy., M.S. C-27, Cincinnati, OH 45226-1998, wjm4@cdc.gov), and Douglas T. Rohrer (DTR Software Design and Development, Cincinnati, OH 45226)

In June 1998, NIOSH researchers recognized the necessity of rebuilding the NIOSH Hearing Protector Device Testing Laboratory. Several design requirements influenced the software development and hardware selection. First, the hardware must be commercially available and easily operated from personal computers that use the Microsoft Windows operating system. Second, software must have the flexibility to support not only the demands of a research laboratory and field studies but also be configurable such that a commercial laboratory could use the software for production testing of protectors. Finally, the complete package must ensure the use of standard operating procedures for the purpose of achieving and maintaining NVLAP certification for the ANSI S12.6 and S3.19 standards for the measurement of real-ear attenuation at threshold. The total system has been operational since May 2002 and has been installed in three laboratories to date. This paper will present how these competing needs were addressed in the development of the NIOSH HPD Lab software.

11:45

3aNS10. Sound propagation in straight, conical, exponential and real external ear canal. Alexsander Fortkamp, Washington J. N. de Lima, and Samir N. Y. Gerges (Federal Univ. of Santa Catarina, Cx.P. 476, Florianopolis, SC, Brazil)

Mathematical and numerical modeling of the noise attenuation of hearing protectors (HP) is a powerful tool for optimizing the HP design. The human ear canal has a complex geometry and therefore can be modeled by numerical methods. Below the cut-off frequency, plane wave propagation can be modeled by simple geometric. In this paper straight, conical, exponential, and real ear canal geometry are considered. The Webster equation is used to model the sound propagation inside these canals. In this paper a comparison of sound field propagation obtained from a plane wave model of the simple geometries and the finite elements model of the real ear are presented.

WEDNESDAY MORNING, 4 DECEMBER 2002 CORAL ISLAND 1 AND 2, 8:30 TO 11:45 A.M.

Session 3aPAa

Physical Acoustics and Signal Processing in Acoustics: The Coda and Other Stochastic Seismic Signals I

Richard L. Weaver, Chair

Department of Theoretical and Applied Mechanics, University of Illinois, 216 Talbot Laboratory, 104 South Wright Street, Urbana, Illinois 61801

Invited Papers

8:30

3aPAa1. Seismic coda. Michael Fehler (Los Alamos Natl. Lab., M.S. D443, Los Alamos, NM 87545)

Observations and analysis of seismic scattering in the heterogeneous earth have grown from initial observations in the 1960s into a well-developed subfield of seismology. The field presents many challenging, interesting, and fruitful areas of research for seismologists. The term seismic coda has been used to describe the overall discipline but the term coda was originally used only for the tail portion of the seismograms recorded within 100 km of earthquakes. Several characteristics of seismic coda have been identified that provide important constraints on models to explain the observed waveforms. It is now generally accepted that coda is a superposition of waves that have been multiply scattered from heterogeneities in the Earth's lithosphere. Coda waves, even if they are not completely understood, have practical use for seismologists because some of their characteristics lead to extremely reliable methods for quantifying relative sizes of earthquakes and relative local site amplification characteristics in seismic hazards investigations. Observations and theoretical modeling studies of coda waves have provided insight into the complexity of wave propagation in the Earth and have yielded new insights into the character of the Earth's lithosphere.

9:00

3aPAa2. Radiative transport for elastic waves. George Papanicolaou (Dept. of Mathematics, Stanford Univ., Stanford, CA 94305)

Radiative transport equations for elastic waves in random media can be derived from first principles in a certain regime of physical parameters. They account for both longitudinal and transverse waves, as well as polarization. Transport boundary conditions can also be derived from first principles. An important application of transport theory is to the analysis of the deep coda of elastic waves, where some rather interesting phenomena of energy distribution emerge. A reference to work in this area is by Ryzhik *et al.* [Wave Motion **24**, 327–370 (1996)]. Another reference, more directly related to seismics, is Ryzhik *et al.* [Bull. Seismol. Soc. Am. **86**, 1107–1115 (1996)].

9:30

3aPAa3. Quasi-2D diffusion of elastic waves. Bart Van Tiggelen, Nicolas Tregoures (Lab. de Physique et Modlisation des Milieux Consenss, CNRS/UJF, Maison des Magistres, BP 166, F-38042 Grenoble, France), Michel Campillo, Ludovic Margerin, and Celine Lacombe (Lab. de Geophysique Interne et Tectonophysique, F-38041 Grenoble, France)

Recent studies on seismic Coda and seismic equipartition have shown the possibility to model the propagation of seismic waves around 5 Hz in the crust using a heterogeneous layered model with a free surface overlying a homogeneous half-space. The depth of the layer is roughly 30 km. If the mean free path of the waves is in excess of this value, and this is probably the case in many places in the world (except perhaps Mexico), the wave diffusion should be described by mode diffusion in a 2D layer, rather than plane waves in a 3D medium. This notion affects many observed and potentially observable quantities, such as equipartition, leakage of energy, and coherent backscattering. [Work supported by the French Ministry of Research.]

10:00–10:15 Break

10:15

3aPAa4. Long codas of coupled wave systems in seismic basins. Thomas H. Seligman (Centro de Ciencias Fisicas, Univ. of Mexico, Cuernavaca, Mexico)

Quite some time ago it was pointed out that the damage patterns and Fourier spectra of the 1985 earthquake in Mexico City are only compatible with a resonant effect of horizontal waves with the approximate speed of sound waves in water [see Flores *et al.*, Nature **326**, 783 (1987)]. In a more recent paper it was pointed out that this indeed will occur with a very specific frequency selection for a coupled system of Raleigh waves at the interface of the bottom of the ancient lakebed with the more solid deposits, and an evanescent sound wave in the mud above [see J. Flores *et al.*, Bull. Seismol. Soc. Am. **89**, 14–21 (1999)]. In the present talk we shall go over these arguments again and show that strong reflection at the edges of the lake must occur to account for the strong magnification entailing necessarily a long coda, and that the mecanism can be understood in the same terms.

Contributed Papers

10:45

3aPAa5. Numerical modeling of seismogram envelopes in 2-D random media. Michael Fehler (Los Alamos Natl. Lab., M.S. D443, Los Alamos, NM 87545)

Several portions of seismograms recorded from regional earthquakes cannot be easily explained as resulting from waves propagating along deterministic paths within the Earth. For example, seismic coda, which is the tail portion of the seismogram of an earthquake recorded at distances of less than 100 km, is considered as resulting from waves that are multiply scattered from random heterogeneities in the Earth's lithosphere. At greater distances, observations that the duration of the initial arriving wave packet is much longer than the source–time duration is explained as being due to multiple forward scattering along the path between the source and the receiver. To investigate these phenomena, we use a finite difference method to numerically simulate 2-D scalar-waves that propagate through random media characterized by a von Karman autocorrelation function. Such media are considered to be appropriate models for the random component of the structure of the Earth's lithosphere. We investigate the characteristics of the resulting wavefields and compare them with those of observed seismograms.

11:00

3aPAa6. Anisotropic diffusion of elastic waves. Joseph A. Turner and Liyong Yang (Dept. of Eng. Mech., W317.4 NH, Univ. of Nebraska-Lincoln, Lincoln, NE 68588, jaturner@unl.edu)

The scattering of elastic waves in heterogeneous, anisotropic media is relevant for ultrasonic and seismic waves for media including textured polycrystalline media, fiber-reinforced composites, and media containing oriented cracks. Elastic waves propagating through such media lose energy due to scattering from the heterogeneities of the material. In anisotropic media, this scattering attenuation is a function of propagation direction. Here, the case of transverse isotropy is examined. The problem is formulated in terms of the Dyson and Bethe–Salpeter equations for the mean and covariance of the anisotropic Green's function, respectively. Solution of the Dyson equation, using the first-order smoothing approximation (FOSA), gives attenuations for shear horizontal, quasicompressional, and quasishear waves as a function of direction and frequency. The FOSA is also used to reduce the Bethe–Salpeter equation to an elastic radiative transfer equation. In the limit of many scattering events, this equation then reduces to an anisotropic diffusion equation. The primary result is the diffusion tensor given in terms of the various scattering functions that characterize the transfer of elastic energy between the different wave types of directions. The results are discussed in terms of applications to textured metals and media with aligned cracks. [Work supported by NSF and DOE.]

11:15

3aPAa7. Diffusion of ultrasonic waves in porous glass bead sinters. John H. Page, James Beck, Russel Holmes, and Jake Bobowski (Dept. of Phys. and Astron., Univ. of Manitoba, Winnipeg, MB R3T 2N2, Canada)

Sintered networks of glass beads form an interesting example of a porous medium, analogous to a very porous rock, in which very strong multiple scattering of elastic waves is observed when the ultrasonic wavelength is comparable with the size of the pores. To investigate the diffusive transport of energy by multiply scattered waves, the transmission of the diffuse energy flux through finite slabs of this material has been measured. From these data, the diffusion coefficient D, as well as the absorption time, was determined by fitting the predictions of the diffusion approximation to the experimental time-of-flight profiles. To accurately measure

D, the internal reflection of diffuse waves at the sample boundaries was taken into account by extending the method used previously for acoustic waves [J. H. Page *et al.*, Phys. Rev. E **52**, 3106 (1995)]. The frequency dependence of the diffusion coefficient was measured over an extended range of frequencies, and compared with estimates of D from ballistic measurements of the scattering mean free path and group velocity. Because of its relatively simple structure, this material may be an ideal system for probing the diffusion of elastic waves, where diffuse waves have mixed character consisting of both longitudinal and transverse polarizations.

11:30

3aPAa8. The coda of artificial shots on volcanos. Ulrich Wegler (Institut für Geophysik und Geologie, Universität Leipzig, Talstraße 35, 04103 Leipzig, Germany, uli@rz.uni-leipzig.de)

The data of active seismic experiments at Merapi volcano and at Vesuvius volcano are used to analyze the scattering of elastic waves in the shallow structure of these volcanos. The seismograms from artificial explosions located on the volcanos are characterized by spindle-like envelopes, small or missing P-onsets, missing S-onsets, and long codas. These unusual shapes of the envelopes can be explained by strong multiple scattering in the shallow heterogeneous eruptive material and are modeled using the diffusion model. As a result diffusivities of $d=0.05$ km^2/s in the case of Merapi and of $d=0.1$ km^2/s in the case of Vesuvius are obtained independent of frequency between 2 to 20 Hz. Assuming the dominance of shear waves in the coda and a typical S-wave velocity of around 1.5 km/s for the shallow volcano this corresponds to a transport mean free path of only 100 m in the case of Merapi and of only 200 m in the case of Vesuvius. The corresponding length scales for intrinsic attenuation, on the contrary, depend on frequency and are at least one order of magnitude larger than the transport mean free path. These results show that multiple scattering is an important effect for the seismic wave propagation in volcanic environments.

WEDNESDAY MORNING, 4 DECEMBER 2002

CORAL GALLERY FOYER, 10:00 A.M. TO 12:00 NOON

Session 3aPAb

Physical Acoustics: Thermoacoustics and Resonators

Matthew E. Poese, Cochair
Graduate Program in Acoustics, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Guadalupe Huelsz, Cochair
Centro de Investigacion en Energia, A. P. 34, Temixco, Morelos 62580, Mexico

Contributed Papers

10:00

3aPAb1. Simulations of acoustic streaming inside and surrounding thermoacoustic stacks. Mark F. Hamilton, Yurii A. Ilinskii, and Evgenia A. Zabolotskaya (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063)

Acoustic streaming generated inside and surrounding thermoacoustic engine stacks by standing waves in resonators is investigated numerically. The study is performed using the nonlinear 2-D code described recently by the authors [J. Acoust. Soc. Am. **111**, 2076 (2002)]. The stack consists of parallel plates of negligible thickness. There is no restriction on plate separation, or on the length or position of the stack in the resonator. Temperature dependence of the fluid viscosity is taken into account. A temperature gradient imposed on the stack drives the sound field. Acoustic streaming patterns, investigated in terms of average mass flow velocity, are calculated by time averaging the momentum density. For typical plate separations on the order of the thermoviscous penetration depth, only the inner, boundary-layer streaming vortices exist between the plates. These inner vortices rotate in directions opposite those of the outer, more familiar Rayleigh streaming vortices encountered in wide channels (wider than about 10 penetration depths). The inner vortices extend beyond the ends of the stack out to distances on the order of the acoustic particle displacement. Flow patterns near the entrances to the stack are investigated as functions of channel width and acoustic amplitude [Work supported by ONR.]

10:15

3aPAb2. Use of complex frequency to analyze thermoacoustic engines. James E. Parker, Mark F. Hamilton (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029), Yurii A. Ilinskii, and Evgenia A. Zabolotskaya (Univ. of Texas, Austin, TX 78712-1063)

Rotts equations of thermoacoustics are expressed in the frequency domain, and as such they describe steady-state conditions. In numerical algorithms such as DELTAE the frequency is considered real, and the solutions correspond to the equilibrium states associated with the critical temperature gradient in the stack. An alternative approach, and the one described here, is to allow frequency to be complex and to determine its value for a specified temperature gradient. A nonzero imaginary part of the frequency indicates a nonequilibrium state, and depending on its sign it determines the rate at which the amplitude of the sound field either increases or decreases initially with time toward equilibrium at the critical temperature gradient. For solutions corresponding to thermoacoustic instability the imaginary part may be interpreted as a negative loss coefficient, which provides an estimate of the nonlinear loss required to offset the instability. This in turn determines the acoustic pressure at saturation. Graphs of the imaginary part versus the position and length of the stack for various resonators provide insight into favorable operating configurations. Consideration is also given to nonconstant temperature gradients along the length of the stack. [Work supported by ONR.]

10:30

3aPAb3. Performance of a small, low-lift regenerator-based thermoacoustic refrigerator. Matthew E. Poese and Steven L. Garrett (Grad. Prog. in Acoust., P.O. Box 30, State College, PA 16804, poese@psu.edu)

A regenerator-based thermoacoustic refrigerator [Swift, Gardner, and Backhaus, ''Acoustic recovery of lost power in pulse tube refrigerators,'' J. Acoust. Soc. Am. **105**(2), 711 (1999)] has been constructed. It is capable of moving about 5 W across a 40 °C temperature span. The machine operates with air at atmospheric pressure and is driven by an off-the-shelf electro-dynamic loudspeaker capable of producing peak-to-mean pressure ratios up to 12%. The thermal core of this research device contains an exhaust-side shell and tube heat exchanger (with water as the secondary heat transfer fluid), a regenerator made of 88 annular stainless-steel screens, and a constantan wire electrical heater that applies a measurable heat load to the cold side of the regenerator. An annular latex diaphragm is placed over the cold side of the regenerator to stop time-averaged mass flow through the regenerator and insulate the cold side [Gedeon, ''DC Gas Flows in Stirling and Pulse Tube Cryocoolers,'' in *Cryocoolers 9*, edited by R. G. Ross (Plenum, New York, 1997)]. Detailed measurements of heat load, temperature span, and exhaust heat flux will be presented and compared to DELTAE. [Work supported by ONR.]

10:45

3aPAb4. Competition between two different resonance acoustic modes in a thermoacoustic prime mover. W. R. Canul-Vargas, I. Vera-Manrique, R. Miranda, Y. Nahmad-Molinari, J. C. Ruiz-Suárez (Departamento de Física Aplicada, Cinvestav-Mérida, Mérida Yucatán 97310, Mexico), and G. Huelsz (UNAM, Temixco, Mor. 62580, Mexico)

When a stack of plates is placed near the closed end of an open tube, and a temperature gradient is applied to the stack (the hotter part next to the sealed tube end), a sound wave is produced. The greater the temperature gradient beyond a critical value, the stronger the sound wave produced. Inversely, to switch off the sound wave one has to decrease the temperature gradient applied to the stack. In the present work we report a new method to turn off and on this thermoacoustic prime mover almost instantaneously, without modifying the temperature gradient. A second resonator of different length and much smaller cross section (a tube with both ends open) is introduced into the first tube until contact with the stack is made. When this new resonator is left in place there is no change in the sound wave, but as soon as the end of the small tube outside the main resonator is closed, the sound wave disappears completely. Under this condition,the system goes rapidly from an acoustic thermodynamic cycle in which heat is transported by the wave, to a simple heat diffusion through the stack. We study the physics of this intriguing phenomenon.

11:00

3aPAb5. Measurement of the time evolution of Rayleigh streaming in high-amplitude standing waves. Michael W. Thompson and Anthony A. Atchley (Grad. Prog. in Acoust., The Penn State Univ., 217 Appl. Sci. Bldg., University Park, PA 16802, mwt126@psu.edu)

The time evolution of the axial component of the Rayleigh streaming velocity field in a cylindrical, air-filled resonator has been measured using laser Doppler anemometry. At low acoustic amplitudes, the measured field is in agreement with classical theory. The axial component varies sinusoidally along the axial direction and quadratically along the radial direction, and the position of maximum streaming velocity on the axis is located midway between the velocity node and velocity antinode. At higher acoustic amplitudes, the streaming velocity field initially resembles the classical result, but becomes progressively distorted as time passes. The position of maximum streaming velocity on the axis shifts towards the velocity antinode, while the streaming velocity outside the boundary layer, near the resonator wall, remains unaffected. Additionally, the radial dependence of the axial streaming velocity deviates from its initially parabolic shape, becoming flatter near the velocity node and steeper near the velocity antinode. The length of time required for the streaming to reach steady state is on the order of several minutes for a fundamental frequency of 310 Hz, a resonator radius of 23 mm, and an antinodal acoustic velocity of 6.0 m/s peak. [Work supported by ONR.]

11:15

3aPAb6. Computational and experimental investigation of minor losses in high amplitude acoustic resonators with varied cross section. Andrew J. Doller (Grad Prog. in Acoust., Penn State Univ., State College, PA 16804), Said Boluriaan, Anthony A. Atchley, and Philip J. Morris (Penn State Univ., University Park, PA 16802)

Minor losses in high amplitude oscillatory flows through changes in resonator cross section are not well understood. The application of numerical simulation in the study of these losses has shown promise over the past few years. A computational model has been developed based on a time accurate numerical simulation of the full Navier–Stokes equations on parallel computers. A rectangular clustered mesh is used to capture flow field details in the boundary layers and in the vicinity of changes in cross section. Results of simulations are compared to measurements of acoustic and time-averaged pressures at multiple points along the length of resonators and are compared at multiple drive amplitudes. The resonators consist of two sections of straight brass pipe of different diameters and joined through either steplike or conical couplers. They are driven at the end of the large diameter pipe with a rigid piston. The end of the small tube is rigidly terminated. [Research supported by ONR.]

11:30

3aPAb7. Effect of mean fluid flow on an acoustic standing wave in an open cavity. Michael Anderson, Kenneth Line, and Ralph Budwig (Dept. of Mech. Eng., Univ. of Idaho, Eng. Phys. Bldg., 324K, Univ. of Idaho, Moscow, ID 83844-0902)

Acoustic radiation pressure can be used to concentrate or remove small particles from an airborne aerosol. In this application, an ultrasonic transducer, mounted flush to one wall of a channel, is used to excite an integer half-wavelength standing wave of high amplitude that propagates perpendicular to the aerosol flow direction. An expression for the Fourier transform of the acoustic pressure in a semi-infinite channel, including the effect of mean fluid flow and finite transducer aperture, has been obtained. A parabolic (laminar) mean flow was assumed. The acoustic pressure was found to be governed by the Mach number of the flow, defined by the projection of the propagation direction relative to the mean flow velocity vector; and the aperture function of the transducer. Near a frequency of 50 kHz, numerical inversions of the acoustic pressure transform showed that the presence of mean flow in the velocity range 0-2 m/s caused changes in acoustic pressure on the order of 1%–4%. Corresponding experimental measurements showed changes in acoustic pressure up to 10%. The highest changes in measured acoustic pressure were found to occur up- and down stream relative to the transducer, and these patterns were in agreement with predictions of the analytical model.

11:45

3aPAb8. Feasibility study of acoustic agglomeration of fly ash in shaped resonators. Bart Lipkens and Theodore J. Conrad (Mech. Eng. Dept., Virginia Commonwealth Univ., 601 W. Main St., Richmond, VA 23284-3015)

An interesting method for controlling fine particulate emission is acoustic agglomeration. Acoustic agglomeration is a process where fine particles coalesce and grow into bigger particles when they are subjected to a high intensity sound field. As a result the particle size distribution shifts to larger sizes. Acoustic agglomeration can thus be used as a preconditioner prior to a conventional filter and improve the overall capture rate of the filter. An experimental study was performed to evaluate the potential of a new high intensity acoustics technology, resonant mac-

rosonic synthesis (RMS), to agglomerate fly ash. RMS allows the generation of very high amplitude shock-free acoustic standing waves within shaped resonators. Experiments were performed in a football shaped resonator at 570 Hz under varying fly ash loadings and sound pressure levels. Particle size measurements were obtained using a laser diffraction particle-sizing instrument. Reduction of fine particulate matter (diameter of particles less than 5 micron) of up to 50% has been achieved. [Work supported by Dominion Corp.]

WEDNESDAY MORNING, 4 DECEMBER 2002 CORAL GALLERY FOYER, 8:00 TO 9:45 A.M.

Session 3aPP

Psychological and Physiological Acoustics: Physiological Acoustics and Cochlear Implants

Linda Polka, Cochair
School of Communication Sciences and Disorders, McGill University, Beatty Hall, 1266 Pine Avenue West, Montreal, Quebec H3G 1A8, Canada

Fernando Elizondo-Garza, Cochair
Cd Universitaria, P.O. Box 28F, San Nicolas 66450, N.L. Mexico

Contributed Papers

8:00

3aPP1. The cochlear outer hair cell motility must be based on a tonic change in the membrane voltage: Implications for the cochlear amplifier. Jont B. Allen (Mimosa Acoust., 382 Forest Hill Way, Mountainside, NJ 07092) and Paul F. Fahey (Univ. of Scranton, Scranton, PA 18510)

It has been widely assumed that the function of the OHC is to increase the sensitivity and frequency selectivity of the cochlea via a phasic OHC voltage, which controls the soma length. This action is called the cochlear amplifier. According to this view the length of the OHC is assumed to follow the stimulus to the upper frequency limit of hearing, in a phasic manner (cycle by cycle), adding power at the signal frequency. We propose an alternative view that the OHC controls the dynamic range in a parametric, or tonic manner, via the cells axial stiffness. In this case the change in gain seen by the IHC does not require a phasic response at high frequencies. The OHC could mediate a fast acting gain control, via impedance changes, that follows the OHC membrane tonic voltage envelope. Given a level-dependent change in dynamic range (i.e., dynamic range compression), the tuning and sensitivity would necessarily change. Our analysis and conclusions are based upon a re-interpretation of existing mammalian outer hair cell (OHC) studies using a generalized admittance matrix formulation of the OHC, that relates the plasma membrane voltage and current to the soma axial force and velocity.

8:15

3aPP2. A measure of internal noise based on sample discrimination of level differences. Walt Jesteadt, Lance Nizami, and Kim S. Schairer (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

In a sample discrimination task, subjects are presented with stimuli drawn from each of two overlapping distributions and asked to identify the distribution from which the stimuli were drawn. Lutfi [J. Acoust. Soc. Am. **86**, 934–944 (1989)] has shown that performance in this task is nearly ideal when the stimuli are single sinusoids differing in level, for distributions with a mean difference of 5 dB and standard deviations of 5 dB. Theoretically, as the mean difference and standard deviation of the distributions are decreased, however, performance should increasingly deviate from ideal, yielding an estimate of internal noise. To test this hypothesis, intensity discrimination was measured for seven subjects' sinusoids drawn from distributions that had mean differences ranging from 0.1–2.2 dB, in steps of 0.3 dB. The standard deviation of each distribution was always half the mean difference, resulting in uniform ideal performance. Estimates of internal noise obtained by fitting a function to the observed *d*-prime values agreed with estimates obtained from psychometric functions measured using the same differences in level. Analyses of decisions on individuals trials indicated that the two observation intervals were weighted equally. [Work supported by NIDCD.]

8:30

3aPP3. Effects of the finite speed of sound in—and surgical modification of—the cochlea. Carrick Talmadge (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Oxford, MS 38677, clt@olemiss.edu) and Arnold Tubis (Univ. of California, San Diego, La Jolla, CA 92093)

Most analyses of the mechanics of the cochlear to date have involved the assumption that the cochlea is a rigid chamber filled with incompressible fluids. However, *in vivo* measurements of the cochlear fluid pressure or basilar membrane motion necessarily must be performed by first surgically modifying the cochlea through removing part of the cochlear bone. This modification violates the assumption of a rigid chamber, and can significantly modify the mechanics of the cochlea. Because of anatomical constraints, many of these measurements are performed near the base of the cochlea. For the species that are used (guinea pig, gerbil, chinchilla, etc.), measurements near the base necessitate the use of stimulus tones with frequencies that are too high for the assumption of an incompressible fluid to be valid. In this paper, we discuss the modifications to the cochlear mechanics necessary to accommodate both the finite speed of sound and the surgical modification of the cochlea. We show that both of these modifications can easily be incorporated into numerical cochlear model codes. Semianalytic solutions based the WKB method are also discussed.

8:45

3aPP4. Dichotic sound interference products, sound localization cues, and wave interaction at a neural level. Pantelis Vassilakis (Dept. of Ethnomusicology, UCLA, 2539 Schoenberg Music Bldg., Box 951657, Los Angeles, CA 90095-1657)

The possibility of sound interference products arising from neural wave interaction is reexamined through five dichotic experiments. It is argued that interference percepts do not arise unless sound waves interact physically. Experiment 1 examines the possibility of beating for diotic sine stimuli. In this case, diplacusis would introduce beating if neural wave interaction could result in interference products. No beating was observed. Experiments 2a, 2b, and 2c examine the effect of dichotic phase changes on sine stimuli in terms of loudness, direction, and stereo-image-width. It is shown that such phase changes lead to perceptions incompatible with

the interference principle and consistent with claims of binaural phase differences as sound-localization cues. Experiment 3 uses two-component dichotic stimuli. The results indicate that, when two sines with slightly different frequencies are presented dichotically, the constantly shifting phase between them affects a constantly shifting localization of the two-component complex stimulus. Depending on the frequency difference, this shift may be interpreted as a perceived rotation of the sound and/or a timbral fluctuation, often confused with loudness fluctuation. The findings do not support the claim of sound interference products arising from neural wave interaction. They are consistent with sound localization studies and indicate that the dichotic beats reported previously must have been a misidentified rotating sensation.

9:00

3aPP5. Computation and analysis of a nonlinear nonlocal cochlear model with applications to multitone interactions in hearing. Jack Xin (Dept. of Mathematics, Univ. of Texas, Austin, TX 78712), Yingyong Qi (Qualcomm, Inc., San Diego, CA 92121), and Li Deng (Microsoft Res., Redmond, WA 98052)

A nonlinear nonlocal cochlear model of the transmission line type is studied to capture the multitone interactions and tonal suppression effects. The model can serve as a module for voice signal processing. It is a one-dimensional (in space) damped dispersive nonlinear partial differential equation (PDE) based on mechanics and phenomenology of hearing. The elastic damping is a nonlinear and nonlocal functional of basilar-membrane displacement. The initial boundary value problem is solved with a semi-implicit second-order finite difference method. Numerical results are shown on two-tone suppression from both high-frequency and low-frequency sides, consistent with known behavior. Suppression effects among three tones are demonstrated by showing how the response magnitudes of the two fixed tones are reduced as the third tone is varied in frequency and amplitude. Qualitative agreement with existing cat auditory neural data is observed. The model is thus simple and efficient as a processing tool for voice signals. Mathematical analysis of global well-posedness of the model PDE and the existence of multitone solutions will also be shown via the method of *a priori* estimates and fixed point theory. [Work partially supported by ARO and NSF.]

9:15

3aPP6. The effect of middle ear stiffness changes on noise hearing thresholds in presence of a high level background pure tone. Hashir Aazh (P.O. Box 17445_177, Tehran, Iran, hash_aazh@hotmail.com)

In this study the interaction of two different sounds, loud low frequency pure tone and noise in threshold level was evaluated during air pressure changes in order to determine the effects of middle ear stiffness changes on noise threshold, with presence of a high level background pure tone. Forty-two normal hearing subjects after otoscopy, audiometry, and tympanometry participated in this experiment. During the presence of a background pure tone (220 Hz, 90 dB), the thresholds of wide band, low pass filtered, and high pass filtered noises were measured (in dB SPL) in three ear canal air pressure conditions (0, +300, −300 dapa) through a probe set. The pump manometer system of an AZ7 impedance audiometer was used to change the air pressure in an external auditory canal. The means of noise thresholds were significantly greater in 0 dapa than +300 and −300 dapa. But there was no significant difference between noise thresholds in +300 and −300 dapa. By increasing the middle ear stiffness, the transmission of background loud tone will decrease and ear overload is omitted, then the noise hearing thresholds will be better.

9:30

3aPP7. Design of an analog cochlear implant. Joel Flores, Jorge Becerra, and Patricia L. Ramrez (Lab. de Acustica, Unidad Profesional Adolfo Lopez Mateos, ESIME Zacatenco, Av. IPN S/N, Gustavo A. Madero, Lindavista 07738, Mexico, joelflores71@hotmail.com)

The present work describes the development of cochlear implants, an effective methodology in the treatment of severe to profound hearing loss. A group in the acoustic section from ESIME Zacatenco is investigating the development of cochlear implants in Mexico. Cochlear implants are known to have had encouraging results for subjects with postlinguistic deafness; however, one of the great drawbacks of these implants is their cost. For this reason, this work described here counts on using interdisciplinary personnel to accomplish the construction of a cochlear implant of the low-cost analog type, without losing sight of its objective: the best possible restoration of hearing.

WEDNESDAY MORNING, 4 DECEMBER 2002 CORAL GALLERY 1, 8:00 TO 11:40 A.M.

Session 3aSA

Structural Acoustics and Vibration: Energy Flow Methods in Vibroacoustic Analysis and Control I

Stephen A. Hambric, Chair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Chair's Introduction—8:00

Invited Papers

8:05

3aSA1. Computing structural energy density and power flow using spectral elements. José Roberto F. Arruda and Khaled M. Ahmida (Dept. of Computational Mech., Univ. Estadual de Campinas, Cidade Univ. Zeferino Vaz, Campinas, SP 13083-970, Brazil, arruda@fem.unicamp.br)

The large deterministic models obtained using numerical methods are frequently inadequate to deal with structural dynamics problems in the audio frequency range. An alternative approach that has been used for many years is the spectral approach. After transformation to the frequency domain, structural dynamics problems are described by partial differential equations that can be solved analytically in an exact way. In the case of one-dimensional waveguides, the spectral approach may be combined with a mobility approach yielding a spectral element methodology that, in many ways, resembles the finite element method. The major drawback of this approach is the difficulty in treating two- and three-dimensional waveguides. Exact solutions exist only for plates and shells with particular geometry and boundary conditions. An approximate method has been proposed for flat plates with arbitrary

conditions, but the element assembling is still restricted to one dimension. In this paper, different formulations for flat-plate spectral elements are presented and their advantages and drawbacks are discussed. Expressions for the computation of energy densities and power flow in frame and plate spectral elements are derived and some examples are shown to illustrate the advantages of this approach in energy flow analysis applications.

8:30

3aSA2. Spectral finite elements combined with energy flow techniques: Hybrid approach for the energy flow parameters evaluation. Mohammed Ichchou and Louis Jezequel (DSS/LTDS, UMR CNRS 5513, ECL, 69130 Ecully, France)

Energy flow parameters involved in many energy flow techniques (statistical energy analysis, or its local formats) are mainly deduced from pure analytical basic theories. However, in practice, application of SEA or alternatives to complex realistic structures suffers from the choice of subsystems and needed relevant inputs. This paper proposes a procedure which reuses existing reduced finite element modeling of the structural component in view of energy flow parameters identification. Precisely, extraction and analysis of dispersion curves for typical structures, wave numbers, and energy velocities, are first concerns in this contribution. The dispersion curves extraction is based on a spectral problem and uses a reduced finite element model. Properties of eigensolutions of the posed spectral problem are first demonstrated and some remarkable aspects in term of energy flow parameters are discussed. Definitions of energy velocity, needed for modal densities expressions associated with the spectral finite element model, are presented in depth. Ultimately, extension of the proposed formulation, in order to deal with coupled complex structural components is given. This leads to the numerical evaluation of diffusion matrix (reflection and transmission parameters), often needed in order to define coupling loss factor or equivalent energy transfer quantities. Some numerical applications are finally presented.

8:55

3aSA3. Analysis and minimization of power flow in a mechanical vibration isolation system using a hybrid (active/passive) approach. Peter C. Herdic,[a] Brian H. Houston, Robert D. Corsaro, and John A. Judge (Naval Res. Lab, Code 7130, Washington, DC 20375)

Implementation of active control techniques in mechanical vibration isolation systems has been a challenging problem for a number of years where numerous physical control laws have been explored. An energy-based approach to the problem involving the energy transfer or power flow through the mount into the base structure is a first-principles approach to developing control laws and evaluating the system performance. A lumped-parameter model of a passive–active hybrid isolation mount has been developed and validated with experimental data. The mount device has a conventional passive compliant spring, embedded force and velocity sensors, and a piezoceramic actuation layer. This study investigates a complete set of possible layer configurations, that is, the optimal placement of sensors and actuator relative to the passive compliant isolator element. A number of different local physical control laws are examined and the level of power flow through the mount is used to evaluate the performance for the matrix of possible implementations. These results will be discussed with particular emphasis placed on the optimal control configuration and laws, and the related physics. [a] Also with SFA, Inc., Largo, MD 20774.

9:20

3aSA4. Is the modal approach appropriate for analysis of energy flow? Goran Pavic (LVA, INSA-Lyon, 20 Av. A. Einstein, 69621 Villeurbanne, France)

Modal superposition is a most commonly used approach in a numerical analysis of vibration. However, the computation requirements of a typical analysis of energy flow limit the attractiveness of the modal approach because, as a rule, a very large number of modes have to be taken into account in order to produce realistic results. The reason for this particularity is that the energy analysis involves not only vibration displacements but also higher derivatives of these which are contributed by higher modes, the higher the derivative order. More careful analysis of structure-borne vibration shows that the modal truncation is not the only inconvenience where the modal approach is used. An equally important factor limiting its use is the representation of vibration dissipation by modal damping. The paper shows comparisons of computed energy flow in plates using modal and wave approaches. The differences between the two are noticeable, in particular where the vectorial functions of energy flow field, divergence and curl, are concerned. The wave approach to vibration analysis is shown to be more physically consistent than the modal approach.

9:45–10:00 Break

10:00

3aSA5. Summary of energy flow measurements and calculations made on the INCE standard test structures. Stephen A. Hambric (ARL/Penn State Univ., P.O. Box 30, State College, PA 16804, sah@wt.arl.psu.edu)

In 1996 a series of standard test structures was conceived and manufactured by members of the Institute of Noise Control Engineering (INCE) [Cuschieri, Burroughs, and Carroll, *Evaluation of Structure-Borne Noise Prediction Techniques Review*, Proceedings of Noise-Con 98, April 1998, pp. 315–320]. The structures include a Lexan T-shaped beam and two ribbed panels of identical geometries but different materials: aluminum and lexan. In subsequent years, a wide variety of investigators from the U.S. and around

the world have conducted experimental and numerical studies on the test structures, particularly on energy flow parameters, such as power input, power dissipation, and power flow. Most of the studies have been performed at low frequencies, and have shown phenomena such as the conversion of flexural wave power to longitudinal wave power at the *T*-beam joint and the nature of the structural intensity fields in the ribbed panels. Measurements and computations compare well. At high frequencies, Statistical Energy Analysis (SEA) techniques have shown that energy tends to become trapped in the drive leg of the T-beam. SEA studies on the ribbed panels show that the Lexan panel transmits less energy across the ribs than the aluminum panel does due primarily to differences in material loss factor.

10:25

3aSA6. Active minimization of acoustic energy density in enclosed sound fields. Scott D. Sommerfeldt and Benjamin M. Faber (Dept. of Phys., Brigham Young Univ., Provo, UT 84602)

Minimization of acoustic energy density has been investigated for active noise control applications for enclosed sound fields. The standard approach of minimizing the squared acoustic pressure has been shown to often lead to localized control of the sound field, which may be undesirable. It has been shown that minimizing the energy density often leads to improved global attenuation of the field, since quantities dependent on both the pressure and particle velocity are minimized. This penalizes active control solutions in which the pressure is minimized and the particle velocity is increased, as often occurs when simply minimizing the squared pressure. Minimizing the acoustic energy density also requires multiple microphones to be integrated as part of the error sensor. However, in many cases the number of microphones can be reduced by strategically placing the sensor in a location where one component of the particle velocity is assumed to be negligible. Active control results will be shown for an enclosed sound field to compare the reduction obtained using energy density with that obtained for minimizing squared pressure. In addition, the attenuation obtained using a full three-dimensional energy density sensor will be compared with the control achieved using an error sensor with fewer microphones.

10:50

3aSA7. Effects of cross-sectional variation in T-beam structures. Alberto Coronado-Matutti and Rubens Sampaio (Depto. de Eng. Mecanica, PUC-Rio, Rua Marques de Sao Vicente 225, 22453-900 Rio de Janeiro, Brazil)

Structural modification of geometry has been proven to be very effective when it was necessary to improve the dynamic performance of some structures [A. J. Keane, J. Sound Vib. **185**(3), 441–453 (1995)]. This technique has the advantage of achieving good performance even without the heavy viscoelastic treatments or the high-tech, high-cost active vibration control schemes. In this paper a T-beam structure [R. P Szwerc, C. B. Burroughs, S. A. Hambric, and T. E. McDevitt, J. Acoust. Soc. Am. **107**, 3186–3195 (2000)] of uniform cross section is modified in order that the power transmitted to some point, due to an external excitation, is minimized. The modification of geometry is made only in one of the legs of the T-beam. The structure will be modeled combining the finite element method and the spectral element method. The dynamic response of the structures (initial and modified configurations) will be carried between 0–350 Hz; the first ten resonances are placed in this frequency range. After defining a specific region to be optimized via genetic algorithms, the options to minimize the frequency response will be shown; in some cases it is possible to move one or more resonances, in others only a reduction on the response between the nonresonance frequencies is permitted.

11:15

3aSA8. Combining statistical energy analysis and finite element analysis in RESOUND mid frequency vibroacoustic analysis. Bryce K. Gardner, Philip J. Shorter, and Paul G. Bremner (Vibro-Acoust. Sci., 12555 High Bluff Dr., Ste. 310, San Diego, CA 92130)

At low frequencies, vibroacoustic systems exhibit a dynamic response characterized by spatially correlated motion with low modal density. These systems are typically modeled with deterministic methods. While at high frequencies, the dynamic response is characterized by weak spatial correlation and a large number of modes with high modal overlap. These systems are typically modeled with statistical methods. However many vibroacoustic systems have some regions with high modal density and some regions with low modal density. Such systems require a midfrequency solution technique. One such method has been developed based on a hybrid approach combining finite element analysis (FE) in the low modal density regions and statistical energy analysis (SEA) in the high modal density regions. This method is called RESOUND [Langley and Bremner, J. Acoust. Soc. Am. **105**, 1657–1671 (1999)]. Recent developments of RESOUND have focused on predicting the appropriate dynamic interactions and mechanisms for energy flow between the FE and the SEA regions. By including these effects, RESOUND can predict the dynamic response of systems having regions with low modal densities and regions with high modal densities. This paper will provide an overview of recent developments.

Session 3aSC

Speech Communication: Voice, Processing and Phonetic Algorithms (Lecture/Poster Session)

Sergio Suarez Guerra, Cochair
Laboratorio de Sistemas Digitales, CIC-IPN, Instituto de Cibernetica, Matematica y Fisica, ICIMAF, Calle 15 #551, Vedado Cuidad de La Habana, Cuba

Maria E. Hernandez-Diaz Huici, Cochair
CEETI, UCLV, C. Camajuani Km 5 1/2, Santa Clara 54830, Cuba

Diane Kewley-Port, Cochair
Indiana University, 501 North Morton, Bloomington, Indiana 47404

Chair's Introduction—8:30

Invited Papers

8:35

3aSC1. Extraction and visualization acoustic contour of the voice signal, its use as a method for the implementation and rehabilitation of voice and speech. Sergio Suarez Guerra (Laboratorio de Sistemas Digitales, CIC-IPN, Instituto de Cibernetica, Matematica y Fisica, ICIMAF, Calle 15 #551, Vedado Cuidad de La Habana, Cuba)

The possibility to see what is said has resulted in something quite new, as a method for the implementation and rehabilitation of voice and speech. The immediate visualization of the principals acoustic parameters of the voice signal, in a graphic form, associating them with images that represent what was said, resulted in an additional alternative very stimulant in the field of Phoniatry, with applications in Special Schools. In researches done in the Special Schools during one year, results denoted a sustained advance in the learning of correct diction, from students, that in addition to the traditional method used an extraction and visualization acoustic parameters system. This system, Exparam V.2.0, has been proposed for its introduction at a national level into the Special School of the Republic of Cuba and the final statistics of its use will be ready at the end of the course 2002/2003. At the CIC/IPN, Mexico D.F., is at conclusion another system, which will support the Phoniatry clinic consulting room management using voice analysis in the same manner involving the acoustic profile representation and the inclusion of the patient's Clinic data base. This system is being tested at the Human Communication Institute. (To be presented in Spanish.)

9:05

3aSC2. Desirable characteristics in algorithms for measurements of features in pathological voices. Maria E. Hernandez-Diaz Huici (CEETI, UCLV, Santa Clara 54830, Cuba)

In this paper a critical review of some reported algorithms is made, regarding the objective measurement of features in pathologic speech. The discussion is grouped in the following topics: prosody, pith related measurements, hoarseness and vowel/consonant production. The prosody topic includes measurements of timing, stress, tempo and intonation. Pitch related measurements involve the algorithms based on pitch period, hoarseness groups measurements of noise and turbulence related to the pharyngeal functioning, while vowel/consonant production includes formant's transitions, voice onset time and nasality perturbations. Every algorithm was analyzed taking into account the following aspects: Domain of analysis, sampling frequency, sample of speech or voice used, amount of patients and kind of disease, correlation with subjective evaluations. Limitations of analyzed algorithms are shown and the desired characteristics for each topic's measures are proposed. (To be presented in Spanish.)

9:35

3aSC3. Speech intelligibility training using automatic speech recognition technology. Diane Kewley-Port (Commun. Disord. Technol., Inc. and Indiana Univ., 501 N. Morton, #215, Bloomington, IN 47404, kewley@indiana.edu), Jonathan Dalby, and Deborah Burleson (Commun. Disord. Technol., Inc., Bloomington, IN 47404)

This paper describes two computer-based speech training systems that have been developed over the past decade of research. This research began at Indiana University and has been continued and commercialized at Communication Disorders Technology, Inc. The Indiana Speech Training Aid (ISTRA) uses a speaker-dependent speech recognizer and is designed for use by speech-language pathologists treating clients with articulation disorders of various sorts. The HearSay system employs a speaker-independent recognizer and is designed to improve the segment-level intelligibility of learners of English as a second language. Both systems employ minimal-pairs contrast training and share the underlying assumption that effective speech training requires intensive individual practice accompanied by evaluative feedback. The key technological achievements of ISTRA and HearSay are the methods employed

for providing valid feedback on the intelligibility of the client's speech. This paper will describe these methods and report the results of new training studies that demonstrate that computer-based speech training can be an effective method for improving the English speech intelligibility of learners of English as a second language. [Work supported by Grants No. DC00893 and No. DC02213 from the NIH–National Institute on Deafness and Other Communication Disorders.]

10:05

3aSC4. String kernels for the classification of speech data. John Ch. Goddard Close, Fabiola M. Martinez Licona, Alma E. Martinez Licona (Dept. of Elec. Eng., Universidad Autónoma Metropolitana Iztapalapa, México City 09340, Mexico), and H. Leonardo Rufiner (Universidad Nacional Entre Ríos, Oro Verde, Entre Ríos 3100, Argentina)

Support Vector Machines (SVM) have been applied to a wide variety of classification problems with excellent results. This has to do with their provable generalization ability derived from Statistical Learning Theory. Recently specialized kernels, such as the Fisher kernel and string kernels, have been introduced in an attempt to apply the same SVM framework to sequential data. Notable results have been obtained on classification tasks related to biosequences and text documents showing that the specialized kernels may provide a viable and interesting alternative to other classifiers, such as those using Hidden Markov Models. String kernels are particularly attractive because of their conceptual simplicity and they also furnish insight into the task of sequential data classification. In the present paper string kernels are applied to a new application area, that of automatic speech recognition. In particular, different string kernels are tested on the task of phoneme recognition and the results obtained are compared with those from several other common classifiers. This comparison reveals the potential of string kernels as a simple and feasible alternative to other established methods for certain speech recognition tasks. [Work supported by CONACYT under Project 31929-A.]

Contributed Papers

All posters will be on display from 10:35 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 10:35 a.m. to 11:15 a.m. and contributors of even-numbered papers will be at their posters from 11:15 a.m. to 12:00 noon.

3aSC5. Technical considerations in the design of a wearable voice dosimeter. Peter S. Popolo, Karen Rogge-Miller, Jan G. Svec, and Ingo R. Titze (Wilbur James Gould Voice Ctr., The Denver Ctr. for the Performing Arts, 1245 Champa St., Denver, CO 80206, ppopolo@dcpa.org)

This paper deals with the technical issues involved in the design of a voice dosimeter, a wearable device to be used by teachers to measure vocal dose on the job, at home and elsewhere during the total waking hours of each day. The concept of vocal dose is introduced and definitions of various dose measures are presented. Practical issues of the design are discussed, including software development, selection of an external transducer, and devising a comfortable, reliable and repeatable means of attaching the transducer to the subject. The dosimeter software was designed to calculate SPL (sound pressure level), $F0$ (fundamental frequency) and voicing time. In addition, an interactive portion was developed that prompts the user to verify the validity of the stored data every two hours, and to perform a rating of vocal effort and quality of soft phonation. The SAL (surface acceleration level) to SPL relation for the prototype dosimeter was determined. The prototype has been field-tested, and SPL, $F0$, and voicing time data have been analyzed. A sensitivity analysis shows that small errors in SPL have a large effect on the accuracy of dose calculations, while $F0$ errors have less of an effect.

3aSC6. Neural-network-based voice-tracking algorithm. Mary Baker, Charise Stevens, Brennen Chaparro (Dept. of Elec. Eng., Texas Tech Univ., M.S. 3102, Lubbock, TX 79409, mcb@coe.ttu.edu), and Dwayne Paschall (Texas Tech Univ., Lubbock, TX 79409)

A voice-tracking algorithm was developed and tested for the purposes of electronically separating the voice signals of simultaneous talkers. Many individuals suffer from hearing disorders that often inhibit their ability to focus on a single speaker in a multiple speaker environment (the cocktail party effect). Digital hearing aid technology makes it possible to implement complex algorithms for speech processing in both the time and frequency domains. In this work, an average magnitude difference function (AMDF) was performed on mixed voice signals in order to determine the fundamental frequencies present in the signals. A time prediction neural network was trained to recognize normal human voice inflection patterns, including rising, falling, rising–falling, and falling–rising patterns. The neural network was designed to track the fundamental frequency of a single talker based on the training procedure. The output of the neural network can be used to design an active filter for speaker segregation. Tests were done using audio mixing of two to three speakers uttering short phrases. The AMDF function accurately identified the fundamental frequencies present in the signal. The neural network was tested using a single speaker uttering a short sentence. The network accurately tracked the fundamental frequency of the speaker.

3aSC7. Diagnostic system for speech articulation and speech understanding. Andrzej Czyzewski, Bozena Kostek, and Henryk Skarzynski (Inst. of Physiol. and Pathol. of Hearing, Pstrowskiego 1, Warsaw, Poland)

With the increase in access to multimedia computers, speech training can be made available to patients with no continuous assistance required from speech therapists. Another function the system can easily perform is screening testing of speech fluency providing directed information to patients who have various speech disorders and problems with understanding speech. The idea underlying the proposed system is a programmed speech therapy training algorithm consisting of diagnostic tools and rehabilitation devices connected with it. The first function the system described in the paper has to perform is data acquisition where information about the patient's medical history is collected. This is done through electronic questionnaires. The next function is analysis of the speech signal articulated by the patient when prompted by the computer followed by some multimedia tests carried out in order to assess the subject's ability to understand speech. Next, the results of the electronic questionnaire, the patient's voice, and the patients' reactions are automatically analyzed, based on that the system automatically diagnoses possible speech disorders and how strong they are. A large number of schoolchildren were tested with this method. In the paper foundations of applied speech testing method and obtained results will be demonstrated.

3aSC8. Choice of speech features for single-channel vibrotactile stimulators. Alessio Trindade Barros, Raimundo Carlos S. Freire, and Gurdip Singh Deep (Dept. of Elec. Eng., Federal Univ. of Paraiba, Av Aprigio Veloso, 882, Campina Grande, PB, Brazil)

The choice of speech features for coding of a single-channel vibrotactile stimulation signal was analyzed to obtain a better discrimination between phonemes that compose each viseme of the Portuguese language. Two different speech features were selected for comparison, based on previous investigations [I. R. Summers and D. Gratton, IEEE Trans. Rehab. Eng. **3**, 117–121 (1993)]: the amplitude envelope—AE and the zero-crossing frequency—ZRX of the speech signal. A codification scheme (PRP) was proposed, with output based on the switching between AE and ZRX, depending on the value of ZRX. The selected voice items were consonants of each viseme, in the context vCv: ata/ada/ana, apa/aba/ama, aka/aga/alha, asa/aza, afa/ava, acha/aja, and ala/ara, for different combinations of vowels. Histograms for each feature were done for each consonant of a viseme, and were compared in the context of each viseme. It was observed that the output of PRP was the same as AE for most of the consonants, but different for fricatives. PRP allowed better discrimination than AE and ZRX alone. The difficulty of single-channel stimulator with discrimination of fricatives is known in the literature, and the proposed scheme showed better results for this work. [Work supported by CAPES.]

3aSC9. Speech coding for fast speech commands recognition. Abel Herrera and Arturo Gardida (Universidad Nacional Autonoma de Mexico (UNAM), Electrica, 2o piso, Edificio Posgrado, Fac. Ingenieria, Circuito Exterior, CU, DF, 04510, Mexico)

Isolated speech recognition could be more appropriate than continuous speech recognition systems for some word command applications. In this paper, some speech coders are explored that can be useful for fast word command recognition using low memory storage. First, decimated critical bands vectors by frame are used as coding features. The accuracy is around 94% for a set of ten commands; they are used at first step. After that, using a reduced DTW classification scheme, the accuracy for a small subset of possible candidates is above 99%. Another coding scheme is using the Karhunen–Loeve Transform (KLT) combined with multisection vector quantization. The KLT is not a fast transform; however, used at the subword level it gives a fast recognizer with accuracy around 99%. The classifiers in both cases are fast, around a second, and require low memory storage.

3aSC10. Adapted speech recognizer to improve English pronunciation of Hispanics. Maria A. Garcia and Rene O. Arechiga (UAM-A, Depto. Electronica, Av. San Pablo #180 Col Reynosa Tamaulipas, Azc. Cp. 02200 Mexico D.F., arechiga@correo.azc.uam.mx)

This work presents advances on an adapted speech recognition system, based on hidden-Markov models (HMMs), to help Hispanics in their pronunciation of English. The main results so far, with English digits, show significant recognition improvements of the adapted recognizer over the not adapted one, both for one speaker (100% vs 94%) and for groups of them (96% vs 74%). The adaptation uses maximum likelihood linear regression (MLLR). Two databases were used in the experiments. The TIMIT database to train the recognizer models with English native speakers, and the Latin-American Spanish database to adapt and test the adapted and not adapted recognizers. The project, of which this work is a part, is an English pronunciation tutor, whose main parts are this adapted speech recognizer, a pronunciation evaluator (phonetic and prosodic) and a dialog manager. English speech recognition systems are trained with native speakers, and most of the recognition errors are attributed to the interference of foreign accents. This is why it is necessary to adapt the models to take into account the characteristic features of a given population; in this case Hispanics. (To be presented in Spanish.)

3aSC11. Phone inventory optimization for multilingual automatic speech recognition. Lynette Melnar and James Talley (Motorola Labs, 7700 W. Parmer Ln., MD: PL26, Austin, TX 78729)

This paper describes a phone inventory optimization procedure for application in multilingual automatic speech recognition (ASR). The optimization procedure is based on three knowledge sources that act collectively to guide phonological reduction and selection processes: (1) abstract (language-independent) phonological universals and tendencies that are used in the construction of a hierarchical structure that specifies phone class reduction paths; (2) language-dependent knowledge that includes information of the targeted languages' phone inventories and individual phone frequencies in language data resources; (3) acoustic data that provides phone discriminability and similarity metrics. Using the optimization procedure, the phone inventories of six languages, American English, Mandarin Chinese, Egyptian Colloquial Arabic, Japanese, German, and Spanish, were merged to create an inventory consisting of 64 distinct cross-phonological units. This reduced phone set was used in all training and testing procedures and resources for the recognition of the six targeted languages. Preliminary recognition results are very encouraging: while purely data-driven approaches to multilingual ASR fail to reach word-recognition rates comparable to monolingual applications, the use of the optimized phone inventory in our multilingual ASR program yields recognition rates approximating that of monolingual ASR.

3aSC12. Energy redistribution speech intelligibility enhancement, vocalic and transitional cues. John G. Harris and Mark D. Skowronski (Computational Neuro-Eng. Lab., Univ. of Florida, Gainesville, FL 32611, harris@cnel.ufl.edu)

Two novel speech enhancement algorithms are presented that automatically increase intelligibility in noisy environments while maintaining the signal power and naturalness of the original speech. These energy redistribution (ER) algorithms move signal energy to targeted regions of relatively high information content that are crucial for intelligibility. The boosted regions are originally of low energy and therefore usually the first segments lost with the addition of environmental noise. The ER voiced/unvoiced (ERVU) method transfers energy from voiced speech to regions of unvoiced speech, while the ER spectral transition (ERST) method moves energy from spectrally stationary regions to spectrally transitional regions. Hand-held cell phones and public address systems are expected to be the dominant applications for these techniques. Standard noise reduction methods such as spectral subtraction are assumed to have already been applied to the voice signal before broadcast. Using human listening tests, it was found that both algorithms boost the intelligibility of speech in noisy environments by nearly 7% over the original unprocessed signals, without degrading naturalness or increasing signal power. Furthermore, both algorithms allow for controlling the trade-off between boost gain and speech naturalness.

3aSC13. The effect of speech compression on second language speech perception learning. Takahiro Adachi (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan and The Grad. Univ. for Advanced Studies, tadachi@atr.co.jp) and Tsuneo Yamada (Natl. Inst. of Multimedia Education, Chiba 261-0014, Japan)

In recent years, second language education is sometimes conducted over the Internet. However, the speech used for training is often compressed in order to decrease the number of packets transmitted, without an evaluation of whether the compression causes a decrease in intelligibility. The current study assessed the intelligibility of compressed speech in order to determine whether it is appropriate for use in Internet-based training. English words minimally contrasting in /r/ and /l/, /b/ and /v/, or /s/ and /th/ were recorded from native speakers of English and compressed using several different methods (MP3, WMA, QCELP, ADPCM, etc.). The intelligibility of these speech files was then evaluated by a group of native English speakers and a group of native Japanese speakers, using a minimal-pair identification task. Results showed that: (1) intelligibility

scores for Japanese speakers were lower than those for the English speakers, as was expected; and (2) intelligibility of /r/–/l/ stimuli was unaltered by the compression, whereas intelligibility of /s/–/th/ stimuli dropped remarkably in several formats, and that of /b/–/v/ stimuli dropped only in QCELP-compressed words. These findings suggest that some compression formats are not appropriate in speech perception training. Implications for Internet educational materials will be discussed.

3aSC14. Invariance of lognormal distribution parameters in speech segment durations. Kristin Rosen and Raymond Kent (Waisman Ctr., Univ. of Wisconsin, 1500 Highland Ave., Madison, WI 53705)

This study re-examines speech segment duration using lognormal distributions and presents a basis for the unification of many previous measurements in connected speech. Distributions of speech segment length in previous studies (Pols, Wang, and ten Bosch, 1996; Crystal and House, 1982, 1988; Duez–Goldman–Eislinger, 1968) were replotted onto lognormal cumulative plots. Speech segments were consistently symmetric on the logarithmic scale and well characterized by the lognormal distribution, with the exception of fricatives. Analysis with respect to the logarithmic scale reveals surprising consistencies within categories. Specifically, geometric standard deviation appears to be essentially preserved across phonemic categories. The nature of the distributions themselves suggests possible clues to the understanding or modeling of the underlying timing mechanisms, and provides a foundation for meaningful comparisons in connected speech. Possible advantages of logarithmic-based measures (i.e., geometric mean, geometric standard deviation) over linear measures (e.g., coefficient of variability) are presented.

3aSC15. Quantifying nonverbal communicative behavior in face-to-face human dialogues. Mustapha Skhiri (Dept. of Computer and Information Sci. (IDA), Linköping Univ., SE-581 83 Linköping, Sweden, and Grad. School of Lang. Technol. (GSLT), Sweden, mussk@ida.liu.se and Loredana Cerrato (Royal Inst. of Technol., Stockholm, Sweden)

The referred study is based on the assumption that understanding how humans use nonverbal behavior in dialogues can be very useful in the design of more natural-looking animated talking heads. The goal of the study is twofold: (1) to explore how people use specific facial expressions and head movements to serve important dialogue functions, and (2) to show evidence that it is possible to measure and quantify the entity of these movements with the Qualisys MacReflex motion tracking system. Naturally elicited dialogues between humans have been analyzed with focus on the attention on those nonverbal behaviors that serve the very relevant functions of regulating the conversational flux (i.e., turn taking) and producing information about the state of communication (i.e., feedback). The results show that eyebrow raising, head nods, and head shakes are typical signals involved during the exchange of speaking turns, as well as in the production and elicitation of feedback. These movements can be easily measured and quantified, and this measure can be implemented in animated talking heads.

WEDNESDAY MORNING, 4 DECEMBER 2002 CORAL GARDEN 2 AND 3, 8:30 TO 11:30 A.M.

Session 3aSP

Signal Processing in Acoustics and Underwater Acoustics: Acoustic Automatic Target Recognition I

Ning Xiang, Cochair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Armando Andrade, Cochair
Calle H. No. 11, Col. Jardines de Santa Clara, Ecatepec Edo. De. CP 55450, Mexico

Chair's Introduction—8:30

Invited Papers

8:35

3aSP1. Automatic target recognition in acoustics: An overview. John R. Sacha (Appl. Res. Lab., The Penn State Univ., P.O. Box 30, State College, PA 16804-0030, jrs9@psu.edu)

Automatic target recognition (ATR) constitutes one of the major uses for acoustical signal processing. ATR is employed in manned systems for operator workload reduction and performance improvement, as well as in autonomous applications. An overview of some of the major components involved in the architecture of such systems is provided. Feature extraction is the most critical step of ATR and is necessarily application specific. Generic feature selection and ranking methods are presented, including heuristic search and information–theoretic measures. Basic pattern recognition definitions and techniques are reviewed. Commonly used classification paradigms include classical statistical formulations, both parametric and nonparametric, and neural nets; support vector machines and nonmetric methods such as decision forests are some alternative techniques that have received recent attention. A few practical issues often encountered when constructing recognition systems, including training data requirements, ground truth labeling, and performance evaluation methodologies and metrics, are also addressed.

9:05

3aSP2. Advances in algorithm fusion for automated sea mine detection and classification. Gerald J. Dobeck and J. Tory Cobb (Naval Surface Warfare Ctr., Coastal Systems Station, Dahlgren Div., Panama City, FL 32407-7001, DobeckGJ@ncsc.navy.mil)

Along with other sensors, the Navy uses high-resolution sonar to detect and classify sea mines in mine-hunting operations. Scientists and engineers have devoted substantial effort to the development of automated detection and classification (D/C) algorithms for these high-resolution systems. Several factors spurred these efforts, including: (1) aids for operators to reduce work overload; (2) more optimal use of all available data; and (3) the introduction of unmanned minehunting systems. The environments where sea mines are typically laid (harbor areas, shipping lanes, and the littorals) give rise to many false alarms caused by natural, biologic, and manmade clutter. The objective of the automated D/C algorithms is to eliminate most of these false alarms while maintaining a very high probability of mine detection and classification (PdPc). In recent years, the benefits of fusing the outputs of multiple D/C algorithms (Algorithm Fusion) have been studied. To date, the results have been remarkable, including reliable robustness to new environments. In this paper a brief history of existing Algorithm Fusion technology and some techniques recently used to improve performance are presented. An exploration of new developments is presented in conclusion.

9:35

3aSP3. A pulse length tolerant neural network-based detector for sector-scan sonar. Stuart W. Perry (Maritime Operations Div. Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont, NSW 2009, Australia, Stuart.Perry@dsto.defence.gov.au) and Ling Guan (Ryerson Polytechnic Univ., Toronto, ON M5B 2K3, Canada)

In this paper we present a neural network-based system to detect small manmade objects in sequences of sector-scan sonar images created using signals of various pulse lengths. The sonar system considered has three modes of operation to create images over ranges of up to 800 m using acoustic pulses of different durations for each mode. After initial cleaning and segmentation to extract objects, features are computed from each object. These features consist of basic object size and contrast statistics, shape moments, moment invariants, and features derived from the second-order histogram of each object. Optimal sets of 15 features from the total set of 31 are chosen using sequential feature selection techniques. Using these features a neural network is trained to detect manmade objects in any of the three sonar modes. The proposed detector is shown to perform very well when compared with detectors trained specifically for each sonar mode and a number of statistical detectors. The proposed detector achieves a 92.4% detection probability at a mean false alarm rate of 10 per frame averaged over all sonar mode settings. Finally, research into Recurrent Neural Network detectors is described and shown to further improve performance.

10:05

3aSP4. Model-based acoustic characterization and classification of irregular-shaped targets: Application to fisheries and zooplankton acoustics. Dezhang Chu, Timothy K. Stanton, and Peter H. Wiebe (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Acoustic scattering by fish and zooplankton is a complicated function of the geometrical and physical properties of the targets, as well as the environmental and sonar system parameters. The shape and anatomy of zooplankton vary significantly from taxa to taxa and their dominant scattering mechanisms can be completely different. As a result, the acoustic classification of such targets is extremely difficult and often nonunique. To reduce the ambiguity and nonuniqueness, a number of model-based methods are presented. These methods use the temporal, spatial, spectral, and statistical signatures of acoustical scattering signals and can be applied to a variety of acoustic systems, including narrow-band, broadband, and multifrequency systems. The methods also depend strongly on whether or not the targets are resolved. Individual targets with different shapes and material properties have their unique characteristics and can be classified acoustically in terms of their size, orientation, scattering mechanisms, as well as their material properties. Results of applying these methods to the laboratory and field data will be presented and analyzed. [Work supported by ONR, NSF, and the Comer Science and Education Foundation.]

10:35–10:45 Break

Contributed Papers

10:45

3aSP5. A neural network approach for data inversion, application to the characterization of marine sediments, and sonar targets. Manell E. Zakharia (IRENAV, Res. Inst. of the French Naval Acad., BP 600 F29240, Brest Cedex, France)

The paper presents an innovative method for inversion based on the use of neural networks. The inversion procedure will be split in three major steps: the understanding and enhancing of the physical phenomena involved, the learning of the physics from extracted parameters, and the inversion itself. This separation allows a very convenient distribution of computation load: the first two steps (which are the most demanding in terms of computation time) can be done in a prior phase, while the third one (which becomes very simple) can be implemented on-line (real time). Another major advantage of the proposed method (with respect to conventional approaches such as conjugate gradient) is that it requires only poor *a priori* information thanks to a judicious choice of the characteristic physical phenomena and the selection of relevant signal processing tools for parameter extraction. Two examples of the proposed procedure are described: the inversion of geoacoustic parameters of seabed and the characterization of sonar targets. The inversion is applied to both simulated and experimental data sets. In both cases, the training will be achieved on simulated data and then applied to the experimental.

11:00

3aSP6. Real time identification of several targets by iterative time reversal. Gabriel Montaldo, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, Universit Paris VII, ESPCI, CNRS UMR 7587, Paris, France)

Finding punctual targets inside an aberrating medium is a problem of interest in nondestructive testing and medical research. Previous works have shown that the iteration of the time reversal method allows us to identify the strongest reflector of the medium in real time but the identification of several targets is not possible with this technique. Using the idea that each strong scatterer is associated to an eigenvalue of the time reversal operator, a method has been developed (the DORT method) to identify punctual diffusers in a medium. This technique requires the experimental acquisition of the whole time reversal operator and its numerical decomposition implies many calculations. We propose here a way to identify in real time all the punctual targets by modifying the time reversal iterative method. In a first step, we identify the strongest reflector by time reversal processing. In a second step, we subtract its contribution in order to find the second reflector contribution and so on. This method allows us to build an ''eigen pulse'' able to focus a pulsed signal on each target. We demonstrate experimentally its real time feasbility by identifying several targets inside a strongly aberrating medium.

11:15

3aSP7. Ultrasonic sensor system to detect solids in a milk pasteurization process. Carlos Barroeta Z., Fernando L. Sanchez M., G. Moreno Fernando R., and Laura Montes P. (Control Lab., ESIME, IPN, Mexico, cbarroet@att.net.mx)

In the food industry, many products require a specific process. In the milk industry, the raw milk passes through several process stages before reaching the end user in a very qualitative and healthy way. One of the problems of the milk is that it can contain solids in suspension, result of contamination of the milk, or inherent to the pasteurization process itself. In order to control these solids, a solid detection system is being developed, which will detect the solids by the reflection and refraction of ultrasonic waves. The sensor must be set in the upper part of the milk containers, and with a grid array to allow the control system to prevent these solids from entering into the pipes of the processing plant. The sensing system may activate an acoustic alarm to indicate that a solid has been detected, and a visual one to indicate the affected part of the process. (To be presented in Spanish.)

WEDNESDAY MORNING, 4 DECEMBER 2002 GRAND CORAL 1, 7:50 A.M. TO 12:00 NOON

Session 3aUW

Underwater Acoustics and Acoustical Oceanography: Littoral Environmental Variability and Its Acoustic Effects I

Finn B. Jensen, Cochair
SACLANT Undersea Research Center, Viale San Bartolomeo 400, 19138 La Spezia, Italy

James F. Lynch, Cochair
Woods Hole Oceanographic Institution, 203 Bigelow Building, Woods Hole, Massachusetts 02543

Chair's Introduction—7:50

Invited Papers

7:55

3aUW1. Results from a series of acoustic variability experiments. Martin Siderius (SAIC, 1299 Prospect St., La Jolla, CA 92037, sideriust@saic.com), Peter Nielsen (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy), and Jurgen Sellschopp (FWG, 24148 Kiel, Germany)

A series of experiments were conducted in recent years by the SACLANT Undersea Research Centre to measure the effect of the environment on acoustic propagation. These broadband, shallow water experiments used a moored vertical array with both towed and fixed sound sources. During the fixed experiments the sound source was mounted on a steel frame tower that was resting on the seabed. This design was used to isolate the acoustic fluctuations caused by the environment from those due to source motion. During transmissions extensive measurements of the environment were made including range-depth profiling of the ocean sound speed along the propagation track using a vertically towed CTD (conductivity, temperature, and depth) chain. The seabed types at the sites varied from soft clay (sound speed less than in the water) to a highly reflective sand-gravel (sound speed about 1750 m/s). The temporal fluctuations and spatial structure of the received acoustic signals varied greatly between the sites. This had a big impact on propagation modeling for transmission loss; matched-field processing; and model-based, geo-acoustic inversion. In this paper an overview of these acoustic variability experiments will be described along with results showing the effect of the environment on model-based processing.

8:15

3aUW2. Acoustic propagation in the coastal environment of the Florida Straits—recent experimental results. Harry Deferrari, Neil Williams, and Hien Nguyen (RSMAS, Univ. of Miami, Miami, FL 33138)

An autonomous source was moored at ranges of 10 and then 20 km from a vertical receiver array with 32 elements in a depth of 145 m of water. *M* sequences were transmitted for 28 days at six center frequencies from 100 to 3200 in one octave increments. Arrivals and paths are identified with models and then fluctuation statistics, coherence, and predictability are examined in a parameter space of frequency, range, and receiver depth. A group of refracted-bottom-reflected (RBR) modes/rays has nearly equal group

velocities and tends to focus in time and depth forming intense arrivals especially at the depth of the transmitted. A second group of surface reflected bottom reflected (SRBR) modes produce arrivals that fan out in time. Coastal areas inside western boundary currents have exceptionally variable sound speed fields owing to dynamical effects such as meanders, shelf waves, eddies, coastal upwelling and energetic internal waves and tides. Sound speed fluctuations are observed to be an order greater than the deep ocean. Very large changes in mean sound speed profiles and extreme gradients occur at subinertial periods. Also, potential energy of the internal wave field varies with the same longer periods as do statistical properties of observed acoustic signals.

8:35

3aUW3. Modeling the interaction of acoustic and internal wave fields in shallow-water environments. Steven Finette (Acoust. Div., Naval Res. Lab., Washington, DC 20375)

Over the past decade, there has been considerable interest in the effect of solitary wave packets on the acoustic field. This talk will present an overview of modeling efforts to describe the interaction of these internal gravity waves with acoustic field propagation in shallow-water waveguides. The emphasis will be on the interaction of solitary wave packets with low-frequency acoustic fields propagating in a continental shelf/slope environment. Topics to be discussed will include both adiabatic and mode-coupling effects, differences between 2-D and 3-D propagation, as well as the influence of wave packets on horizontal array beamforming. Most research to date involves solitary wave propagation over flat bathymetry, but does not address the generation of internal tides over variable bathymetry and their subsequent evolution into wave packets. In this regard, recent nonlinear fluid dynamic simulations of both the generation and propagation of internal tides and wave packets are illustrated and their influence on acoustic propagation briefly addressed. [Work supported by ONR.]

8:55

3aUW4. Rapid fluctuations of the channel impulse response at midfrequencies in shallow water. W. S. Hodgkiss, W. A. Kuperman, and D. E. Ensberg (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0701, wsh@mpl.ucsd.edu)

A fixed source, fixed receiving array experiment was carried out to measure the stability of forward transmissions in shallow water ($\sim$100 m deep) over a 6 km forward propagation path off San Diego, CA. The source was moored 6 m above the seafloor and the 12 aperture, 64-element vertical receiving array was deployed with the lowest element 4 m off the bottom. The source transmissions of interest here are the 2 kHz bandwidth, 1 s duration FM chirps which were transmitted continuously for 5 min at a time and have been matched filtered to yield the channel impulse response. In addition to CTDs taken in the region between the source and receiving array, a thermistor string at the receiving array site provided measurements of water column temperature fluctuations. The time-evolving structure of the channel impulse response clearly shows significant, environmentally induced fluctuations which also are evident in an arrival angle vs travel time spatial decomposition at the array. [Work supported by ONR.]

Contributed Papers

9:15

3aUW5. High frequency sonar variability in littoral environments: Irregular particles and bubbles. Simon D. Richards (QinetiQ, Winfrith Technol. Ctr., Dorchester, Dorset DT2 8XJ, UK), Timothy G. Leighton, and Paul R. White (Univ. of Southampton, Highfield, Southampton SO17 1BJ, UK)

Littoral environments may be characterized by high concentrations of suspended particles. Such suspensions contribute to attenuation through visco-inertial absorption and scattering and may therefore be partially responsible for the observed variability in high frequency sonar performance in littoral environments. Microbubbles which are prevalent in littoral waters also contribute to volume attenuation through radiation, viscous and thermal damping and cause dispersion. The attenuation due to a polydisperse suspension of particles with depth-dependent concentration has been included in a sonar model. The effects of a depth-dependent, polydisperse population of microbubbles on attenuation, sound speed and volume reverberation are also included. Marine suspensions are characterized by nonspherical particles, often plate-like clay particles. Measurements of absorption in dilute suspensions of nonspherical particles have shown disagreement with predictions of spherical particle models. These measurements have been reanalyzed using three techniques for particle sizing: laser diffraction, gravitational sedimentation, and centrifugal sedimentation, highlighting the difficulty of characterizing polydisperse suspensions of irregular particles. The measurements have been compared with predictions of a model for suspensions of oblate spheroids. Excellent agreement is obtained between this model and the measurements for kaolin particles, without requiring any *a priori* knowledge of the measurements.

9:30

3aUW6. Sediment tomography in the East China Sea: Compressional wave speed and attenuation inversions from Airy phase dispersion measurements and time series correlation. Colin Lazauski, James Miller, Gopu Potty, Chuen-Song Chen (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882), and Peter Dahl (Univ. of Washington, Seattle, WA 98105)

This paper discusses ongoing data analysis results from the acoustic bottom interaction experiment conducted in May–June 2001 in the East China Sea as part of the Asian Seas International Acoustics Experiment (ASIAEX-2001). Using time-frequency scalograms of broadband signals, the modal arrivals and group speed minimums (Airy Phase) of several modes are clearly observed. The structure of the Airy Phase signal is used to match the dispersion curves which forms the basis of this inversion technique. Utilizing the Airy Phase group speed minimums and corresponding pressure amplitudes of each observable mode, the sediment compressional wave speed and attenuation as a function of depth are derived. The group speed minimum for each mode provides additional information on the compressional wave speed in the modal sediment depth penetration interval. To refine the sediment parameters, synthetic and measured time series are correlated for goodness of fit and used in the inversion process. The synthetic time series is generated from the scalogram values corresponding to the times and frequencies of the calculated dispersion curves. Inverted speeds and estimated modal penetration depths are then used to develop the sediment profile. The estimated resolution is dependent on the number and frequency span of the observable modes. Estimated sediment properties from several areas are presented with verification from coring results. [Work supported by ONR.]

9:45

3aUW7. Inversion of geoacoustic parameters in the South China Sea. Gopu R. Potty, James H. Miller, Colin J. Lazauski (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882), James F. Lynch, and Arthur Newhall (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

During the ASIAEX Experiment in the South China Sea, light bulb sources were deployed in the vicinity of the receive array in order to provide array element localization. We present inversions for bottom geoacoustic properties using these broadband light bulb data. The light bulbs imploded about 40 s after launch and were deployed approximately 1 km from the receive array. The receive array was at a depth of 125 m of water. The ground truth measurements include bathymetry data from ship board measurements, chirp sonar data, and sediment cores. SeaSoar surveys and CTD measurements provide sound-speed data in the water column. The inversion is carried out using a coherent broadband matched field processing technique [Chapman *et al.,* Oceans 97, MTS/IEEE Conference Proceedings, Vol. 2, pp. 763–768]. The thickness of the top layer of the sediment and the compressional wave speeds in the top and bottom layers are estimated in the vicinity of the receive array and compared with ground truth measurements. [Work supported by ONR.]

10:00–10:15 Break

10:15

3aUW8. Horizontal wavenumbers from a fluctuating waveguide. Robert Field, Joal Newcomb, James Showalter, and Jacob George (Naval Res. Lab., Stennis Space Center, MS 39529)

From September 26 through November 2, 2000 the U.S. and Japan conducted a joint experiment in the New Jersey Bight off the U.S. East Coast. The experiment was conducted in three legs, the first of which investigated low frequency acoustic fluctuations in the presence of a shelf/slope front and internal waves. Legs II and III investigated acoustic scattering at 5.5 kHz and methods of acoustic inversion for sea floor properties, respectively. In this paper we will focus on the experimental results of leg I and discuss implications for doing acoustic inversions in a strongly fluctuating medium. Theoretical results based on the Pekeris waveguide show that a large-scale, slowly fluctuating sound speed can cause rapid fluctuations in the acoustic field in the form of acoustic harmonics. Pekeris waveguide calculations are compared with a subset of the acoustic data. Theoretical results and parabolic equation simulations are used to show the impact of the water column sound speed on the horizontal wavenumbers and hence on acoustic inversion methods that utilize them. [Work supported by ONR, Program Element PE62435N.]

10:30

3aUW9. A broadband ocean sediment acoustics model for signal processing applications. Nicholas P. Chotiros and Marcia J. Isakson (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, misakson@arlut.utexas.edu)

It has been shown that fluid and viscoelastic solid approximations cannot accommodate the observed sound speed dispersion and enhanced reflection loss over sandy shallow water sediments. A plausible poroelastic model has been developed for the high frequency band (>50 kHz) using measurements from several sources. This model, with constant coefficients, is unable to track the observed sound speed dispersion at lower frequencies. It is hypothesized that one parameter, the frame bulk modulus, varies with frequency in a relaxation process associated with squirt flow at the grain–grain contact. This hypothesis has the potential to be a critical component in broadband acoustic models of granular ocean sediments. It will link measurements at high frequencies to propagation modeling at low frequencies, provide accurate, physics based, models of propagation loss, and a means to invert for bottom properties over a broad range of frequencies. [Work supported by ONR, Undersea Signal Processing.]

10:45

3aUW10. Coupled electrokinetic-Biot theory and measurement techniques in sediment acoustics. Gareth Block and Nicholas Chotiros (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, gblock@arlut.utexas.edu)

Biot theory will no doubt play an important role in modeling wave propagation in ocean sediments. One prediction of Biot theory is the existence of the Biot slow wave, described by out-of-phase motion of the pore fluid and porous grain structure. Unfortunately, slow wave motion has been extremely difficult to verify using acoustical transducers. Electrokinetic (EK) transduction techniques shed new light on this effort. EK parameters, such as the surface chemistry of the grains and pore fluid salinity, are coupled to Biot theory at the microscopic level, and therefore provide another method of determining the microstructural properties of sediments. One application of this is the electrokinetic transmission of acoustic pressure waves. This occurs when an applied voltage drives ionic (and thus fluidic) currents in electrolyte-saturated porous media. Experimental and theoretical results for electrokinetic motion in sediments will be reported. [Work supported by ONR, Ocean Acoustics.]

11:00

3aUW11. Environmental influences on the frequency dependence of effective bottom attenuation. James D. Nickila (Adv. Sonar Technol. Div. Naval Undersea Warfare Ctr., Newport, RI 02841) and Kevin B. Smith (Naval Postgrad. School, Monterey, CA 93943)

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 correlated with results from similar calculations without environmental variability but with attenuation as a free parameter used to maximize the correlation. By performing such comparisons over a band of frequencies, the effective frequency dependence of the bottom attenuation can be determined. The dominant environmental influences will then be identified and quantified.

11:15

3aUW12. Sparse acoustic gridding in an azimuthally complex environment. Erik R. Rike and Donald R. DelBalzo (Naval Res. Lab., Stennis Space Center, MS 39529)

A very sparse acoustic gridding method, the Objective Grid/Radials using Environmentally sensitive Selection (OGRES) algorithm, for sonar system performance predictions was reported for a realistic synthetic square annular ridge environment [E. R. Rike and D. R. DelBalzo, Proc. 6th ECUA (2002)] and a complex real-world environment in the Sea of Japan [E. R. Rike and D. R. DelBalzo, Proc. IEEE Oceans (2002)]. The sparse OGRES grids significantly outperformed uniform grids (matched for calculation time) in accuracy. Both cases were based on ground-truth uniform grids high in spatial density, but moderate-to-low in azimuthal density. The present study is a very sparse gridding of a realistic synthetic spiral ridge based on a ground-truth uniform grid high in both spatial and azimuthal density. This analysis extends the previous work by exploring the relationship between the spatial curvature of bathymetric features and optimal azimuthal grid density. The trade-offs between calculation time and accuracy of sonar system predictions for both very sparse and uniform grids are discussed. [Work supported by ONR/NRL under PE 62435N.]

11:30

3aUW13. Acoustic modal analysis in an anisotropic internal wave field. Roger Oba and Steven Finette (Acoust. Div. Naval Res. Lab, Washington, DC 20375)

Simulations show that an anisotropic, internal wave field in shallow water can cause significant horizontal redirection of acoustic propagation, requiring true three-dimensional sound modeling. Calculations are presented for a sound speed profile that is perturbed by both a nonlinear internal wave packet, and a linear background internal wave field that is homogeneous and isotropic. A 3D parabolic equation code is used to propagate a 400 Hz signal through the spatially and temporally varying sound speed distribution. A single mode starter field at the source is propagated to a receiver location, and there the field is decomposed into modes to determine the degree of modal coupling. If the compressional wave number vector is perpendicular to that of the nonlinear internal wave packet wave number vector, the modeled acoustic field is nearly adiabatic and can be approximated using horizontal mode decomposition. In this case the pressure wave shows horizontal refraction, and ducting. For propagation paths not perpendicular to the internal wave number vector, the acoustic field can b enonadiabatic, and mode coupling can be significant. Modal analysis shows similar refraction and coupling results for lower frequencies. [Work supported by ONR.]

11:45

3aUW14. Simulations of time spreading in shallow water propagation. Eric I. Thorsos, W. T. Elam, Dajun Tang, Frank S. Henyey, Kevin L. Williams, and Stephen A. Reynolds (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Pulse propagation in a shallow water wave guide leads to time spreading due to multipath effects. Results of PE simulations will be described for pulse propagation in shallow water with a rough sea surface and a flat sandy sea floor. The simulations illustrate that such time spreading may be significantly less at longer ranges than for the flat surface case. Pressure fields are simulated in two space dimensions and have been obtained using a wide-angle PE code developed by Rosenberg [A. D. Rosenberg, J. Acoust. Soc. Am. **105**, 144–153 (1999)]. The effect of rough surface scattering is to cause acoustic energy initially propagating at relatively high angles but still below the critical angle at the sea floor to be eventually shifted to grazing angles above the critical angle. This energy is then lost into the bottom, effectively stripping higher propagating modes. The surviving energy at longer ranges is concentrated in the lowest modes and shows little effect of time spreading. Thus, the effect of rough surface scattering is found to produce a simpler temporal field structure than if the surface were treated as flat. [Work supported by ONR.]

WEDNESDAY MORNING, 4 DECEMBER 2002 MERIDA ROOM (HYATT), 9:00 TO 11:00 A.M.

Meeting of Accredited Standards Committee (ASC) S2 Mechanical Vibration and Shock

to be held jointly with the

ANSI-Accredited U.S. Technical Advisory Group (TAG) Meetings for:
ISO/TC 108 Mechanical Vibration and Shock
ISO/TC 108/SC 1 Balancing, including balancing machines
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

R. J. Peppin, Chair S2
5012 Macon Road, Rockville, Maryland 20852

D. J. Evans, Vice Chair S2 and Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108 Mechanical Vibration and Shock and Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 3 Use and calibration of vibration and shock measuring devices
National Institute of Standards and Technology (NIST), 100 Bureau Drive, Stop 8221, Gaithersburg, Maryland 20899-8221

R. Eshleman, Acting Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 1 Balancing, including balancing machines
Vibration Institute, 6262 Kingery Highway, Ste. 212, Willowbrook, Illinois 60514

A. F. Kilcullen, Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 2 Measurement and evaluation of mechanical vibration and shock as applied to machines, vehicles and structures
5012 Woods Road, Hedgesville, West Virginia 25427

R. Eshleman, Vice Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 2 and Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 5 Condition monitoring and diagnostics of machines
Vibration Institute, 6262 Kingery Highway, Ste. 212, Willowbrook, Illinois 60514

G. Booth, Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 6 Vibration and shock generating systems
44 Bristol Street, Branford, Connecticut 06405-4842

Accredited Standards Committee S2 on Mechanical Vibration and Shock. Working group chairs will report on the status of various shock and vibration standards currently under development. Consideration will be given to new standards that might be needed over the next few years. There will be a report on the interface of S2 activities with those of ISO/TC 108 and its subcom-

mittees including plans for future meetings of ISO/TC 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 fields of mechanical vibration and shock, and condition monitoring and diagnostics of machines, but excluding those aspects which pertain to biological safety, tolerance and comfort.

WEDNESDAY AFTERNOON, 4 DECEMBER 2002 GRAND CORAL 2, 1:15 TO 2:30 P.M.

Session 3pAA

Architectural Acoustics: Halls, Theaters and Cathedrals

David E. Marsh, Chair
Pelton Marsh Kinsella, 1420 West Mockingbird Lane, Suite 400, Dallas, Texas 75247-4932

Contributed Papers

1:15

3pAA1. Virtual auditorium concepts for exhibition halls. Jack Evans, Chad Himmel, and Sarah Knight (JEAcoustics/Eng. Vib. Acoustics & Noise Solutions, 5806 Mesa Dr., Ste. 380, Austin, TX 78731-3742)

Many communities lack good performance facilities for symphonic music, opera, dramatic and musical arts, but have basic convention, exhibition or assembly spaces. It should be possible to develop performance space environments within large multipurpose facilities that will accommodate production and presentation of dramatic arts. Concepts for moderate-cost, temporary enhancements that transform boxy spaces into more intimate, acoustically articulated venues will be presented. Acoustical criteria and design parameters will be discussed in the context of creating a virtual auditorium within the building envelope. Physical, economic, and logistical limitations affect implementation. Sound reinforcement system augmentation can supplement the room conversion. Acceptable control of reflection patterns, reverberation, and to some extent, ambient noise, may be achieved with an array of nonpermanent reflector and absorber elements. These elements can sculpture an enclosure to approach the shape and acoustic characteristics of an auditorium. Plan and section illustrations will be included.

1:30

3pAA2. Municipal Theater of Lima: Reconstruction after fire. Jorge Moy, Mario Segami, Jose Ferrari (Universidad Peruana de Ciencias Aplicadas., Dep. de Arquitectura., Av Prolongacion Primavera 2390, Surco, Lima, Peru), and Richard A. Moscoso (Pontificia Universidad Catolica del Peru, Lima, Peru)

The Municipal Theater of Lima was inaugurated on 28 July 1920 as the Forero Theater and was acquired by the Metropolitan Municipality of Lima in 1929. On 2 August 1998 it was partially destroyed by fire. The damage to the audience area was minimal, allowing complete restoration. However, the stage was destroyed, which required a complete reconstruction. The Municipality of Lima launched a project competition for its reconstruction and restoration the results of which were known in 2001. One of the objectives was the correction of well-known acoustical defects, which had been never solved. The results of computer simulations of the original theater with a commercially available ray-tracing-type computer program will be presented, showing its acoustical evaluation and available options to correct the acoustical defects. The acquisition of five lots around the theater site will allow increases in the fly tower, the dressing rooms, the pit, and the shops. There are plans for the modernization of the technical and mechanical services, the accesses, the air-conditioning system, and security as well. These modifications will allow all types of presentations of contemporary arts. Retaking its position as first in the national scene, the Municipal Theater will be able to house sophisticated national and international performances.

1:45

3pAA3. Reuse of a cylinder shape lecture hall as a recital hall. Weihwa Chiang, Weiping Wu, Wenling Chih, and Choye Lee (Natl. Taiwan Univ. of Sci. and Technol., 43, Sec. 4, Keelung Rd., Taipei, Taiwan)

A cylinder shape lecture hall in Danchiang University, Taiwan was reused as a 300-seat recital hall with recording facilities. Acoustical design was analyzed with computer models and a scale model. Besides the acoustical defects due to the cylinder plan, the low ceiling height of 3.3 m caused a low reverberation with ceiling preserved but poor sound propagation with the ceiling removed because of the 1.2 deep girders. Three major convex-curved walls were used, two on both sides of the stage and the other as the partition between the hall and the control room. Two side boxes on each side of the hall were used with the box facing slightly tilted down to provide lateral reflection. Installing oval shaped supply ducts diffused the remaining concave wall surfaces. Triangular overhead panels in various sizes were used to diffuse first order ceiling reflection. Scale modeling demonstrated that most of the design strategies were not only effective in enhancing uniform energy distribution spatially but also randomized the reflection pattern for individual receivers. A low strength factor of 11 dB was beneficial for large chamber groups. On going study is conducted regarding the field tuning and measurements of the hall.

2:00

3pAA4. Acoustic conditioning of the metropolitan cathedral of Porto Alegre, RS, Brazil. Flavio M. Simoes, Luciano B. Nabinger, and Aline I. Ramalho (Curso Arquiteura Urbanismo-Ritter dos Reis, FAIR, R. Orfanotrofio, 555 Porto Alegre, Brazil)

In the acoustic study of the Metropolitan Cathedral of Porto Alegre, RS, Brazil, initially background noise and reverberation time were measured. A digital model was built using acoustic simulation software AcustaCadd, applying the values of the measured reverberation time. Then reverberation time, speech intelligibility, and geometric acoustics were analyzed. As a result the Project of Acoustic Conditioning was developed to correct the high reverberation time, by increasing absorption with the installation of 65 000 m of panels of glass wool (100 mm, 60 kg/m). Advantage was taken of existing details in the plaster to embed the panels in the walls. Also the volume of the choir and of the lateral balcony to the

altar was reduced and the interior of this was covered with the same glass wool. Special care was taken to minimize alterations to the architectural characteristics of the place, because it is a construction of historical importance. The measured values of background noise were also analyzed and appropriate acoustic isolation considered. The final measure of the reverberation time showed an average reduction of 5 seconds and better speech intelligibility, long demanded by the users. [Work supported by FAIR/FUNDATEC, BR; IUCC-US, SP.]

2:15

3pAA5. Subjective evaluation of a concert hall's acoustics using a free-format-type questionnaire and comparison with objective measurements. Toshiyuki Okano (Takenaka Res. and Development Inst., 1-5-1, Ohtsuka, Inzai, Chiba 270-1395, Japan, okano.toshiyuki@takenaka.co.jp) and Leo L. Beranek (Cambridge, MA 02138)

A free-format type of audiences' judgment of the acoustical properties of a hall and music critics' writings were used as the basis for this study. These subjective responses are related to the Dai-Ichi Seimei Hall in Tokyo. This hall is an oval-shaped, one-balcony space, seating 767 persons. Its primary use is for various types of chamber music and solo-instrument performances. Eight acoustical attributes were investigated, ''reverberation,'' ''clarity,'' ''loudness,'' ''intimacy,'' ''spaciousness,'' ''balance,'' ''localization,'' and ''timbre,'' plus ''general impression.'' Subjective comments about these attributes were obtained. Objective measurements were made in the hall and are compared with those made in several similar-sized halls of two shapes. In the rear seats of two oval-shaped halls the strength factor GE (determined in the first 80 ms of the impulse response) was greater than the GE found in the rear seats of similar-sized rectangular halls. The subjective results and the objective measurements were closely correlated, especially for reverberation, clarity, and warmth (a subcomponent of timbre). It was suggested that the greater strength GE in the rear seats made the hall seem smaller and thus more intimate. The subjective comments also confirmed the hall's wide applicability, indicating that the acoustical characteristics used for its design were well chosen.

WEDNESDAY AFTERNOON, 4 DECEMBER 2002 CORAL KINGDOM 2 AND 3, 1:15 TO 3:05 P.M.

Session 3pAB

Animal Bioacoustics and Psychological and Physiological Acoustics: Electrophysiological Investigations of Animals II

Eduardo Mercado III, Chair
Center for Molecular and Behavioral Neuroscience, Rutgers University, 197 University Avenue, Newark, New Jersey 07102

Invited Papers

1:15

3pAB1. Evoked-potential study of echolocation in whales and dolphins. Alexander Ya. Supin (Inst. of Ecology and Evolution, 33 Leninsky Prosp., 119071 Moscow, Russia), Paul E. Nachtigall, and Whitlow W. L. Au (Univ. of Hawaii, Kailua, HI)

During echolocation, whales and dolphins perceive faint echoes very shortly after loud ongoing pulses. How does the auditory system avoid the masking of the echo by the preceding pulse? To investigate mechanisms of release from masking, auditory brainstem evoked responses (ABR) were recorded in bottlenose dolphins and a false killer whale both during active echolocation and during paired external acoustic stimulation. When the animal ensonified a target, both the ongoing pulse and echo (2-ms delay) evoked ABRs of comparable amplitudes, in spite of echo intensity of 40 dB relative the ongoing pulse near the animal's head. When a phantom echo was used (emitted pulses were replayed through an electronic channel with the same delay of 2 ms), the ABR to the echo was comparable to that to the emitted pulse at the echo intensity of 20 dB, but the ABR decreased when the echo intensity diminished to 40 dB relative to the ongoing pulse. An explanation of higher efficiency of the real than the phantom echo may be a difference between spectra of the real echo and the outgoing pulse. Paired-pulse stimulation confirmed that a moderate spectral difference between two pulses resulted in release from masking at short interpulse intervals.

1:35

3pAB2. Frequency and directional selectivity in primary auditory afferents in fishes. Richard R. Fay (Parmly Hearing Inst., Loyola Univ. of Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, rfay@luc.edu)

Coding of frequency and particle motion directionality is described for two fishes; goldfish (a sound-pressure specialist), and oyster toadfish (a hearing generalist). The goldfish hears in a wide bandwidth (>2 kHz) and the toadfish in a narrower one (<400 Hz). Nevertheless, the mechanisms and degree of peripheral frequency analysis appear to be similar. Reverse correlation (revcor) studies of frequency filtering in the auditory nerve reveal two fundamental filter types in both species; one at low and one at higher frequencies. The details of the two filter functions are similar, and one can be transformed into the other by introducing a simple spectral tilt to one of the filters. This suggests that only one fundamental filter exists for each species, and that the second filter type may be derived from it through alterations in the mode of coupling between the hair cell cilia and restraint structures. In spite of this simple filtering, goldfish effectively analyze the frequency components in complex and simple sounds. In both species, primary afferents show directional preferences in 3-D space that correspond to the directionalities of hair cells. Frequency analysis and directional processing in fishes share many fundamental features with those of most other vertebrates.

1:55

3pAB3. Auditory brainstem response (ABR), a noninvasive electrophysiological method for the study of hearing ability of fishes. Hong Yan (School of Biological Sci., Univ. of Kentucky, Lexington, KY 40506-0225)

Over the past 100+ years several methods have been used to study hearing ability of fish. Behavioral (psychophysical) methods including food reward, barrier crossing, and electrical shock conditioning are used to train fish to respond to sound. Invasive-type of electrophysiological methods include single unit and microphonics recordings. In 1996, my laboratory developed auditory brainstem response (ABR) recording technique, an electrophysiological far-field recording method, to record acoustically evoked brainstem potentials from fish. This noninvasive method overcomes the major limitations of traditional behavioral (prolonged training period; erratic behavioral response) and electrophysiological (invasive surgery) methods. The ABR recording can be achieved easily with assemblage of conventional electrophysiological devices with the aid of computer programs. Since the development of this noninvasive electrophysiological recording method, it has been widely adopted by many laboratories around the world to study fish hearing. The role of gasbladder in hearing enhancement, the ontogenetic development of hearing ability of embryo and larval fish, the impacts of noise exposure on threshold shifts, the effect of aging on hearing all have been understood with the use of ABR method. This quick pace of achievement cannot be obtained without the use of this noninvasive recording technique. [Work supported by NIH, NOHR, IMLS.]

2:15

3pAB4. Sea turtle auditory behavior. Martin Lenhardt (Prog. in Biomed. Eng., Virginia Commonwealth Univ., Box 168 MCV, Richmond, VA 23298-0168)

Little is known of sea turtles' auditory behavior. Startles (neck contractions) were observed in 25 loggerheads (*Caretta caretta*) to tones from 35 to 1000 Hz, when the turtles were near the bottom of holding tanks at a depth of 1 meter. A composite audiogram revealed lowest thresholds in the 400–500 Hz range (106 dB SPL *re* 1 μm). Thresholds at 735 and 1000 Hz were 117 and 156 dB, respectively. Thresholds in the 100–200 Hz range were $\sim$124 dB, with lower frequencies being 10–12 dB higher. Tank diving behavior was elicited with 30 Hz at 164 dB. ABR thresholds to vibration clicks with peak energy at 500 Hz were 113 dB. Seismic air guns (Bolt 600) were employed in a large net enclosure. Turtles increased swimming speed for exposures in the 151–161 dB levels. Avoidance $\sim$175 dB was common in initial trials, before habituation. ABRs pre- and post-air gun exposures revealed TTS of more than 15 dB in one animal with recovery in 2 weeks. Air guns in depths $>$10 m may result in more energy in the low frequencies with unknown bioeffects. Turtle repelling with sound is possible and can be made practical (Lenhardt, US Patent No. 6388949).

Contributed Papers

2:35

3pAB5. Some constraints on the use of ABR in hearing studies: Investigation in dolphins. Vladimir Popov and Alexander Supin (Inst. of Ecology and Evolution, Russian Acad. of Sci., 33 Leninsky Prosp., 119071 Moscow, Russia)

The main difficulty in using ABR in frequency-tuning measurements arises from the fact that these responses are evoked only by transient acoustic stimuli. This limits their use as probes for measuring frequency selectivity. Even with a slowly rising and falling stimulus, only a shorter part of the stimulus may be effective to evoke the response, and the spectrum of this short part is broader than the spectrum of the whole burst. It was shown in tone–tone masking study in dolphins, that tuning curves at lower frequencies reflect not the real frequency tuning but broad spectral bandwidth of the probe's effective part. In the bottlenose dolphin, this effective part is not longer than 0.5 ms, irrespective of the actual probe duration. Another problem in the use of the ABR technique is that it is not clear which spectral part of the wideband stimulus participates in ABR generation, and what is the contribution of various parts of the cochlear partition to ABR. ABR to clicks and noise bursts of various frequency bands were recorded in dolphins in order to find these contributions. ABR amplitude increased with increasing the stimulus frequency, thus indicating a higher contribution of basal cochlear parts.

2:50

3pAB6. Experience-dependent cortical processing of complex sounds by rats. Itzel Orduna, Mark A. Gluck (CMBN, Rutgers Univ., 197 University Ave., Newark, NJ 07102, orduna@axon.rutgers.edu), Eduardo MercadoIII (SUNY at Buffalo, Buffalo, NY 14260), Jennifer F. Linden, Shaowen Bao, Michael M. Merzenich (UCSF, San Francisco, CA 94143), and Maneesh Sahani (UCL, London, UK)

Auditory cortex is thought to play a critical role in the processing of species-specific vocalizations and other acoustically complex sounds. Although evolutionary processes strongly constrain cortical sensitivities to sound, cortical processing is not fixed by biology, but rather is shaped by the auditory experiences of each individual. Auditory cortical neurons in adult rats respond selectively to spectrotemporal features of complex sounds. These selective responses are predictive of rats' behaviorally measured perceptual sensitivities. With extensive training, the abilities of rats to discriminate frequency-modulated sounds improve. Recordings from cortical neurons in trained rats show increased sensitivities to features of the sounds used in training. These results demonstrate that discrimination training with biologically irrelevant complex sounds can change how cortical neurons process those sounds. Changes in cortical processing of complex sounds can also be induced by controlling activity in neuromodulatory neurons while a rat experiences the sounds. Pairing the presentation of complex sounds with electrical stimulation of basal forebrain neurons radically changes cortical sensitivities to spectrotemporal features of those sounds. These data suggest that auditory cortical representations of acoustic events can be flexibly adjusted throughout the lifespan of an individual rate. Presumably, similar or greater flexibility exists in animals with more sophisticated auditory systems.

Session 3pBB

Biomedical Ultrasound/Bioresponse to Vibration: Lithotripsy II

Pei Zhong, Chair

Department of Mechanical Engineering and Materials Science, Duke University, 1 Science Drive, Box 90300, Durham, North Carolina 27708

Contributed Papers

1:00

3pBB1. A gypsum-based artificial stone for shock wave lithotripsy research. James A. McAteer, James C. Williams, Jr., Andrew P. Evan (Dept. of Anatomy & Cell Biol., Indiana Univ. School of Medicine, 635 Barnhill Dr., Indianapolis, IN 46202-5120, mcateer@anatomy.iupui.edu), Robin O. Cleveland (Boston Univ., Boston, MA 02115), Michael R. Bailey, and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

Natural kidney stones are heterogeneous in structure, composition, material properties and fragility, and as such are problematic for use in determining the mechanisms of SW-action in SWL. A variety of model stones have been developed. We have adopted Ultracal-30 gypsum [Dahake and Gracewski, J. Acoust. Soc. Am. **102**, 2138 (1997)] for *in-vitro* and *in-vivo* studies. U-30 stones (7.5×6.5 mm) cast in polystyrene molds were liberated with chloroform and stored in water. Drop-impact testing of dry stones showed a linear relationship between increase in surface area of fragments and energy applied. Breakage of hydrated stones in a research-electrohydraulic lithotripter, likewise showed a linear increase in fragment area with increased SW-number and SW-voltage. The density (1800 kg/m^3) and transverse (1520 m/s) and longitudinal (3100 m/s) wave speeds of U-30 stones place them in the range determined for natural stones. U-30 stones implanted in pig kidneys exhibited cavitation erosion and spall fracture similar to stones *in-vitro*, and U-30 stones *in-vitro* and *in-vivo* showed equivalent response to SW-rate (200% higher fragmentation at 0.5 Hz compared to 2 Hz). U-30 stones softened with prolonged exposure to water and degraded during long-term implantation *in-vivo*. With these cavaets U-30 stones provide a useful model for SWL research. [Work supported by NIH P01-DK43881.]

1:15

3pBB2. Effect of the diameter and the sound speed of a kidney stone on the acoustic field induced by shock waves. Robin Cleveland and Juan Tello (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215)

A time domain finite difference solution to the acoustic wave equation was used to model the propagation of lithotripsy shock waves in kidney stones. The stones were modeled as cylindrical objects and the impact of stone sound speed, stone diameter, and the internal structure on the pressure field inside the stones was calculated. The sound speed was varied from 2500 to 3500 m/s and had a minimal effect on the peak pressures within the stone. However, reducing the stone diameter from 15 mm to 3 mm reduced the peak negative pressure by 67%. For cases where the sound speed of the stone was inhomogeneous (inner core different from outer ring), the amplitudes of the peak pressures varied slightly, however the spatial distribution of the peak negative pressure varied significantly. In particular configurations with a lower outer sound speed lead to the peak negative pressure occurring at the boundary between the layers. The results indicate that the pressure field in a kidney stone is very sensitive to the size and internal composition of a kidney stone. This effect could be partly responsible for the large variance in fragility observed in human stones. [Work supported by the Whitaker Foundation and NIH-DDK.]

1:30

3pBB3. Progression of crack formation in artificial kidney stones subject to shock waves. Javier Van Cauwelaert and Robin Cleveland (Aerosp. and Mech. Eng. Dept., Boston Univ., 110 Cummington St., Boston, MA 02215)

We used micro computed tomography (CT) imaging to follow the progressive development of cracks in artificial kidney stones. The artificial stones were made from U30 cement with a cylindrical shape (6.5 mm diameter and 7.5 mm long). The stones were held within a polypropylene vial in one of three different orientations: vertical, horizontal, and angled at 45 deg. The stones were treated with an electromagnetic lithotripter using between 50 and 150 shock waves. The initiation and growth of cracks was observed using microCT. We found that crack formation in the U30 stones was influenced by the orientation of the stone with respect to the shock wave (SW) propagation direction. Vertical stones developed a spall-like crack near the distal surface; horizontal stones had little internal damage for the number of shock waves applied; and angled stones were damaged primarily in the vicinity of the leading corner. The position of the cracks were in qualitative agreement with the numerical solutions of the pressure field inside the stones. The elastic properties of the U30 stones are being measured which will alow a quantitative analysis of crack growth to be performed and compared to the experimental data. [Work supported by the Whitaker Foundation.]

1:45

3pBB4. Role of cavitation in stone fragmentation by shock wave lithotripsy. James A. McAteer, James C. Williams, Jr., Andrew P. Evan (Dept. of Anatomy & Cell Biol., Indiana Univ. School of Medicine, 635 Barnhill Dr., Indianapolis, IN 46202-5120, mcateer@anatomy.iupui.edu), Robin O. Cleveland (Boston Univ., Boston, MA 02115), Michael R. Bailey, and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

Several strategies were used to assess the importance of cavitation in the breakage of stones by an electrohydraulic lithotripter *in vitro*. (1) Stones exposed to SWs at atmospheric pressure broke readily. However, stones treated at high overpressure (OP$\sim$125 atm) sufficient to eliminate cavitation did not break into fragments even with twice the number of SWs. Stones at OP did, however, develop transverse fractures typical of spall. This suggests that cavitation contributes to stone fragmentation, but is clearly not the only mechanism involved in stone breakage. (2) Cylindrical model stones positioned vertically in the acoustic field of a research-EHL showed proximal erosion and spall. However, placement of a mylar disk against the flat leading face of the stone eliminated cavitation-erosion, and spall did not occur. This suggests that cavitation may contribute to stone fracture by spall. (3) Time reversal of the lithotripter wave form using a pressure release reflector (Prel) also prevented stone fragmentation. With the Prel insert the tensile phase of the SW preceding the compressive wave bubble growth is interrupted by P+ and, thus, cavitation is suppressed. Together, these results suggest that cavitation plays an important role in the breakage of stones by lithotripter shock waves. [Work supported by NIH P01-DK43881.]

2:00

3pBB5. Evidence of oxidative stress in both kidneys after shock-wave lithotripsy to one renal pole. Lynn R. Willis, James E. Klaunig, Yong Xu (Dept. of Pharmacology and Toxicology, Indiana Univ. School of Medicine, 635 Barnhill Dr., Indianapolis, IN 46202), Bret A. Connors, Andrew P. Evan (Indiana Univ. School of Medicine, Indianapolis, IN 46202), and James E. Lingeman (Methodist-Clarian Hospital, Indianapolis, IN 46202)

Shock-wave lithotripsy (SWL) damages renal tissue and reduces renal perfusion. These studies asked whether SW-induced injury or reduced perfusion might cause oxidative stress. Two groups of 3 anesthetized pigs received 2000 shock waves (SWs) at 24 kV (Dornier HM3, 2 Hz) to one renal pole. Renal hemodynamics were measured 1 h before and 1 and 4 h after SWL. Samples of renal tissue from upper and lower poles of each kidney were analyzed for malonaldehyde (lipid peroxidation) and 8-OH-deoxyguanosine (OH8dG, DNA damage). SWL reduced perfusion in both kidneys, but injured only the shocked kidneys. Malonaldehyde levels were elevated in both poles of the shocked kidneys, but were not elevated in the contralateral unshocked kidneys. In contrast, OH8dG levels were elevated in the treated pole of the shocked kidneys, and in both poles of the contralateral unshocked kidneys. In conclusion, (1) The elevated levels of malonaldehyde seen in the shocked kidneys but not in the unshocked kidneys are consistent with lipid peroxidation occurring after SWL-induced tissue injury; (2) elevated OH8dG levels in contralateral unshocked kidneys denotes oxidative stress and may reflect reperfusion injury associated with SWL-induced renal vasoconstriction. [Work supported by a grant from NIH P01 DK43881.]

2:15

3pBB6. The reduced renal blood flow observed after exposure to shock wave lithotripsy involves intact renal nerves. Bret A. Connors, Andrew P. Evan, Lynn R. Willis, Jay R. Simon (Indiana Univ. School of Medicine, Indianapolis, IN 46202, connors@anatomy.iupui.edu), Arieh L. Shalhav, Ryan F. Paterson, Ramsay L. Kuo, and James E. Lingeman (Methodist Hospital Inst. for Kidney Stone Disease, Indianapolis, IN 46202)

Treating one kidney with high-energy acoustic shock waves reduces blood flow 65% in that kidney and 33% in the contralateral unshocked kidney. We tested the hypothesis that renal nerves were involved in this response. Six-week old pigs underwent unilateral renal denervation. Nerves along the renal artery of one kidney were cut and the artery was painted with 10% phenol. After 2 weeks, the pigs were anesthetized and bilateral renal function was determined using inulin and PAH (para-aminohippurate) clearance. Glomerular filtration rate (GFR, inulin clearance) and renal blood flow (PAH clearance) were measured both before and after lithotripsy to the lower pole of the innervated kidney (2000 shocks, 24 kV, unmodified Dornier HM-3) or sham lithotripsy. Both kidneys were then removed to measure norepinephrine content in the tissue. Norepinephrine levels were significantly reduced (more than 90%) in denervated kidneys indicating complete denervation. As expected, blood flow and GFR fell (approximately 50%) after lithotripsy in the innervated shocked kidneys, but did not fall significantly in denervated contralateral kidneys or in sham lithotripsy animals. Results suggest that renal nerves mediate the fall in contralateral renal function after lithotripsy. [Work supported by NIH P01 DK43881.]

2:30

3pBB7. Effects of the viscosity of the liquid in the microcapsule on deformation process of a bubble for developing DDS using shock waves. Masaaki Tamagawa, Ichiroh Yamanoi, and Toshiaki Matsuo (Grad. School of Life and Systems Eng., Kyushu Inst. of Technol., Wakamatsu-ku, Fukuoka 808-0196, Japan, tama@life.kyutech.ac.jp)

This paper describes the fundamental investigations for developing new DDS using shock waves and bubbles, especially effects of the viscosity of the liquid in the capsules on the deformation process of the bubbles. The viscosity of the liquid in the microcapsules requires controlling the bubble position and motion, deformation process to optimize the disintegration of the capsules. Using plane shock wave in the water by shock tube apparatus, a bubble near the curved gelatin wall was observed by high-speed framing camera. The deformation process for a bubble was analyzed by image processing. From these experiments, it is found that the viscosity of the liquid around the bubble suppress the amplitude of the bubble oscillations and collapsing the bubble. This means that the viscosity of the liquid in the capsule should be as small as possible for developing DDS using shock waves. In spite of this result, the effects of curvature and elasticity of the wall on the deformation process is almost the same as low viscosity. [Work supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan.]

Session 3pMU

Musical Acoustics: Hands-On Explorations in Acoustics

James P. Cottingham, Cochair
Physics Department, Coe College, Cedar Rapids, Iowa 52402

Leonardo Fuks, Cochair
Escola de Musica, Universidade do Brazil UFRJ, Rue do Passero 98, Rio de Janeiro 20021-290, Brazil

Hands-on activities in acoustics, open to all meeting attenders, but designated primarily for accompanying persons, both adults and children

Session 3pNS

Noise: Hearing Protection II

Daniel P. Salomon, Cochair
Comaudi, Patriotismo 706, Colonia Mixcoac 03730, D.F. Mexico

Elliott H. Berger, Cochair
EAR/Aearo Company, 7911 Zionsville Road, Indianapolis, Indiana 46268-1657

Chair's Introduction—1:00

Contributed Papers

1:05

3pNS1. Current status of standards for testing electroacoustic hearing protectors. William J. Murphy and John R. Franks (NIOSH Hearing Loss Prevention Section, 4676 Columbia Pkwy., M.S. C-27, Cincinnati, OH 45226-1998, wjm4@cdc.gov)

Electroacoustic hearing protectors encompass several classes of electronics: level limiting, amplitude compression and active noise control (ANC) devices. Each class of protector seeks to enhance performance by overcoming insertion loss and/or improving low-frequency attenuation. Level-limiting devices turn off the amplification when the external sound level exceeds a preset threshold. Amplitude compression devices apply variable gain and limit amplification of signals above the threshold. ANC devices create an out-of-phase signal under the protector to improve the low-frequency attenuation of earmuffs. ANC devices typically perform best in continuous noise that has strong harmonic components that can be predicted and cancelled. The different methods of enhancing protector performance require separate standards to understand both the static and dynamic performance. Testing of ANC devices has further requirements which require knowledge of both passive and active electronic conditions and in noise with different crest factors. This talk will review a variety of tests of the different protector classes.

1:20

3pNS2. Objective measurements for the assessment of hearing protectors attenuation at high level impulsive noise. Felipe Vergara, Samir N. Y. Gerges, Washington J. N. de Lima (Federal Univ. of Santa Catarina, Mech. Eng., Cx.P. 476, Florianopolis, SC, Brazil, CEP:88040-900), and Robert Birch (Univ. of Liverpool, Liverpool L69 3H, UK)

The evaluation of hearing protectors for the attenuation of high amplitude impulsive noise cannot be carried out using the conventional subjective Real Ear Attenuation at Threshold (REAT) technique. In the case for high level impulsive noise it is not possible to conduct subjective type tests with volunteer listeners, therefore, other methods based on using

artificial human head forms need to be considered. In this paper we present an objective technique that uses an artificial head, ear-canal simulator and shock tube to assess earplug and earmuff protectors. The shock tube is used as a means of producing controlled repeatable high amplitude pressure pulses (>140 dB) of varying rise time, amplitude, and duration. The artificial head and ear simulator with a protector are mounted inside the shock tube and subjected to a high level sound pulse. Pressure levels are measured simultaneously outside of the protector (incident pulse) and inside at the ear simulator at the eardrum position. The attenuation of peak pressure levels in the time domain and the corresponding spectra of the pulses are determined. The technique offers pointers towards future standards of performance hearing protectors for applications with high level impulsive noise.

1:35

3pNS3. Attenuation of high-level acoustic impulses by hearing protectors. Jan Zera (Central Inst. for Labour Protection, Czerniakowska 16, 00-701 Warsaw, Poland)

Attenuation of acoustic impulses by hearing protectors was measured in peak level ranges of 115–135 dB and 145–170 dB. The impulses were generated by a loudspeaker system in the low range of levels and by a blast of air expanding from a cylinder in the high level range. Transmission loss method was used to determine the difference between the peak level under the earmuff and outside the earmuff for over 30 different types of earmuffs. The measurements confirm that attenuation of high-level impulses depends on their peak level. In the 115–135 dB range the difference between the peak level outside and under the earmuff is approximately constant. In the 145–170 dB range, as the impulse level is increased, the level under the earmuff increases to a lesser extent. The peak-to-peak level characteristics determined for the 115 to 170 dB level range substantially differ for various types of protectors and may be a useful indicator of the effectiveness of protectors for impulse noise. [Work supported by the State Committee for Scientific Research Grants Nos. III-7.03 and III-6.07.]

1:50

3pNS4. Communications earplug performance in tracked-vehicle crews. William A. Ahroon, Dale A. Ostler (U.S. Army Aeromedical Res. Lab., P.O. Box 620577, Fort Rucker, AL 36362-0577), Ben T. Mozo (Commun. & Ear Protection, Inc., Enterprise, AL 36331-1174), and Cynthia M. Crossley (Martin Army Community Hospital, Fort Benning, GA 31905)

Tracked-vehicle crewmembers operate in an acoustical environment that is one of the most inhospitable encountered in the U.S. Army. Overall noise levels in the M1A2 Abrams tank and M2A2 Bradley Fighting Vehicle can be well over 100 dBA and can exceed 130 dB in some octave bands. Due to the nature of the crewmembers' duties, speech communication is essential for effective operations and this capability is seriously degraded because of the nature of the noise environment. The Communications Earplug (CEP), consisting of miniature receivers attached to replaceable foam earplugs, is used in Army aviation to provide superior noise attenuation with improved speech communication. The noise attenuation of the DH-132A Combat Vehicle Crewmans helmet (CVC) when used in conjunction with CEP (measured in accordance with ANSI S12.6-1997) is reported with an NRR of 27 dB vs 15 dB for the CVC alone. CEPs were installed in the helmets of 77 tracked-vehicle crewmembers. Following a 6-month trial period, crewmembers ($n=36$) reported increased noise attenuation, improved speech communication, and reduced stress when using the CVC-CEP integration. The CEP is a cost-effective alternative to expensive active noise reduction helmets currently being fielded for use by tracked-vehicle crews. [Work supported in part by USA-CHPPM.]

2:05

3pNS5. System for measurement of headband force in hearing protection devices and audiometric equipment. Osvaldo Llamas-Llamas and Jose-Noe Razo-Razo (Centro Nacional de Metrologia, Div. de Vibraciones y Acustica, Carr. a Los Cues km 4,5 76241 El Marques, Qro., Mexico, ollamas@cenam.mx)

Application force influences audiometric results and hearing protection devices (HPD) attenuation performance. In HPD attenuation testing it is neccessary to know the application force. Audiometric studies are being conducted with standardized values of the application force. Design and results of a measurement device for bands and headbands application force are presented. Error analysis in the range from 1N to 30N is carried out and the associated measurement uncertainty is estimated. The device provides settings for test distances from 70 mm to 140 mm (vertical axis), and 115 mm to 195 mm (horizontal axis), as required in most of the available standards related with HPD and audiometric equipment.

2:20

3pNS6. Overpressure and noise due to multiple airbag systems in a passenger car. Robert Hickling (Sonometrics, Inc., 8306 Huntington Rd., Huntington Woods, MI 48070, sonomet@aol.com), Peter J. Henning, and Gary Newton, Jr. (Bruel & Kjaer Sound and Vib. Measurement A/S, Livonia, MI 48154)

Multiple airbag systems in passenger cars can generate overpressure and noise that may be hazardous to human hearing. Overpressure is compression of the air inside a closed compartment caused by deployment of the bags. Noise results from the action of the gas inflating the bags. SAE J247 provides a standard for measuring the combination of overpressure and noise in a passenger compartment. A special microphone has recently been developed that meets this standard, which operates down to a fraction of a hertz. Details of the microphone are given. Little appears to have been published on the overpressure and noise of modern multiple airbag systems, but early results [R. Hickling, ''The noise of the automotive safety air cushion,'' Noise Control Eng., May–June, 110–121 (1976)] provide a basic understanding of the phenomenon. Spectral data shows that peak overpressure occurs at about 2 to 3 Hz. A significant reduction in overpressure and noise can be achieved with an aspirating airbag, originally developed at General Motors, whose outer structure is inflated with gas from the inflator, and whose inner structure draws in air from the passenger compartment through one-way cloth valves. Tests have shown that such bags function well when impacted.

Session 3pPA

Physical Acoustics and Signal Processing in Acoustics: The Coda and Other Stochastic Seismic Signals II

Joseph A. Turner, Chair
Department of Engineering Mechanics, University of Nebraska-Lincoln, Lincoln, Nebraska 68588

Invited Papers

1:00

3pPA1. Diffusing acoustic wave spectroscopy. John H. Page, Michael L. Cowan (Dept. of Phys. and Astron., Univ. of Manitoba, Winnipeg, MB R3T 2N2, Canada), and David A. Weitz (Dept. of Phys. and DEAS, Harvard Univ., Cambridge, MA 02138)

Diffusing Acoustic Wave Spectroscopy (DAWS) is an ultrasonic technique that has been developed to measure the dynamics of heterogeneous media from the temporal fluctuations of multiply scattered waves. This technique is similar to more recent developments in field fluctuation spectroscopy, called coda wave interferometry, that use variations in the seismic coda to infer changes in the medium with time. After reviewing the basic principles on which Diffusing Acoustic Wave Spectroscopy is based [M. L. Cowan, J. H. Page, and D. A. Weitz, Phys. Rev. Lett. **85**, 453 (2000); Phys. Rev. E **65**, 066605 (2002)], its potential as a sensitive method for probing the dynamics of strongly scattering materials will be illustrated with recent experiments on fluidized suspensions of particles. In this type of system, DAWS measures the local relative motion of the scatterers (or strain rate) on a length scale determined by the transport mean free path of the multiply scattered waves. When combined with the complementary technique of Dynamic Sound Scattering using singly scattered waves, DAWS can also determine the instantaneous velocity correlation length of the moving scatterers, thus giving a quite complete picture of the system dynamics over a wide range of length and time scales.

1:30

3pPA2. Correlations in the seismic coda and elastic Green's functions. Michel Campillo and Anne Paul (LGIT UJF-CNRS BP53 38041, Grenoble, France)

The conjecture that the correlation of coda records at two stations may provide the Green's function between these two stations, or an approximation of it, is discussed. Under the assumption of equipartition and/or considering that we use a set of sources that is evenly diffracted in the whole space, it can be shown that the average cross correlation between the records of every earthquake at the two stations is an approximation of the Green's function between the two stations. Records from stations of the Mexican network are used to test this idea. The stack of the cross correlations indicates that a low-frequency coherent signal is present with a signal-to-noise ratio that was ~ 0.2 for a single correlation. The Green's function between two points at the surface is expected to be widely dominated by the Rayleigh wave. The signal that we extracted presents the characteristics expected for a Rayleigh wave: elliptical polarization in the radial–vertical plane and adequate group velocity. The coherent signal extracted from coda correlations has the symmetry properties of the Green tensor. In spite of the *a priori* limitations of this approach in seismology, the preliminary results presented here demonstrate its potential.

2:00

3pPA3. Coda wave interferometry, a new method for monitoring change. Roel Snieder, Alexandre Gret, Huub Douma, and John Scales (Ctr. for Wave Phenomena and Dept. of Geophys., Colorado School of Mines, Golden, CO 80401-1887)

Multiply scattered waves are extremely sensitive to small changes of the medium through which these waves have propagated. Coda Wave Interferometry [Science **295**, 2253–2255 (2002)] is a new technique that utilizes multiply scattered waves in the time domain to monitor small changes in media. This is applied to ultrasonic waves that were recorded in a granite sample that was subjected to a change in temperature. Velocity perturbations of about 0.1% can be detected with this technique with an accuracy of about 0.02%. A multiply scattered wave in an elastic medium has traveled part of its trajectory as a *P*-wave and part as an *S*-wave. A model for the equilibration of *P*- and *S*-waves is presented. This model is used to extend the theory of coda wave interferometry to include elastic wave propagation. [Work was partially supported by the NSF (EAR-0106668 and EAR-0111804), by the U.S. Army Research Office (DAAG55-98-1-0070), and by the sponsors of the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena.]

Contributed Paper

2:30

3pPA4. Coda wave interferometry in finite solids, recovery of the *P* to *S* conversion rate. Richard L. Weaver and Oleg I. Lobkis (Dept. of Theoret. & Appl. Mech., 104 S. Wright St., Univ. of Illinois, Urbana, IL 61801, r-weaver@uiuc.edu)

Diffuse fields, which appear incoherent, nevertheless retain their deterministic character. Even wave fields that have scattered thousands of times are highly repeatable if temperature is held fixed. The temperature dependence of elastic wave speeds, and the great age of these signals can, however, lead to significant fluctuations. Our 1 MHz transient point source in a 10 cm aluminum block gives rise to a complex waveform that was observed to undergo an almost pure dilation of 0.0262% per degree C over the range from room temperature to 40°. The dilation rate was determined from the cross-correlation function between time-windowed responses at different temperatures. Theoretical calculations of the shift of the correlation function were constructed in terms of a weighted average of the temperature dependencies of the individual *P* and *S* waves that compose

the full field. That the *P* and *S* wave speeds have different dependence on temperature leads to a degradation of the strength of the correlation function maximum. The rate of degradation should be a measure of the mean lifetime of *P* and *S* rays against mode conversion; comparisons with theory corroborate that suggestion. [Work supported by the NSF, Grant No. CMS 99-88645.]

WEDNESDAY AFTERNOON, 4 DECEMBER 2002 CORAL GALLERY 1, 1:00 TO 2:20 P.M.

Session 3pSA

Structural Acoustics and Vibration: Energy Flow Methods in Vibroacoustic Analysis and Control II

José R. Arruda, Chair
Department of Computational Mechanics, University of Estadual de Campinas, Cidade University, Zeferino Vaz, Campinas, SP-13083-970, Brazil

Invited Papers

1:00

3pSA1. Turbulent boundary layer induced vibration up to high frequencies by means of local energy methods. Pierre Hardy, Louis Jezequel, Mohammed Ichchou, and Yves Jacques (Equipe Dynamique des Systemes et des Structures, LTDS, UMR CNRS 5513, Ecole Centrale de Lyon, 69130 Ecully, France)

The local energy method developed in the last years revealed appropriate in medium and high frequencies and supplies an accurate description of the spread of vibration and acoustic fields up to high frequencies. Our aim in the paper is to provide a complete description of the turbulent boundary layer (TBL) induced vibration by means of this method, for a simply supported thin plate. The first step in the energy method proof is the characterization of energy input from a given model of the TBL pressure interspectrum. Then, is deduced the uncoherent structural response of the panel, and the uncoherent normal mean square velocity. The latter provides, using the acoustic radiation resistance, a prediction of noise radiating by the panel up to high frequencies. Accuracy of the local energy analysis versus the usual random normal modes decomposition is demonstrated. Ultimately, a numerical parametric survey is given for various internal loss level. Precisely, the link between results provided here and SEA predictions of TBL structural induced vibration is discussed.

1:25

3pSA2. Absorption coefficient and energy flow path identification by means of inverse local energy method. Pierre Hardy, Louis Jezequel, and Mohammed Ichchou (Equipe Dynamique des Systemes et des Structures, LTDS, UMR CNRS 5513, Ecole Centrale de Lyon, 69130 Ecully, France)

Estimation of the single absorption rate and of acoustical power input is currently made in reverberant rooms where a diffuse field is established. In this paper we aim at describing a method dedicated to absorption coefficients and energy flow path identification within all type of acoustic fields in medium and high frequencies, by means of an inverse local energy method. Making use of an energy integral equation with diffuse reflection, an estimator of the wall various absorption rates is built up, while the cavity is excited by a standard spherical acoustic source. A similar formulation is used to characterize a wall continuous excitation. Then it is possible to go through the measurements (pressure, intensity) and supply a detailed analysis of the wall input energy flow. This study also includes a numerical comparison between optimization methods used when trying to match the calculated field and the reference field. Among those methods, the spheric gradient proves to be efficient when compared with estimators based on other cost-functions. Sensitivity to discretization, robustness versus statistic measurement errors, and relevance of identified parameters are dealt with.

Contributed Papers

1:50

3pSA3. Predicted and measured structural intensity in a plate excited by a diffuse acoustic field. Michael Daley and Stephen Hambric (Grad. Prog. in Acoust. and Appl. Res. Lab., Penn State, Appl. Sci. Bldg., State College, PA 16801)

Structural intensity (S-I) fields indicate energy source and sink regions in structures as well as energy flow paths through structures. Various S-I-based studies have been performed for structures excited at a point or group of points. Little attention, however, has been paid to S-I fields in structures under distributed random loads, such as those due to turbulent boundary layer pressure fields and diffuse acoustic fields. Such excitations occur in common engineering problems. This study presents experimental S-I results for a thin, simply supported, rectangular plate with an attached damper excited by a diffuse acoustic field. The plates velocity response to the excitation is measured using a scanning laser Doppler vibrometer. Finite differencing techniques are applied to the measured velocity field to yield S-I vectors. The experiments results are shown to be similar to predicted S-I fields. [Work supported by ARL Penn State.]

2:05

3pSA4. Causal energy absorption by point-attached substructures. J. Gregory McDaniel and Xianhui Li (Dept. of Aerosp. and Mech. Engr., Boston Univ., 110 Cummington St., Boston, MA 02215)

One means of attempting global vibration control when a structure is excited by an unknown transient force is to attach a substructure that optimally absorbs energy by matching the structural dynamics. Frequency-domain analysis allows one to determine the substructural admittance for optimal absorption, however the resulting admittance is often noncausal. A classic noncausal example is when the substructural admittance is chosen

as the complex conjugate of the structural admittance. This presentation will present a methodology for solving this problem by employing an implicitly causal Fourier series in frequency for the substructural impedance. The Fourier coefficients of this series are chosen to maximize the energy absorption of the attachment. This methodology also allows one to seamlessly incorporate a passivity requirement for the substructure.

WEDNESDAY AFTERNOON, 4 DECEMBER 2002 GRAND CORAL 3, 1:00 TO 3:00 P.M.

Session 3pSC

Speech Communication and Signal Processing in Acoustics: Speech and Signal Processing (Poster Session)

Maria Garcia, Cochair
VAN-A, Dept. Electron., Av. San Pablo #180, Col. Reynosa Tamaulipas, Azc. CP 02200, D.F. Mexico

John G. Harris, Cochair
Center QEFP, Northwestern University, 2137 North Sheridan Road, Evanston, Illinois 60208-3020

Contributed Papers

All posters will be on display from 1:00 p.m. to 3:00 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 2:00 p.m. and contributors of even-numbered papers will be at their posters from 2:00 p.m. to 3:00 p.m.

3pSC1. Automatic phone segment alignment using statistical deviations from manual transcriptions. Toru Hayakawa, Katsuhiko Shirai (School of Sci. and Eng., Waseda Univ., 3-4-1 Okubo, Shinjuku, Tokyo 169-8555, Japan, toru@shirai.info.waseda.ac.jp), Hiroaki Kato (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan), and Yoshinori Sagisaka (GITS/GITI, Waseda Univ., Tokyo 189-0051, Japan)

For precise temporal characteristic description, disagreements between manual labeling and automatic labeling were quantitatively analyzed with respect to the spectral feature extraction, adoption of acoustic matchers (HMM models), and acoustic matcher by itself. Error analysis shows that boundaries are shifted at phone boundaries where the speech spectrum changes quite rapidly. This disagreement results from the spectral feature extraction averaged over a given window. For the adoption of model, big errors are found at phone boundaries where the spectrum changes slowly. The third model-dependent errors are seen at phones whose duration cannot be shorter than the frame increment period times the HMM state number. To take into account these error factors individually to reduce the amount of alignment errors, we modified the automatic alignment results context-dependently using statistical characteristics of phone boundary displacement. This post-processing of boundary modification reduces boundary errors from 14.79 ms to 11.07 ms. Supplementary experiment shows that this improvement of about 4 ms corresponds to eight times of error reduction obtained by speaker adaptation of acoustic matchers. [Work supported by TAO, Japan.]

3pSC2. Duration normalization for improved automatic speech recognition. Jon P. Nedel and Richard M. Stern (Dept. of Elec. and Computer Eng. and School of Computer Sci., Carnegie Mellon Univ., Pittsburgh, PA 15213, jnedel@cs.cmu.edu)

While hidden Markov models (HMMs) serve as the basic acoustic modeling framework for many automatic speech recognition systems, they are known to model the duration of sound units poorly. Phone duration normalization can be accomplished by adding and reconstructing missing frames when a phone is shorter than the desired duration, and by deleting frames when a phone is longer than the desired duration. If phone segmentations are known *a priori*, this technique achieves relative reductions in word error rate (WER) of up to 35%, confirming the conjecture that speech with normalized phone durations may be modeled better and discriminated more accurately using standard HMM acoustic models. Unfortunately, duration normalization using imperfect automatically generated phone segmentations has not yielded significant recognition improvements. A modification of the duration normalization approach has been developed. Three different feature streams are generated for each utterance using various combinations of expansion and contraction of hypothesized phone segments. Each stream is recognized using an acoustic model trained for that stream. While the resulting recognition hypotheses themselves are not significantly better than baseline, these hypotheses can be automatically combined to produce relative improvements in WER of up to 7.7% over several speech databases. [Work supported by DARPA and Telefónica.]

3pSC3. Fast on-line speaker/environment adaptation using modified maximum likelihood stochastic matching. Shubha L. Kadambe (HRL Labs., LLC, 3011 Malibu Canyon Rd., Malibu, CA 90265) and Marcus Iseli (UCLA, Westwood, CA)

The problem of speaker/environment adaptation to improve the recognition accuracy and thus making recognizers robust is addressed here. For this, a fast on-line adaptation algorithm that does not need a separate adaptation training data and that adapts acoustic models fast enough to achieve near real-time recognition is developed. This technique is based on stochastic matching in the model space similar to [A. Shankar and C.-H. Lee, IEEE Trans. Signal Process. **4**, 190–202 (1996)]. For fast adaptation only the models and the mixture components that need to be adapted are selected based on the cluster formation and Euclidean distance. This adaptation algorithm is implemented as part of a GMM based continuous speech recognizer. It is tested using a non-native speakers dataset. For example, the five best hypotheses output of the speech recognizer before and after applying the adaptation technique indicated that the right answer corresponding to an utterance ''none of the earth'' before adaptation did not correspond to the best hypothesis; however, it corresponded to the third best. After adaptation all of the five-best hypotheses converged to the right answer. The test results of this technique on a larger non-native speakers' dataset shows 70% to 75% relative WER improvement.

3pSC4. Cues for question intonation in Arabic: Disambiguation techniques for use in automatic speech recognizer systems. Leslie Barrett (Transclick, Inc., New York, NY 10021) and Kazue Hata (Univ. of California, Santa Barbara, CA 93106)

The focus of this study is to determine the extent to which prosodic characteristics can contribute to the improvement of speech recognition in Arabic. $F0$ rising rate was chosen to disambiguate yes–no question from declarative sentences. In Arabic, as in English, a rising intonation indicates a yes–no question whether the question takes lexical question markers or uses an inverted word order or whether the sentence takes just a declarative form. We conducted a production study with 55 yes–no question sentences uttered by a female native speaker of Arabic. Two types of measurements were taken for $F0$ rise rate. First, we visually obtained the best-fit rise in sentence-final position. Second, we computed the rate based on minimum and maximum $F0$ values within the sentence-final 500 ms. The results show that although the rise obtained from the final 500 ms (0.41 Hz/ms) is different from the best-fit rise rate (0.49 Hz/ms) ($p < 0.05$), when examining two different $F0$ rising shapes and considering JND for the rising rate (Nabelek and Hirsh, 1962), a threshold of 0.4 Hz/ms can be considered a threshold indicator of a yes–no question intonation for this speaker. Thus, a sufficient rise rate was extracted automatically from a fixed sentence-final duration.

3pSC5. A descriptive analysis between syllables in English and Spanish. Fabiola M. Martinez Licona, John Ch. Goddard Close, and Alma E. Martinez Licona (Dept. of Elect. Eng., Universidad Autónoma Metropolitana Iztapalapa, Mexico City, Mexico 09340, fmml@xanum.uam.mx)

In recent years different attempts have been made to incorporate temporal information longer than the phoneme into automatic speech recognizers (ASR) for English. The reason for these approaches is related to limitations which arise with existing systems based on phonemes, such as the degradation in performance of ASRs under noisy conditions and variations in pronunciation due to phoneme omission. It is conjectured that humans naturally use longer time periods, corresponding, for example, to syllables, to perceptually integerate information. In the case of Spanish, little seems to have been done in this direction for ASRs. In the present paper, a preliminary comparison is made between syllables in Spanish and English with a view to their factibility in an ASR for Spanish. In particular, a descriptive statistical analysis is conducted with a Spanish speech database to derive the most common structures of the syllables and the most common monosyllables. This is contrasted with previously found results in English. Spectrograms are also used to illustrate pertinent characteristics of Spanish. These results suggest that syllables may indeed provide useful information for an ASR in Spanish and could provide greater success than their counterparts in English. [Work supported by CONACYT under Project 31929-A.] (To be presented in Spanish.)

3pSC6. Low bit rate speech coding using spectral trajectory modeling. Sorin Dusan and James Flanagan (Ctr. for Adv. Information Processing, Rutgers Univ., 96 Frelinghuysen Rd., Piscataway, NJ 08854, sdusan@caip.rutgers.edu)

In the U.S. Federal Standard coder for 2400 bps, a data frame containing 54 bits of encoded signal is transmitted every 22.5 ms. In each frame, 25 bits encode the spectral features (10 Line Spectrum Frequencies—LSF). In this paper we describe a method for reducing the transmission rate while preserving most of the quality and intelligibility. This method is based on modeling the spectral trajectories with polynomial functions and on encoding these functions for segments of speech extending over multiple frames. Here 10 polynomials are computed by fitting them to the 10 LSF trajectories in the least-squares sense. Then the polynomial coefficients are encoded for the whole segment instead of directly encoding the LSF vectors. The spectral parameters are thus reduced (compressed) to $[(P+1)/N]\times 100\%$, where P represents the order of the polynomials and N the number of frames for each segment. Different compression rates can be achieved. For example, for $P=5$ and $N=10$ the spectral features are encoded using 40% less bits than those required to encode $N=10$ LSF vectors. For this example, the overall transmission rate is reduced to 1956 bps, yielding an average Itakura–Saito spectral distance of 0.14 between the original and reconstructed LPC parameters computed for a typical utterance.

3pSC7. Conveying discourse structure in synthetic speech. Peter C. Gordon and Harry Halpin (Univ. of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3270, pcg@email.unc.edu)

A pre-processing system for assigning prosodic characteristics of speech was created in order to investigate how synthesized speech can capture the global characteristics of discourse structure. The system allows manipulation of a variety of prosodic characteristics that have been tied to the structure of discourse above the level of the sentence, in particular: pitch fluctuation, pitch range, speaking rate, and pauses. These characteristics can be used to highlight the onset of a discourse segment, defined as a group of utterances that contribute to a single discourse purpose. After highlighting the onset of the discourse segments, these speech characteristics are progressively modified to indicate the continuity of the discourse segment. The system was evaluated by comparing it to a corpus of natural speech, the Boston Directions Corpus. That corpus has been analyzed both in terms of the informational content of the discourse and the acoustic manifestations that appear in natural speech for conveying that content. Ways in which those acoustic manifestations can be realized in natural speech are discussed.

3pSC8. Intelligibility tests for synthetic speech subjective evaluation: The semantically unpredictable sentences approach for European Portuguese. Daniela Braga, Luís Coelho, António Moura, and Diamantino Freitas (Faculty of Eng. of Univ. of OPorto, R. Dr. Roberto Frias, OPorto, Portugal)

In this paper a test proposal for European Portuguese (EP) synthetic speech quality evaluation as well as some of its application are presented. For this purpose we have built a semantically unpredictable sentences (SUS) corpus of 25 sentences with different lengths that were chosen according to the standard EP phonological, syllabic and prosodic specific features. The goal is to create a standard test linguistically focused that constitutes a base for a more accurate comparison and ranking of different synthesis techniques. The tests were conducted in two distinct conditions: one performed in a controlled environment with listeners ranging from 16 to 60 years old; the other performed through the web by a large group of listeners of different ages, each one in their own environment. The listeners had only one chance of listening to each sentence. Of course in the second case this could not be controlled. The evaluation parameters considered were the number of (1) totally well identified sentences; (2) non-identified sentences; (3) misunderstood words in each sentence; (4) misunderstood words in the whole corpus; (5) misunderstood words according to the length of the sentence; (6) errors in the perception of a word according to its position in the sentence.

3pSC9. An estimation method for fundamental frequency and voiced segment in infant utterance. Tomohiro Nakatani, Shigeaki Amano, and Toshio Irino (NTT Commun. Sci. Labs., NTT Corp., 2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan)

Fundamental frequency ($F0$) and unvoiced/voiced segment (U/V) estimation of infant utterances are important for investigating humans perception of prosodic information in an early stage of speech communication. However, this estimation process is difficult as infant utterances have several features that differ to those of adults: (1) $F0$ has a wide range in value (200 to 2000 Hz); (2) $F0$ is unstable, for example, it discontinuously changes to its double or half value; and (3) voiced segments may have high energy in the higher frequency regions degrading U/V decisions of

existing methods. Additionally, infant utterance data is often collected in daily child care settings which lowers the signal-to-noise ratio (SNR). To cope with these problems, a robust $F0$ estimation method based on instantaneous frequency [Nakatani and Irino, ICSLP2002] is introduced, and a new U/V detection method is proposed. The former has a mechanism to extract accurate $F0$ avoiding double and half pitch errors in low SNR environments. Once accurate $F0$ is obtained, the latter method can reliably detect U/V just by examining the harmonic structure corresponding to the $F0$. The effectiveness of this method is examined using a database devised from infant utterances in daycare settings [Amano, Kato, and Kondo, ICSLP2002].

3pSC10. Dynamic constraints on the inverse problem in speech. Khalil Iskarous (Haskins Labs., 270 Crown St., New Haven, CT 06511)

The problem of determining the area function from the speech signal has long been known to be an ill-posed problem—there are many area functions that correspond to the same speech spectrum. To make the problem well-posed, the inversion process must be constrained. In this work, inversion is performed for a dynamic formant pattern of a CV or a VV transition as a whole, rather than for a single static pattern. The input to the process are formant values along with the first and second time derivatives of each formant for each frame of the formant pattern. This provides more information about the possible area function solutions, and thereby acts as a constraint. Also the area function change during the transition is constrained to be stationary at one location in the vocal tract, while maximal change occurs at only two other locations, with the area increasing at one of the locations and decreasing at the other. This constraint is derived from an empirical study of area function change. These constraints are implemented as constraints on a Riccati recursion for the reflection coeficients. Comparison with other work on dynamic constraints on speech inversion will be provided. [Work supported by NIH.]

3pSC11. Auditorily motivated elastic spectral distance and its application to emotional morphing of portrayal speech. Hisami Matsui and Hideki Kawahara (Faculty of Systems Eng., Wakayama Univ., 930 Sakaedani, Wakayama 640-8510, Japan)

An elastic spectral distance measure based on a $F0$ adaptive pitch synchronous spectral estimation and smoothing, that is developed for a high-quality speech modification procedure STRAIGHT [Kawahara *et al.*, Speech Commun. **27** (1999)], is introduced to provide a basis for emotional morphing. The proposed measure is based on a smoothed nonlinear frequency mapping between the target and the original speech spectra on the ERB axis. A portrayal emotional speech database, which was recorded using professional actors and a recording studio for professional use, was designed for developing and evaluating morphing functions. Target words were recorded under four different contextual conditions (preceding, following and franking carrier sentences and isolated pronunciation), two sentence types (declarative and interrogative). Perceptual effects of constituent parameters of the proposed distance measure will be discussed based on morphing experiments using STRAIGHT procedure. [Work supported by JSPS and NIME, Japan.]

3pSC12. Evaluation of a strategy for automatic formant tracking. Terrance M. Nearey (Dept. of Linguist., Univ. of Alberta, Edmonton, AB T6G 0A2, Canada), Peter F. Assmann (Univ. of Texas, Dallas, Richardson, TX 75083), and James M. Hillenbrand (Western Michigan Univ., Kalamazoo, MI 49008)

Variations on an automatic formant tracking strategy developed at Alberta will be compared to manual formant measurements from two databases of vowels spoken by men, women, and children (in Texas or Michigan). "Correct" vowel formant candidates for $F1$, $F2$, and $F3$ may be found roughly 85–90 percent of the time for adult male speakers using autocorrelation LPC with the following settings: $F3$ maximum at 3000 Hz, LPC order of 14, sampling rate of 10 kHz [J. Markel and A. Gray, *Linear Prediction of Speech* (Springer, New York, 1975)]. Experience shows good results are also often found with females' and children's speech, provided the sampling rate and $F3$ maximum are scaled appropriately for each speaker. Our new basic strategy involves analyzing each utterance at several distinct sampling rates and coordinated $F3$ cutoff frequencies with a fixed LPC order. Each scaling choice provides an independent set of candidates that is post-processed by a simple tracking algorithm. A correlation measure between a spectrogram of the original signal and one resynthesized from each estimated track set is defined. This measure is combined with other heuristic figures of merit (based on, e.g., continuity, formant ranges, and bandwidths) to choose the "best" analysis.

3pSC13. Speech enhancement with beamforming enhanced by parametric spectral subtraction. Jaeyoun Cho and Ashok Krishnamurthy (Dept. of Elec. Eng., The Ohio State Univ., 2015 Neil Ave., Columbus, OH 43210, krishnamurthy.1@osu.edu)

This paper describes a novel method of speech enhancement that combines multichannel beamforming techniques with single channel spectral subtraction. Spectral subtraction methods have been widely used for speech enhancement because they require only a single microphone and provide moderate improvements in the signal-to-noise ratio (SNR). But a major drawback of spectral subtraction is that it inevitably introduces *musical noise*, which is nonstationary and very annoying. Generally, the musical noise is reduced by smoothing its spectral variance in the frequency domain. Multimicrophone beamforming techniques, on the other hand, enhance the SNR by using spatial and temporal filters to introduce spatial selectivity towards the desired signal. It is shown that these two techniques are complementary, and can be combined to provide better performance than either method alone. One approach, for example, is to use spectral subtraction as a preprocessor in each channel before beamforming. In this case, the summing process of beamforming reduces the musical noise generated by spectral subtraction. We provide examples of using this approach in an automotive application.

3pSC14. Adaptive beamformer based on average vowels/consonant spectrum weights for noisy speech recognition. Masato Nakayama, Takanobu Nishiura, and Hideki Kawahara (Grad. School of Systems Eng., Wakayama Univ., 930 Sakaedani, Wakayama 640-8510, Japan)

Background noise and reverberations seriously degrades the sound capture quality. A microphone-array is an ideal candidate for capturing distant-talking speech. With a microphone array, a desired speech signal can be acquired selectively by steering the directivity. The AMNOR (Adaptive Microphone-Array for Noise Reduction) is an adaptive beamformer proposed by Kaneda *et al.* In addition, as the beamformer for speech capture, S-AMNOR, the AMNOR with a long time speech spectrum was also proposed by Okada *et al.* However, the performance of the S-AMNOR may be further improved, if each adaptive filter for vowel and consonants could be designed with average vowels/consonants spectrum. Therefore, we propose the new AMNOR with adaptive filters for vowels/consonants, in order to improve the signal capturing performance. We evaluated the ASR (Automatic Speech Recognition) performance with the enhanced desired signal using the adaptive filters for vowels/consonants after detecting vowels and consonants on each phoneme. As a result of evaluation experiments, by comparing the results from the proposed AMNOR and the conventional AMNOR/S-AMNOR, we could confirm that the ASR performance was improved with proposed AMNOR. [Work supported by JSPS.]

3pSC15. Cross spectral measurement of head related speech transfer functions using speaker's own voice. Masumi Nukina and Hideki Kawahara (Grad. School of Systems Eng., Wakayama Univ., 930 Sakaedani, Wakayama 640-8510, Japan)

A cross spectrum method is applied to measure sound pressure variations around the head using the speaker's own speech sounds. The variations are represented as transfer functions from the mouth reference point to a set of measuring points. Preliminary tests indicated that there are systematic frequency response variations depending on vowel colors. This vowel color dependency was not replicated in the classical measurement of speech radiation characteristics by J. L. Flanagan. However, taking into account the large (sometimes exceeding 20 dB) amount of variations, it is not likely to be negligible. A set of calibration and normalization procedures were introduced to reduce artifacts due to background noise, room acoustics, zeros in the speech spectra. A series of M-sequence based transfer function measurements were also conducted using a head and torso simulator to evaluate intrinsic errors in the cross spectral measurements. It was found that the standard errors in the cross spectral measurements using recorded speech sounds are around 1 dB. Based on these reference data and confidence interval calculations based on coherence, it is safe to conclude that the vowel color dependency is significantly modifying the transfer functions. [Work supported by JSPS.]

3pSC16. Real-time magnetic resonance imaging for the study of speech production. Alain Soquet, Didier Demolin (Laboratoire de Phonologie, Universit Libre de Bruxelles, 50 av. F. D. Roosevelt, 1050 Brussels, Belgium), Peter Branderud (Stockholm Universitet, Stockholm, Sweden), Bjorn Linblom (Univ. of Texas, Austin, TX 78712), and Thierry Metens (Hopital Erasme, Universit Libre de Bruxelles, Bruxelles, Belgium)

We have shown [D. Demolin *et al.*, C. R. Biol. **325**, 111 (2002)] that it is possible to record in real-time the MR images and the speech signal. This technique opens new perspectives for the study of speech production. For acquisition of the speech signal, we have used an optical microphone. The principle of the optical microphone is as follows. A first optical fibre carries an incident light produced by a light source. The light is then reflected by a diaphragm and transmitted through a second optical fibre to a photodetector. When the diaphragm moves according to the ambient sound, it modulates the amount of light reaching the photodetector. This kind of microphone has two major advantages. First, it does not contain any metallic part; it can therefore be placed close to the lips without any danger or artefact. Second, the signal is carried with optical fibres; this allows the electronic and the recording device to be placed outside the acquisition room. Real-time MRI technique allows to explore movement of articulators involved during normal speech, while dynamic MRI relies on numerous repetition of the same sequence to reconstruct the impression of movements in time.

3pSC17. An evaluation of talker localization based on direction of arrival estimation and statistical sound source identification. Takanobu Nishiura (Faculty of Systems Eng., Wakayama Univ., 930 Sakaedani, Wakayama 640-8510, Japan) and Satoshi Nakamura (ATR Spoken Lang. Translation Res. Labs., Seika-cho, Soraku-gun Kyoto 619-0288, Japan)

It is very important to capture distant-talking speech for a hands-free speech interface with high quality. A microphone array is an ideal candidate for this purpose. However, this approach requires localizing the target talker. Conventional talker localization algorithms in multiple sound source environments not only have difficulty localizing the multiple sound sources accurately, but also have difficulty localizing the target talker among known multiple sound source positions. To cope with these problems, we propose a new talker localization algorithm consisting of two algorithms. One is DOA (direction of arrival) estimation algorithm for multiple sound source localization based on CSP (cross-power spectrum phase) coefficient addition method. The other is statistical sound source identification algorithm based on GMM (Gaussian mixture model) for localizing the target talker position among localized multiple sound sources. In this paper, we particularly focus on the talker localization performance based on the combination of these two algorithms with a microphone array. We conducted evaluation experiments in real noisy reverberant environments. As a result, we confirmed that multiple sound signals can be identified accurately between ''speech'' or ''non-speech'' by the proposed algorithm. [Work supported by ATR, and MEXT of Japan.]

WEDNESDAY AFTERNOON, 4 DECEMBER 2002 CORAL GARDEN 2 AND 3, 1:00 TO 3:05 P.M.

Session 3pSP

Signal Processing in Acoustics and Underwater Acoustics: Acoustic Automatic Target Recognition II

Ning Xiang, Cochair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Armando Andrade, Cochair
Calle H, No. 11, Col. Jardines de Santa Clara, Ecatepec Edo. De. CP 55450, Mexico

Chair's Introduction—1:00

Invited Papers

1:05

3pSP1. Outdoor experiment results for tracking airborne high-speed broadband acoustic sources. William G. Frazier, Chad Williams, Jay E. Williams (Miltec Res. and Technol., Inc., NCPA, Coliseum Dr., University, MS 38677), and Kenneth E. Gilbert (The Univ. of Mississippi, University, MS 38677)

Recent results of tracking airborne high-speed (subsonic) broadband acoustic sources with an experimental system are presented. The system consists of three acoustic arrays, wireless telemetry systems, real-time signal processing and tracking processors, and a tactical operation center-style display unit. The acoustic signal processor consists of a coherence-based detector and generalized cross

correlator. The real-time tracking processor uses a modified extended Kalman filter algorithm. In this particular work, the calculation of acoustic sensor-to-source-bearing information will be briefly discussed along with the operation of the source detector, which is adaptive and dependent upon both amplitude and signal coherence. Special emphasis will be placed on how the tracking algorithm is compensated to accommodate unknown and variable data latencies produced by acoustic signals propagating from distant sources that are moving at a significant fraction of the speed of sound. Experimental results include estimates of source position, speed, and heading versus time. Estimates of source state uncertainty and comparison to truth data are also provided. [Work supported by the U.S. Army Space and Missile Defense Command.]

1:35

3pSP2. Audible and visual representations of the signals from a seismic landmine detection system. Waymond R. Scott, Jr. (School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332, waymond.scott@ece.gatech.edu), Gregg D. Larson, and James S. Martin (Georgia Inst. of Technol., Atlanta, GA 30332)

A system using high frequency seismic waves to detect buried landmines is being investigated using both experimental and numerical models. The simplest detection cue for the localization of buried landmines has been found to be resonances of the trigger mechanism and overlying soil as these resonances create substantial localized surface displacements over the buried landmine in comparison to other buried objects such as rocks, sticks, and manmade clutter. The prototype system utilizes a noncontact radar sensor to detect surface–normal displacements and a remotely-located surface-contacting source to generate seismic waves in the ground. Processing of the measured data shows the location of mines both graphically and audibly, either of which could be used as inputs for automatic target recognition algorithms. Visual images created by a wavenumber-domain signal processing algorithm and audible representations will be compared for landmines and typical clutter objects. [Work supported by ONR and ARO.]

2:05

3pSP3. Automatic acoustic mine detection using morphological perceptions. Gerhard X. Ritter, Paul D. Gader, A. Koksal Hocaoglu, and Laurentiu Iancu (Dept. of CISE, Univ. of Florida, Gainesville, FL 32611)

Recent developments in landmine detection based on the acoustic-to-seismic coupling phenomenon have demonstrated the feasibility to detect both metallic and nonmetallic mines. In this method, a loudspeaker above the ground surface insonifies the target region of the surface. Acoustic energy is coupled into the ground producing ground vibrations. The ground vibration velocity at the ground surface is measured with a laser Doppler vibrometer producing a ground surface image whose pixel values are the measured velocity amplitudes. Vertical particle velocity amplitudes directly above a mine contrast with those of the background (''away from the mine''). Image processing techniques are used in order to automatically detect regions of interest (''possible mines''). Further image analysis methods extract geometric as well as amplitude features in order to produce four-dimensional feature vectors. These vectors are input to a novel neural network based on mathematical morphology which classifies the regions of interest into mines and false alarms. After training the network, the network correctly identified all the mines on a given test set with an extremely low alarm rate.

2:35

3pSP4. (Spectral) pattern recognition as a versatile tool towards automatic landmine detection: A new European approach. Volker Klein (Kayser-Threde GmbH, Munich, Germany, kv@Kayser-Threde.de), Peter Lutzmann (FOM Ettlingen, D-76275, Ettlingen, Germany), and Thomas Mechnig (Polytech GmbH, D-76337 Waldbronn, Germany)

A mobile acousto-optical sensor (Laser Vibrometer) is being used for the detection and discrimination of buried landmines. Analysis of measurement data, obtained in a number of field tests, reveals that buried mines (anti-tank mines as well as anti personnel mines) can reliably be discriminated from nonlethal clutter objects, such as stones, wood, cans, etc. due to their individual shapes and spectral properties. Increasing operational demands on these acoustic soundings (speed, reliability, and spatial/spectral resolution) implies the introduction of new procedures and strategies to identify and suppress misleading features as well as improved correlation of useful data. These procedures will include improved techniques for real time: (a) processing of the spectrally resolved image of the

soil surface as well as (b) spectral analysis of the on-target soil vibrations. Image processing is comprising background linearization and suppression, shape recognition and gradient detection. Spectral analysis is including the use of look-up tables for cross-correlation and identification of characteristic features. This paper is presenting an introduction to the detection technique, sounding strategies as well as subsequent spectral data evaluation toward automated target recognition.

WEDNESDAY AFTERNOON, 4 DECEMBER 2002 GRAND CORAL 1, 1:00 TO 3:15 P.M.

Session 3pUWa

Underwater Acoustics and Engineering Acoustics: Underwater Acoustic Measurement Laboratories: New Global Perspectives of Automatic Control and Management

Carlos Ranz-Guerra, Chair
Instituto de Acoustica, CSIC. C Serrano 144, 28006 Madrid, Spain

Chair's Introduction—1:00

Invited Papers

1:05

3pUWa1. Remote full control, by an Internet link, of an underwater acoustics laboratory. Carlos Ranz-Guerra, Pedro Cobo-Parra, Manuel Siguero-Guerra, and Alejandro Fernandez-Fernandez (Instituto de Acustica, CSIC. C/ Serrano 144, 28006 Madrid, Spain)

The Underwater Tank Laboratory located at the Instituto de Acustica, CSIC, Madrid, has been fully reshaped. Now, the two bridges (emission and reception) have full automatic motion control by the operator. These capabilities were complemented by a new management of signal generation, signal acquisition, processing and storing of data. This new framework makes many of the tasks to be performed in this kind of facility easier by putting at the hands of the operator specific friendly software programs that attend to the main aspects of the ongoing experiment. In one step forward, the remote control of all the functionalities was considered feasible. The potentialities of the Internet were thought to provide a new dimension to the laboratory by lowering the difficulties of taking over the full control of the installation, by any user around the world. Here is one real example of how this achievement can be carried out. The Underwater Acoustics Laboratory at the Instituto de Acustica, CSIC, is now ready to be run by any one interested. The main lines, over which this problem has been considered, are described in this paper. [Work supported by PN on Science and Technology and CSIC, Spain.]

1:25

3pUWa2. Hydroacoustic station network for monitoring the Comprehensive Nuclear-Test-Ban Treaty (CTBT). Martin W. Lawrence, Marta Galindo Arranz, Patrick Grenard, and John Newton (CTBTO, Vienna Intl. Ctr., Vienna, Austria, martin.lawrence@ctbto.org)

The Comprehensive Nuclear-Test-Ban Treaty (CTBT) provides for monitoring of the whole globe by a network of stations, using various technologies, in order to verify the absence of nuclear explosion tests. The hydroacoustic component of this network, which monitors the major world oceans, is currently under construction. When complete it will consist of 11 stations located with an emphasis on the vast ocean areas of the Southern Hemisphere. Presently, three stations have been completed and work is underway on all of the remaining stations. The stations transmit real-time continuous data to the CTBT Organization headquarters in Vienna, Austria. The hydroacoustic network uses two different types of stations. One type is based on hydrophones floated from the sea floor to the SOFAR axis depth, arranged horizontally in a triplet configuration. The other type is based on the use of seismometers located on small islands to detect hydroacoustic signals after conversion to seismic signals at the flanks of the island. During the time since completing the first stations, many interesting acoustical phenomena have been observed in the data.

1:45

3pUWa3. A network, via www, linking laboratories of underwater acoustics. Vicente Gallego (ETSI de Armas Navales, Arturo Soria, 287, Madrid 28033, Spain)

The High Technical School for Naval Weapons Engineers is a teaching institution of the Spanish Navy dedicated to training advanced engineers in the related technologies to the systems of naval weapons. Some scientific disciplines, like the underwater acoustics, are also included in its curriculum. The tasks of research are being developed by means of collaborations with civil organisms such as Institutes of Scientific Research and departments of some universities. In this context, a proposal has been prepared for creating a network for laboratories of underwater acoustics. It was initially formed by the interconnection via www of the Laboratory of Underwater Acoustics in the Naval Weapon School and the Laboratory of hydroacoustics in the Institute of Acoustics, a center of scientifical research pertaining to the Science and Technology Ministry. The joint exploitation of the diverse capabilities from different laboratories, the rational integration of activities and the saving of efforts and costs are the basic ideas for this project. In this paper are presented the general lines that have been considered in the development of this proposal, the future possibilities of using for a network like this as well as the practical limitations addressed by this new concept of networks of laboratories.

2:05

3pUWa4. Acoustical contamination: Effects on the population of mular dolphin and on their habitats. Ricardo Hernandez, Renau De Stephanis, Jose Luis Cueto, and Rivas Neus (Acoust. and Vib. Lab., Univ. of Cadiz, lav@uca.es)

In this conference we present the part of the research which has been carried out since July 2002, inside the European Project LIFE-Nature 2002 Ref. LIFE 02/ NAT/E/8610. The research aims at the anthropogenic activities existing in the research area that could affect the studied species in a different way. One of the threats is the acoustical contamination, very strong in the research area due to intensive maritime traffic, especially in the area of Gibraltar. It would be extremely useful to establish the common data base in order to permit the researchers to exchange the data obtained, tests details, and experience. In our case it would be very interesting to follow this type of marine species.

2:25

3pUWa5. Ocean acoustic laboratory at the Pacific Missile Range Facility. Peter Stein, Jason Rudzinsky, Subramaniam Rajan (Sci. Solutions, Inc., 99 Perimeter Rd., Nashua, NH 03063, pstein@scisol.com), and James Lewis (Sci. Solutions, Inc., Kalaheo, HI 96741)

An ocean acoustic laboratory (OAL) is being implemented at the Pacific Missile Range Facility (PMRF). The range covers 1100 sq nm off the west coast of Kauai, HI, with water depths ranging from very shallow to abyssal plane. There are 172 bottom-mounted hydrophones and 15 bottom-mounted sources permanently cabled to shore. The range is used extensively for training and test and evaluation. The general OAL concept is to improve the range products by providing an accurate depiction of the four-dimensional ocean environment. A high-resolution ocean model has been implemented for the region with 48-h forecasts available on the Scientific Solutions web site (www.scisol.com/hawaii). We are currently implementing a tomographic imaging capability using the bottom-mounted sources and receivers. These sound-speed images will be assimilated into the ocean model to improve accuracy. The long-term goals include providing a real-time picture of this well-described ocean environment over the global network. Applications of the ocean acoustic laboratory include oceanographic research, ocean acoustic research, system evaluation, training, and virtual war fighting. [Work supported by ONR and CEROS.]

Contributed Paper

2:45

3pUWa6. New underwater acoustic tank facility at Georgia Tech. Michael Gray, Ralph Herkert, George McCallII, Gary Caille (Georgia Tech Res. Inst., Atlanta, GA 30332, michael.gray@gtri.gatech.edu), Van Biesel, John Bogle, Jayme Caspall, Steven Hahn, Adam Lamb, Thomas Logan, James Martin, Peter Rogers, and David Trivett (Georgia Tech, Atlanta, GA 30332)

A large underwater acoustic tank facility located in the Woodruff School of Mechanical Engineering at Georgia Tech has recently been completed. The facility includes a rectangular concrete water tank 25 feet deep, 25 feet wide, and 34 feet long containing around 160,000 gallons of water. There are three computer-controlled positioners: an x-y-z-θ positioner and a z-θ positioner mounted on carriages and a bottom mounted rotator. The facility has a large rectangular nearfield array which can be used either as a receiver or a transmitter. A single vertical nearfield line array can be translated by the x-y positioner to synthesize a cylindrical nearfield receiving array. The rectangular nearfield transmitting array and the synthesized cylindrical receiving array were designed to be used with the bottom mounted rotator to measure the true farfield bistatic target strength of any target up to one meter in length as a function of the target aspect angle. Such measurements can be done from 2 kHz to over 10 kHz. The tank is being used for transducer development, materials, and flow noise studies in addition to structural acoustics. Several available multi-channel data acquisition systems will be described. [Work supported, in part, by a DURIP grant from ONR.]

3:00–3:15
Panel Discussion

Session 3pUWb

Underwater Acoustics: General Topics in Underwater Acoustics I

George E. Ioup, Chair
Department of Physics, University of New Orleans, New Orleans, Louisiana 70148

Contributed Papers

1:00

3pUWb1. Use of the internet in maintaining and operating the Intermediate Scale Measurement System (ISMS) underwater test facility. Duane Nightingale, Steve Troxel, and Bruce Libby (Naval Surface Warfare Ctr., Acoust. Res. Det., 33964 N. Main St., Bayview, ID 83803, nightingaledl@nswccd.navy.mil)

The Intermediate Scale Measurement System (ISMS) is the Navy's premier underwater measurement facility for performing target strength and structural acoustic evaluations on 1/4 scale submarine models. Located in Bayview, Idaho, ISMS reaches out to its customers through extensive use of the internet and open source technologies. Central to the operation of the system is an apache web server used to monitor and control systems within the models and the land-based facilities. Many aspects of ISMS, including data acquisition can be performed remotely over the internet. Specific topics include overcoming firewall issues, remote monitoring and control, collaborative test development and execution, and open source tools.

1:15

3pUWb2. Generic file format. Nick Felgate (Defence Sci. and Technol. Lab., Farnborough, Hampshire GU14 OLX, UK)

The Generic File Format (GFF) is a file format developed within the UK ASW community for the interchange and storage of underwater sonar data. Originally developed for the interchange of time-series data between analysis systems, it has been extended to provide for storage of processed acoustic data (e.g., power and DEMON spectrum, lofargram grey-scale), nonacoustic data (e.g., own-ship dynamics, sensor configuration) and event data (e.g., tracker output, sonar intercepts). The format employs the chunk concept, as used in the WAV and AIFF file formats, to provide extendability (including local variants) while providing a measure of backward compatability. However, the basic concept has been adapted to allow for the mixing in the one file of multiple channels of different sample-rates and data-types through the inclusion of a data frame concept and multiple data blocks. Chunk cross-referencing has been employed to ensure data consistency. A provision is made in the header of the file to store details of the sensor and processing for the data (e.g., the number of hydrophones, beam direction, FFT size) so that an analysis system does not need to know about the sensor or other system from which the data originated.

1:30

3pUWb3. Wavelet denoising of underwater acoustic data. Juliette W. Ioup and George E. Ioup (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148)

The Littoral Acoustic Demonstration Center (LADC) deployed three Environmental Acoustic Recording System (EARS) buoys in the northern Gulf of Mexico during the summer of 2001. The buoys recorded frequencies up to 5859 Hz continuously for 36 days. The acoustic signals measured include sperm whale vocalizations and seismic airguns. The data are analyzed using time series, Fourier transforms, and spectrograms. Noise removal or denoising of these data at selected times using wavelets is investigated. The effectiveness of the technique is influenced by the choice of wavelet, the decomposition level, and the threshold amplitude. Results will be presented of wavelet denoising of the LADC Summer 01 data and compared with Fourier noise removal methods including bandpass filtering and thresholding. [Research supported by ONR.]

1:45

3pUWb4. Organic sonobuoy ranging. Nick Felgate (Defence Sci. and Technol. Lab., Farnborough, Hampshire GU14 OLX, UK)

It is important that military vessels periodically check their passive signatures for vunerabilities. Traditionally, this is undertaken on a fixed range (e.g., AUTEC, BUTEC) with low noise conditions. However, for

operational and cost reasons it is desirable to be able to undertake such measurements while the asset is operating in other areas using expendable buoys deployed by the vessel itself. As well as the wet-end hardware for such organic sonobuoy ranging systems (e.g., calibrated sonobuoys, calibrated data uplink channels), careful consideration is needed of the signal-processing required in the harsher environmental conditions of the open ocean. In particular, it is noted that the open ocean is usually much noisier, and the propagation conditions more variable. To overcome signal-to-noise problems, techniques such as Doppler-correction, zero-padding/peak-picking, and noise estimation/correction techniques have been developed to provide accurate and unbiased estimates of received levels. To estimate propagation loss for source level estimation, a model of multipath effects has been included with the ability for analysts to compare predicted and observed received levels against time/range and adjust modeling parameters (e.g., surface loss, bottom loss, source depth) to improve the fit.

2:00

3pUWb5. A resonant scattering formalism for a fluid loaded elastic spherical near a planar boundary. Garner Bishop (Naval Undersea Warfare Ctr. Div. Newport, Newport, RI 02840, bishopgc@npt.nuwc.navy.mil)

Free field scattering from stationary fluid loaded elastic targets is well known and understood in terms of free field resonant scattering theory (RST), that it is given by the superposition of resonant and background components, and that the resonant components can be isolated by subtraction of an appropriate background. Backgrounds for isolating free field resonance spectra are well known. However, RST for scattering from a fluid loaded elastic target near a planar boundary where target scattering is accompanied by target-boundary scattering is less well known. A *T*-matrix formalism for plane wave scattering from rigid and soft spheres near a planar penetrable boundary is generalized to include a fluid loaded elastic spherical shell and RST is used to replace the free field *T*-matrix for the elastic shell that occurs within the formalism. It is shown that target-boundary scattering couples resonant and background components, couples modal resonances, and produces super-resonances. A background is introduced and it is shown that while background scattering is suppressed, coupled background and resonant scattering is not. Numerical results demonstrate some of the effects of target boundary scattering on the free field resonant spectrum of the shell as well as the dependence of target-boundary scattering on target.

2:15

3pUWb6. Pressure and velocity fields produced by an underwater explosion. Kendall S. Hunter and Thomas L. Geers (Dept. of Mech. Eng., UCB427, Univ. of Colorado, Boulder, CO 80309)

A model for a moderately deep underwater explosion bubble has recently been developed that integrates the initial shock wave phase and the subsequent oscillation phase of the motion [T. L. Geers and K. S. Hunter, J. Acoust. Soc. Am. **111**, 1584–1601 (2002)]. A hyperacoustic volume-acceleration model for the shock wave phase employs an empirical pressure-profile expression for the far field to determine bubble radius as a function of time, thereby providing initial conditions for the subsequent oscillation phase. The model for the oscillation phase yields bubble-response equations that incorporate first-order wave effects in both the external liquid and the internal gas. In this paper, a radiation model consistent with the bubble model is formulated and evaluated. The far-field pressure and velocity fields produced during the shock wave phase derive from the previously mentioned pressure-profile expression. Those produced during the oscillation phase derive from a matched-asymptotic-expansion solution for a translating acoustic source/dipole [F. G. Leppington and H. Levine, Proc. R. Soc. London, Ser. A **412**, 199–221 (1987)]. Computed far-field pressure and velocity histories and snapshots are shown. Comparisons with existing data exhibit good agreement. [Research sponsored by ONR.]

2:30

3pUWb7. Utilizing acoustic suspended-sediment measurement techniques in laboratory flumes. Daniel Wren, Srikanth Vadakapurapu, James Chambers (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677), Roger Kuhnle (USDA-ARS Natl. Sedimentation Lab., Oxford, MS 38655), and Brian Barkdoll (Univ. of Mississippi, University, MS 38677)

Automated measurement of suspended sediments is crucial to the study of sediment transport. The short duration, high-intensity flows that are responsible for a large fraction of sediment movement are best observed by continuous monitoring systems. Acoustic systems are ideal for continuous monitoring and add the advantage of nonintrusively measuring through the water column. In the present work, a single-frequency acoustic system for measuring suspended-sediment concentration in fluvial environments is described. The equipment and procedures used in developing hardware and software for the acoustic technique in two sets of laboratory flume experiments will be presented. Both implicit and explicit methods were used to convert backscatter data into sediment concentrations. The implicit method yielded 39% error and the explicit, 20%.

Plenary Session and Awards Ceremony

Presentation of Certificates and Awards by Acoustical Society of America

Richard Stern, President

Presentation of Certificates to New Fellows

Richard H. Campbell
Laurel H. Carney
Bruce D. Cornuelle
Peter H. Dahl
George E. Ioup
Leon M. Keer
Hugh J. McDermott
Colette M. McKay
William C. Moss
Philip A. Nelson
Marshall H. Orr
Jack E. Randorff
Sean F. Wu
Ning Xiang
George Zweig

Presentation of Awards

Science Writing Award in Acoustics for Journalists to Bennett Daviss for his article *Snap, Crackle and Pop* published in *New Scientist*, 21 July 2001

Science Writing Award for Professionals in Acoustics to Sharem Vaezy for his article (coauthored with Roy W. Martin and Lawrence A. Crum) *Acoustic Surgery* in *Physics World*, August 2002

Pioneers of Underwater Acoustics Medal to Frederick D. Tappert

Silver Medal in Noise to Louis C. Sutherland

Wallace Clement Sabine Medal to Alfred C. C. Warnock

Honorary Fellowship to Michael Longuet-Higgins

Presentation of Awards by Mexican Institute of Acoustics

Sergio Beristain, President

John William Strutt, 3rd Baron of Rayleigh Medal to Leo L. Beranek

John William Strutt, 3rd Baron of Rayleigh Medal to Per V. Brüel

Herman Ludwig Ferdinand von Helmholtz Medal to José Luis Rodríguez García

Presentation of Certificates by Iberoamerican Federation of Acoustics

Samir N. Y. Gerges, President

Session 4aAAa

Architectural Acoustics and Speech Communication: Speech Intelligibility in Buildings and Metrics for the Prediction and Evaluation of Intelligibility

Kenneth P. Roy, Cochair
Innovation Center, Armstrong World Industries, P.O. Box 3511, Lancaster, Pennsylvania 17604

Peter A. Mapp, Cochair
Peter Mapp Associates, 101 London Road, Copford, Colchester CO3 4JZ, United Kingdom

Chair's Introduction—8:00

Invited Papers

8:05

4aAAa1. Is the Speech Transmission Index (STI) a robust measure of sound system speech intelligibility performance? Peter Mapp (Peter Mapp Assoc., 101 London Rd., Copford, Colchester C03 4JZ, UK, Petermapp@btinternet.com)

Although RaSTI is a good indicator of the speech intelligibility capability of auditoria and similar spaces, during the past 2–3 years it has been shown that RaSTI is not a robust predictor of sound system intelligibility performance. Instead, it is now recommended, within both national and international codes and standards, that full STI measurement and analysis be employed. However, new research is reported, that indicates that STI is not as flawless, nor robust as many believe. The paper highlights a number of potential error mechanisms. It is shown that the measurement technique and signal excitation stimulus can have a significant effect on the overall result and accuracy, particularly where DSP-based equipment is employed. It is also shown that in its current state of development, STI is not capable of appropriately accounting for a number of fundamental speech and system attributes, including typical sound system frequency response variations and anomalies. This is particularly shown to be the case when a system is operating under reverberant conditions. Comparisons between actual system measurements and corresponding word score data are reported where errors of up to 50 implications for VA and PA system performance verification will be discussed.

8:25

4aAAa2. Correlation study of predictive and descriptive metrics of speech intelligibility. Abigail Stefaniw, Yasushi Shimizu, and Dana Smith (Prog. in Architectural Acoust., School of Architecture, Rensselaer Polytechnic Inst., Troy, NY 12180)

There exists a wide range of speech-intelligibility metrics, each of which is designed to encapsulate a different aspect of room acoustics that relates to speech intelligibility. This study reviews the different definitions of and correlations between various proposed speech intelligibility measures. Speech Intelligibility metrics can be grouped by two main uses: prediction of designed rooms and description of existing rooms. Two descriptive metrics still under investigation are Ease of Hearing and Acoustical Comfort. These are measured by a simple questionnaire, and their relationships with each other and with significant speech intelligibility metrics are explored. A variety of rooms are modeled and auralized in cooperation with a larger study, including classrooms, lecture halls, and offices. Auralized rooms are used to conveniently provide calculated metrics and cross-talk canceled auralizations for diagnostic and descriptive intelligibility tests. Rooms are modeled in CATT-Acoustic and auralized with a multi-channel speaker array in a hemi-anechoic chamber.

8:45

4aAAa3. The social effects of poor classroom acoustics on students and The District of Columbia Public Schools demonstration of support through mandating the ANSI Classroom Acoustics standard. Donna Ellis (The District of Columbia Public Schools, 1709 3rd St. NE, Washington, DC 20002)

The effects that poor acoustics have on students extend beyond the classroom. This paper is to discuss the immediate and long-term results that inadequate acoustical design in the educational setting has on academic and social development and how the District of Columbia Public Schools (DCPS) are contributing to the classroom acoustic movement. DCPS is taking a pro-active stance in educational acoustics by mandating the ANSI Draft S12.60-200X classroom acoustic standard in the transformation of ten schools

a year for the next ten to fifteen years. Synthesizing the ANSI S12 standard with the DCPS Design Guidelines describes explanation of how to design for appropriate acoustics in all core-learning spaces. Examples of the existing conditions of the facilities and acoustical remediation for new and historical preservation projects will be demonstrated. In addition, experience will be shared on the International Building Code Council hearings for classroom acoustics.

Contributed Papers

9:05

4aAAa4. Assessment of language impact to speech privacy in closed offices. Yong Ma Ma, Daryl J. Caswell (Dept. of Mech. & Manufacturing Eng., Univ. of Calgary, Calgary, AB T2N 1N4, Canada), Liming Dai (Univ. of Regina, Regina, SK S4S 0A2, Canada), and Jim T. Goodchild (Smed Intl., Calgary, AB T2C 4T5, Canada)

Speech privacy is the opposite concept of speech intelligibility and can be assessed by the predictors of speech intelligibility. Based on the existing standards and the research to date, most objective assessments for speech privacy and speech intelligibility, such as articulation index (AI) or speech intelligibility index (SII), speech transmission index (STI), and sound early-to-late ratio (C50), are evaluated by the subjective measurements. However, these subject measurements are based on the studies of English or the other Western languages. The language impact to speech privacy has been overseen. It is therefore necessary to study the impact of different languages and accents in multiculturalism environments to speech privacy. In this study, subjective measurements were conducted in closed office environments by using English and a tonal language, Mandarin. Detailed investigations on the language impact to speech privacy were carried out with the two languages. The results of this study reveal the significant evaluation variations in speech privacy when different languages are used. The subjective measurement results obtained in this study were also compared with the objective measurement employing articulation indices.

9:20

4aAAa5. An analysis and retrofit of the acoustics at Image Creators Health and Beauty Salon. Donna Ellis (521 Manor Rd., Severna Park, MD 21146)

This paper discusses the analysis and retrofit of the acoustics in a high-volume beauty salon in Severna Park, MD. The major issues in what was designed to be a serene environment are reverberation times of 1–1.68 s in the mid- to upper-frequency range. Employee and customer complaints include heightened stress, vocal strain, headaches, and poor intelligibility. Existing analysis and acoustical retrofit solutions will be demonstrated.

9:35

4aAAa6. A speech perception test for children in classrooms. Sergio Feijóo, Santiago Fernández, and José Manuel Álvarez (Dept. de Fisica Aplicada, Fac. de Fisica, Univ. de Santiago, 15782 Santiago de Compostela, Spain)

The combined effects of excessive ambient noise and reverberation in classrooms interfere with speech recognition and tend to degrade the learning process of young children. This paper reports a detailed analysis of a speech recognition test carried out with two different children populations of ages 8–9 and 10–11. Unlike English, Spanish has few minimal pairs to be used for phoneme recognition in a closed set manner. The test consisted in a series of two-syllable nonsense words formed by the combination of all possible syllables in Spanish. The test was administered to the children as a dictation task in which they had to write down the words spoken by their female teacher. The test was administered in two blocks on different days, and later repeated to analyze its consistency. The rationale for this procedure was (a) the test should reproduce normal academic situations, (b) all phonological and lexical context effects should be avoided, (c) errors in both words and phonemes should be scored to unveil any possible acoustic base for them. Although word recognition scores were similar among age groups and repetitions, phoneme errors showed high variability questioning the validity of such a test for classroom assessment.

9:50

4aAAa7. Speech intelligibility in the community mosques of Dhaka City. Sheikh Muhammad Najmul Imam (Dept. of Architecture, Bangladesh Univ. of Eng. and Technol. (BUET), Dhaka 1000, Bangladesh, najmul@buet.ac.bd)

A mosque facilitates a Muslim community through different religious activities like congregational prayers, recitation and theological education. Speech in a mosque usually generates through bare voice though sound amplification system is also applied. Since no musical instrument is used in any liturgy, a mosque involves only speech acoustics. The community mosques of Dhaka city, the densely populated capital of Bangladesh, are usually designed and constructed by common people inspired from religious virtues. Seeking consultancy for acoustical design is almost never done. As an obvious consequence, there is a common crisis of speech intelligibility in different mosques, except those saved for their smaller volume and other parameters generated by chance. In a very few cases, a trial and error method is applied to solve the problem. But in most of the cases, the problem remains unsolved, putting the devotees in endless sufferings. This paper identifies the type and magnitudes of the prevailing crisis of speech intelligibility of these community mosques through instrumental measurements and questionnaire survey. This paper is also intended to establish certain research rationale and hypothesis for further research, which will propose certain parameters in acoustical design for mosques of Dhaka city in particular and of Bangladesh in general.

Session 4aAAb

Architectural Acoustics, Musical Acoustics and Physical Acoustics: Ancient Acoustics II Vases and Theater Spaces

Per V. Brüel, Cochair
Brüel Acoustics, Gl Holtevej 97, Holte DK-2840, Denmark

Sergio Beristain, Cochair
IMA ESIME, P.O. Box 75805, 07300, D.F. Mexico

Invited Papers

10:30

4aAAb1. Models of ancient sound vases. Per V. Bruel (Bruel Acoust., GI Holtevej 97, Holte DK-2840, Denmark, pvb@bruela.dk)

Models were made of vases described by Vitruvius in Rome in about the year 70 A.D. and of sound vases (lydpotter) placed in Danish churches from 1100–1300 A.D. Measurements of vase's resonant frequencies and damping (reradiation) verified that the model vases obeyed expected physical rules. It was concluded that the excellent acoustical quality of many ancient Greek and Roman theaters cannot be ascribed to the vases placed under their seats. This study also found that sound vases placed in Nordic churches could not have shortened the reverberation time because there are far too few of them. Moreover, they could not have covered a broad frequency range. It remains a mystery why vases were installed under the seats of ancient Greek theaters and why, 1000 years later, Danes placed vases in their churches.

10:50

4aAAb2. Acoustics of Chinese traditional theatres. J. Q. Wang (Inst. of Acoust., Tongji Univ., Shanghai 200092, PROC)

Chinese traditional theatre is unique in the world as an architectural form. The Chinese opera evolved into maturity as early as the Song and Yuan Dynasties, 11th–14th centuries, and Chinese theatrical buildings developed accordingly. As the Chinese opera plays on the principle of imaginary actions, no realistic stage settings are required. But Chinese audiences have placed great demands on vocal performance since ancient times. Pavilion stages that are small in area, open on three sides, and thrusting into the audience area are commonly found in traditional theatres, both the courtyard type and auditorium type. The pavilion stage is backed with a wall and a low ceiling (flat or domed). The stage functions as a reflecting shell, which enhances the sound to the audience and provides self-support to the singer. Numerous theatres of this kind exist and function in good condition to the present time. Acoustical measurements show that the sound strength and clarity in audience areas is satisfactory in moderate size courtyard traditional theatres. [Work supported by NSF.]

11:10

4aAAb3. An acoustical performance space in ancient India: The Rani Gumpha. C. Thomas Ault and Umashankar Manthravadi (Dept. of Theatre and Dance, Indiana Univ. of Pennsylvania, Waller Hall, Indiana, PA 15705)

The Rani Gumpha, or Queen's Cavern, was built by artist–king of Kalinga, Kharavela (*ca*. 200–100 B.C.). It is a rock cut structure, carved into Udayagiri hill. As in ancient Greek and Roman theaters, the entire performance space of the Rani Gumpa is backed by a decorated facade, and it is remarkably similar to Greek theaters of the Hellenistic period, having both an upper and lower level for playing. There are acoustical chambers behind each level as well as on either side, and a special ''cantor's chamber'' stage left on the lower level. The effect on the voice is astonishing. This is a rock cut acoustical installation analogous to that described by Vitruvius in Book V, Chaps. 5 and 8, of his de Architectura, where he speaks of vessels placed in Greek and Roman theaters for the same purpose. We have created a computerized model of the Ranim Gumpha, using CATT Acoustic. We have taken acoustic measurements of the site, using Aurora Sofware package. Our results indicate that the Rani Gumpha is an acoustical performance site, sharing characteristics of the classical Greek and Roman theaters of approximately the same period.

Contributed Papers

11:30

4aAAb4. Ethnographic model of acoustic use of space in the southern Andes for an archaeo-musicological investigation. Jose Perez de Arce (Museo Chileno de Arte Precolombino, Bandera 361, Santiago, Chile, jperezdearce@museoprecolombino.cl)

Studies of ritual celebrations in central Chile conducted in the past 15 years show that the spatial component of sound is a crucial component of the whole. The sonic compositions of these rituals generate complex musical structures that the author has termed ''multi-orchestral polyphonies.'' Their origins have been documented from archaeological remains in a vast region of southern Andes (southern Peru, Bolivia, northern Argentina, north-central Chile). It consists of a combination of dance, space walk-through, spatial extension, multiple movements between listener and orchestra, and multiple relations between ritual and ambient sounds. The characteristics of these observables reveal a complex schematic relation between space and sound. This schema can be used as a valid hypothesis for the study of pre-Hispanic uses of acoustic ritual space. The acoustic features observed in this study are common in Andean ritual and, to some extent are seen in Mesoamerica as well.

11:45

4aAAb5. General considerations regarding the importance of a study of the acoustic conditions of churches with historic and artistic value in Mexico. The case of baroque architecture. Fausto Rodriguez Manzo and Eduardo Kotasek Gonzalez (Div. de Ciencias y Artes para el Diseno, Universidad Autonoma Metropolitana–Azcapotzalco, Mexico, D.F., Mexico, rfme@correo.azc.uam.mx)

This paper deals with the problem of old churches in Mexico with historic and artistic value. Some churches in the baroque period in Mexico are examples of a kind of building used as concert and conference halls. However there is a lack of acoustic knowledge of these kinds of architectural spaces in Mexico. This knowledge will also augment that of the historic, artistic and architectural point of view, and be a foundation for high-quality renovation and preservation projects for this type of building. The particular characteristics of some of the most significant churches of this period are presented as well as a first architectural acoustics review of them.

THURSDAY MORNING, 5 DECEMBER 2002 CORAL KINGDOM 2 AND 3, 8:30 TO 11:35 A.M.

Session 4aAB

Animal Bioacoustics: Topics in Animal Bioacoustics

Whitlow W. L. Au, Chair
Hawaii Institute of Marine Biology, P.O. Box 1106, Kailua, Hawaii 96734

Contributed Papers

8:30

4aAB1. A general spike coding model for natural echoes. Rolf Müller (Biosonar Lab, Dept of. Animal Physiol., Tübingen Univ., Morgenstelle 28, D-72076 Tübingen, Germany)

Bats navigating in natural environments face a challenging task in identifying landmarks (e.g., vegetation) that are composed of many reflecting facets and hence generate random echo waveforms. The biological substrate for these tasks is a spike code. Informative code features were identified by a computational approach which encoded a natural stimulus ensemble (84 800 echoes from four deciduous tree species) using a parsimonious model of spike generation. Interspike intervals spanning more than one cycle at the auditory bandpass filter's center frequency proved convenient classification substrates because of the following properties: (a) almost certain presence in each echo representation; (b) simple probabilistic description (uncorrelated); (c) undemanding constraints on firing threshold locations and spike time resolution. An m-ary sequential probability ratio test using a first-order characterization of these intervals (number plus sample mean of duration and threshold location) demonstrated good classification performance (0.1% to 1% error probability at <10 expected echoes) as well as remarkable robustness against changes in the spike-model parameters. Distant-cycle intervals are therefore hypothesized to be a primal sketch building block analogous to visual edges, with the notable difference that they do not delineate deterministic shapes [Work supported by DFG, SFB 550 B6.]

8:45

4aAB2. Narrowband *Phocoena*-like clicks produced by a delphinid species. Marc O. Lammers and Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, lammers@hawaii.edu)

A distinction is traditionally made between the echolocation clicks of the whistling members of the family *Delphinidae* and those of nonwhistling smaller odontocetes, which include the family *Phocoenidae* and members of genus *Cephalorhynchus* and *Kogia*. The larger, whistling delphinids typically emit clicks that are short (40–70 μs) and broadband (30–60 kHz) with peak frequencies ranging from 40 to 130 kHz. Clicks of the smaller nonwhistling odontocetes are generally longer (>100 μs), more narrowband (15–25 kHz) with peak frequencies between 90 and 160 kHz. Here we report findings of long (~ 135 μs), narrowband (10–25 kHz) clicks produced by a whistling species of delphinid, the Hawaiian spinner dolphin (*Stenella longirostris*). These clicks typically have peak frequencies between 90 and 100 kHz and predominantly occur in a doublet pattern with clicks consistently spaced 13 ms apart within a doublet and 23 ms between doublets. The vast majority of free-ranging spinner dolphin clicks conform to the more traditional delphinid form. Narrowband clicks appear to be produced rarely, having been observed on only 4 out of 25 recording sessions. It is presently not clear whether these clicks represent echo location or communication signals.

9:00

4aAB3. Features of echoes bottlenose dolphins use to perceive object properties. Caroline M. DeLong, Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, delong@hawaii.edu), and Heidi E. Harley (New College of Florida, Sarasota, FL 34243)

Echolocating bottlenose dolphins can discriminate among objects that vary in size, shape, and structure. Our purpose in this project is to investigate the features of echoes dolphins may use to determine each of these object properties (e.g., target strength, number, and position of highlights). A dolphin performed a cross-modal matching task where he was presented with an object in one modality (e.g., vision) and then asked to choose the same object from among a group of three objects using another modality (e.g., echolocation). The dolphin was presented with object sets in which the objects within the set varied along one feature (size or shape or material or texture). These objects were later ensonified with dolphin-like clicks and the object echoes were collected with a binaural (two hydrophone) system. To mimic the dolphin's ability to scan across objects, echoes were captured as the objects were rotated. Differences in echoic features within object sets were examined along with the dolphin's error. In addition, an artificial neural network (ANN) was used to classify objects using the object echoes. The ANN was utilized to explore the importance of spectral versus time domain features, and single versus multiple object orientations in object classification.

9:15

4aAB4. Target strengths of two gillnets and predictions of biosonar detection ranges. T. Aran Mooney, Paul E. Nachtigall, and Whitlow W. L. Au (Hawaii Inst. of Marine Biol., Univ. of Hawaii, P.O. Box 1106, Kailua, HI 96734, mooneyt@hawaii.edu)

Small cetaceans are often caught as bycatch in gillnet fisheries. One potential method that has been shown to significantly reduce harbor porpoise incidental take is the use of barium sulfate enhanced gillnets. One hypothesis states barium sulfate nets reduce bycatch because they are acoustically more reflective than monofilament nets, thus easier to detect by echolocating odontocetes. This premise remains untested. To determine acoustic reflectivity, this study compared the target strength of a barium sulfate gillnet with a comparable monofilament gillnet at 0, 10, 20, 30, and 40. At normal incidence (0) the monofilament target strength measured slightly greater than the barium sulfate net and biosonar detection ranges are estimated to be greater for the monofilament. As the angle of incidence increased from normal to 20, the barium sulfate net echo strength remained the same while the nylon net echo strength decreased significantly at 10. Biosonar detection ranges are estimated to be greater for the barium sulfate net at increased angles. Echolocating odontocetes may utilize acoustic reflections of both nets but due to the weak target strength of both nets, other mechanisms may also mitigate reduced odontocete bycatch. Current psychoacoustic research is investigating the net detection ranges of captive bottlenose dolphins.

9:30

4aAB5. Chiroptera-inspired robotic cephaloid (CIRCE): A next generation biomimetic sonar head. Rolf Müller, Andrew J. Mubeezi-Magoola (Biosonar Lab, Dept. of Animal Physiol., Tübingen Univ., Morgenstelle 28, D-72076 Tübingen, Germany), Herbert Peremans (Antwerp Univ., B-2000 Antwerpen, Belgium), John Hallam (Univ. of Edinburgh, Edinburgh EH1 2QL, Scotland), Simon Jones, James Flint (Loughborough Univ., Loughborough LE11 3TU, UK), Dominiek Reynaerts, Herman Bruyninckx (Katholieke Universiteit Leuven, B-3001 Leuven, Belgium), Reinhard Lerch, and Alexander Streicher (Friedrich Alexander Univ., D-91052 Erlangen, Germany)

The dynamic interplay between signal processing in the acoustic and neural domain renders the biosonar system of bats, a prime example of intelligent, integrated systems in biology. Previous attempts at technical reproductions of this system have studied a few functional aspects but did not produce convincing replicas of the integrated system. In particular, none of them could provide for beamforming pinnae, ear mobility, signal bandwidth and an adequate number of elements in the primary signal representation at the scale of the biological paragon. The CIRCE (http://www.circe-project.org) project brings together a multidisciplinary consortium with all competences necessary for addressing these aspects in conjunction. Well-integrated solutions to the following technological key challenges are therefore sought: (a) Evolution of pinna shapes to optimize performance in natural biosonar tasks. (b) Adaptation of state-of-the-art ultrasonic transducer technology for matching bats in terms of bandwidth and sound pressure level/sensitivity. (c) Actuation of mobile pinnae (including the possibility of nonrigid transformations) within the scale constraints of a bat's head. (d) Custom-designed DSP hardware providing a qualitative and quantitative reproduction of the neural signal representation at the level of the auditory nerve. [Work supported by the European Commission, LPS Initiative.]

9:45

4aAB6. Target range discrimination by the bottlenose dolphin. L. Gennadi Zaslavskiy (Univ. Authority for Appl. Res., Tel-Aviv, P.O. Box 3359, Hadera 38100, Israel)

In order to measure range accuracy of the dolphin's sonar, a range discrimination experiment was conducted. To prevent simultaneous exposure of both targets to each dolphin's click, targets were submerged under water at different ranges in sequence with a 0.3–0.5-sec delay. Otherwise, if presented simultaneously, the targets could be distinguished by difference in the echo arrival times, without measuring the target ranges. Couples of identical spheres as well as targets of a different size and shape were tested. Nevertheless, it proved to be impossible to make the dolphin's measure the distance from themselves to targets. Both dolphins solved the task by measuring a time interval between the target echo and an echo from a reference object, which could be the pool walls or even a net. There was no consistent change in the range accuracy with the change of the target range whereas it regularly increased when the targets were moved closer to the reference object. The dolphins did not keep the distance from the targets as short as possible. Instead, by backing off from the targets they tried to position the target echoes as close as possible to the reference echo. It appeared that the dolphins were incapable of measuring a target range.

10:00–10:20 Break

10:20

4aAB7. Click train encoding in a song bird: A psychometric function showing two mechanisms of pitch perception. Jeffrey Cynx (Dept. of Psych., Vassar College, 124 Raymond Ave., Poughkeepsie, NY 12604)

In humans, the pitch of click trains composed of pulses with the same and alternating polarity can vary depending on whether the pulse sequences are below 100 Hz or above 200 Hz. Below 100 Hz, the pitch is equal to the pulse rate regardless of polarity. Above 200 Hz, the pitch is determined by the fundamental frequency and the pattern of click polarity [J. L. Flanagan and N. Guttman, J. Acoust. Soc. Am. **32**, 1308–1319 (1960)]. European starlings (*Sturnus vulgaris*), a species of songbird, were trained on an operant procedure with bipolar stimuli on a two tone discrimination, then tested with unipolar stimuli. We then built a psychometric function, especially covering the ambiguous region between 100 and 300 Hz, and the upper frequency limits.

10:35

4aAB8. Hearing is an active process in mosquitoes. Daniel Robert and Martin C. Goepfert (Biological Sci., Woodland Rd., Univ. of Bristol, Bristol BS8 1UG, UK, d.robert@bristol.ac.uk)

Auditory sensitivity and frequency selectivity are enhanced by active mechanics. Known from vertebrates, active audition is established for an invertebrate. The auditory mechanics of antennal sound receivers in mosquitoes exhibits key diagnostic features of active mechanics: (a) physiologically vulnerable sensitivity and tuning. The mechanical sensitivity and tuning sharpness of the receiver decrease after death. (b) Vulnerable nonlinear mechanical response. Nonlinearity is expressed as amplitude-dependent damping: lower stimulus intensities result in larger response gain and sharper resonance peaks. This band-limited nonlinearity disappears post-mortem. (c) Hypoxia sensitivity. In females, transient exposure to CO_2 reversibly reduces tuning sharpness. In males, the effect is opposite: CO_2 sharpens tuning and elicits large vibrations in the absence of acoustic stimulation. Hypoxia-induced vibrations are unrelated to external forces; they are autonomous and self-sustained, and forcibly generated by an internal motor. (d) Autonomous antennal vibration (AV). AV is also induced by the injection of dimethylsufoxide (DMSO). DMSO-induced AV persists (90 min) at large amplitudes (>400 nm). Antennal deflection shapes rule out the involvement of muscles, narrowing down the source of mechanical activity to the auditory sense organ at the antennas base. [Work supported by the Leopoldina Academic Society, the Swiss National Science Foundation, BBSRC and the University of Bristol.]

10:50

4aAB9. Micromechanics of *Drosophila* audition. Martin C. Goepfert and Daniel Robert (Biological Sci., Woodland Rd., Univ. of Bristol, Bristol BS8 1UG, UK, m.gopfert@bristol.ac.uk)

In *Drosophila melanogaster*, hearing is mediated by the antenna, vibrations of which are transduced by Johnston's organ. The receiver's mechanical response exhibits key features of active mechanics. Limit cycle oscillations (LCO) can be measured in untreated animals and after the injection of dimethylsulphoxide. LCO are entrained by acoustic stimuli and reversibly vanish during hypoxia. The mechanical response is nonlinear, exhibiting an increase in resonance frequency with stimulus intensity. This nonlinear stiffness is reversibly affected by hypoxia and disappears post mortem. The origin of mechanical nonlinearity and LCO was traced in auditory mutants no-mechanoreceptor-potential A. NompA protein links the receptor neurone dendrites to the extracellular caps. Mutations in the nompA gene thus mechanically isolate the auditory receptors from antennal structures, resulting in conductive hearing loss. An antennal response in nompA flies lacks LCO and the nonlinear stiffness alteration of controls. Also, LCO and intensity-dependent stiffness depend on functional mechanotransduction channels: mutations in nompC linearize antennal mechanics. Similar effects are measured in touch-insensitive-larva B (tilB) and beethoven (btv) mutants, indicating that mechanical activity in Drosophila ensues from the interplay between nompC channels and receptor cilia, possibly involving ciliary motility. [Work supported by the Leopoldina Academic Society, the Swiss Science Foundation, BBSRC, and the University of Bristol.]

11:05

4aAB10. Acoustic communication by ants. Robert Hickling (Sonometrics, Inc., 8306 Huntington Rd., Huntington Woods, MI 48070, sonomet@aol.com) and Richard L. Brown (Mississippi State Univ., MS 39763)

Effective communication is essential in highly organized ant societies. Currently myrmecologists believe that pheromones play the central role in ant communication and that relatively little is accomplished with sound. However many ant species make sounds by stridulating, i.e., running a scraper over ridges on their abdomen. Hence, it seems likely that ants communicate using a combination of pheromones and sound. Because ants appear insensitive to airborne sound, myrmecologists have concluded that stridulatory signals are transmitted through the substrate. However, this conclusion is untenable for a number of reasons. How then can ants transmit signals through the air when apparently they are insensitive to airborne sound? The answer lies in the fact that ants are much smaller than the wavelength of the sound they generate, so there is a relatively significant near field surrounding a stridulating ant. The near field is characterized by a major increase in the particle velocity of sound, which can be detected by hair-like sensilla on the apical segments of the ants' antennae. By sensing the relative difference in particle sound velocity between the antennae, an ant can detect stridulation signals from a nearby ant, yet appear deaf to extraneous airborne sound.

11:20

4aAB11. Behavioral responses to acoustical stimulation of individuals and colonies of the termite *Reticulitermes virginicus*. JulieAnn M. Chambers and Roger Hasse (Univ. of Mississippi, Natl. Ctr. for Physical Acoust., University, MS 38677)

This study examined the effects of acoustical stimulation on *Reticulitermes virginicus* termites and describes the results of experiments conducted to analyze behavioral responses to various frquencies of sound. Termites were studied for individual response to sound stimulation and for colony response. Individuals were mounted to rods and responses were recorded to sound stimulations and magnetic vibrations. Colonies were established and allowed to acclimate in simulated nesting environments containing fine porous sand and rotten wood shims. A transducer was buried in the sand close to the colony and driven by a function generator using variable frequencies at fixed amplitude. Two types of sound stimuli were applied to each colony: tone bursts (TB) and continuous waves (CW). Behaviors indicated that the termites were agitated by the sound stimulus and termite activity was interpreted as an avoidance response. Once stimuli were removed the termites returned to normal activity near the aquarium walls. Observed responses are being examined for consistency and characterization of behaviors. Further study to be presented will quantify the nature of acoustic transmission in the aquaria as well as to understand the termite responses to separate aspects of the stimuli.

Session 4aBB

Biomedical Ultrasound/Bioresponse to Vibration: Ultrasound-Mediated Drug Delivery and Gene Transfection

Junru Wu, Chair
Department of Physics, University of Vermont, Burlington, Vermont 05405

Invited Papers

8:00

4aBB1. Ultrasound mediated transdermal drug delivery. Ahmet Tezel, Ashley Sens, and Samir Mitragotri (Chemical Eng. Dept., UCSB, Santa Barbara, CA 93106)

Low-frequency sonophoresis (LFS) is a noninvasive method of transdermal drug delivery and diagnostics. In this method, a short application of ultrasound is used to permeabilize skin for a prolonged period of time. During this period, ultrasonically permeabilized skin may be used as a permeable interface to the body. This interface may be used for the extraction of analytes or the delivery of drugs. Our main objective in this study is to understand the mechanisms of skin permeability due to LFS. Although the principal mechanism of LFS is known to be cavitation, little is known about how cavitation does actually enhance skin permeability. We are specifically interested in understanding the interactions of cavitation bubbles with the skin surface. We are also interested in assessing the structural alterations in the skin induced by ultrasound. We hypothesize that ultrasound generates defects that are responsible for the enhanced delivery of large hydrophilic solutes such as proteins (e.g., insulin). In this study we sought to determine the dependence of LFS transport pathways on ultrasound parameters such as application frequency and energy density.

8:20

4aBB2. Transdermal drug delivery by sono-macroporation. Ludwig Weimann (Ultra-Sonic Technologies, LLC, P.O. Box 319, St. Albans, VT 05478, ljw@together.net) and Junru Wu (Univ. of Vermont, Burlington, VT 05405)

A feasibility study of using high amplitude ultrasound to deliver large molecules transdermally was undertaken. Ultrasound (20 kHz) of intensity in the range between 2 to 50 W/cm^2 was used to increase the permeability of skin *in vitro* to poly-l-lysines as models of drug molecules. When 20 kHz, 5% duty cycle ultrasound at the spatial average and pulse average intensity $I_{SAPA}=19$ W/cm^2 was applied for 10 min the skin permeability was calculated to be 0.5 ± 0.2 cm/h and 8.5 ± 4.2 cm/h, respectively, for poly-l-lysine-FITC (51 kDa) and octa-l-lysine-FITC (2.5 kDa). Without application of ultrasound the skin permeability of the above-mentioned molecules was essentially zero. A transdermal flux enhancement occurred during the process reported here was much higher than that due to sonophoresis ($I_{SAPA}<2$ W/cm^2) as reported in literature (Mitragotri *et al.,* 1995). Reported here are experimental results from transdermal flux kinetis, and confocal microscopic cross sectional and optical images suggest the formation of pores in the stratum corneum whose size varies with skin samples and may be in the range of 1–100 μm. The confocal images also suggest the formation of micron-size pathways in epidermis during ultrasound exposure.

8:40

4aBB3. Ultrasound-enhanced gene transfer: Comparison of contrast agents and ultrasound modalities *in vitro* and *in vivo*. Sorin V. Pislaru, Gulati Rajiv, Cristina Pislaru, Randall R. Kinnick, Ripudamanjit Singh, James F. Greenleaf, and Robert D. Simari (Mayo Clinic and Foundation, Rochester, MN 55905)

Ultrasound (US) enhancement of plasmid-based gene transfer is an emerging technique. Our hypothesis was that two contrast agents (Optison and PESDA), and two US exposure modalities (dedicated continuous wave system and diagnostic scanner) may have different effects. Luciferase plasmid with or without contrast agent was added to vascular smooth muscle cells and endothelial cells, followed by US exposure. Luciferase activity was measured 24 h later. US exposure consistently induced higher transfection rates than all controls. PESDA was superior to Optison in both cell lines. *In vitro,* continuous wave and diagnostic US were not significantly different. *In vivo,* Lux and PESDA were injected into skeletal muscles of rats (IM or intra-arterial) followed or not by US exposure. In separate animals, adenovirus encoding for luciferase was injected IM and was not followed by US exposure. Gene transfer efficacy was 8–10 fold higher with US and PESDA than with plasmid alone, but 2 fold lower than with adenovirus. However, as opposed to adenovirus, US-enhanced plasmid gene transfer was highly localized to the injected muscle, with no expression at distal sites. Our results support the hypothesis that contrast agents and exposure modalities are not equivalent with regard to gene transfer efficacy.

9:00

4aBB4. Gene transfection by echo contrast agent microbubbles. Katsuro Tachibana (Dept. of Anatomy, Fukuoka Univ. School of Medicine, 7-45-1 Nanakuma, Jonan Fukuoka, Japan, k-tachi@cis.fukuoka-u.ac.jp)

In vitro and *in vivo* experiments have demonstrated that various echo contrast agent microbubbles can be intentionally ruptured by diagnostic and therapeutic ultrasound. Violent microstreaming are produced during microbubble collapse. Researchers have hypothesized that these microjets or microstreaming could be applied to promote diffusion of drugs into various tissues and lesions. The most exciting application of this method is probably delivery of genes into cells. As various genes are currently under investigation for the

purpose of treating diseases, ultrasound and microbubbles may be used as a modality to promote better outcome for gene therapy. Recent studies have shown that different gases contained within the bubbles greatly influence the degree of gene transfection. Also, the outer layer of the microbubbles can be custom-made for binding to target tissue. Recent advance on this topic will be discussed.

9:20

4aBB5. Bubbles, membranes and molecules: sequence of events from sonication to intracellular delivery. Mark R. Prausnitz, Pavel P. Kamaev, Robyn K. Schlicher, and Vladimir G. Zarnitsyn (School of Chemical Eng., Georgia Tech, Atlanta, GA 30332-0100, mark.prausnitz@che.gatech.edu)

To elucidate mechanisms and control bioeffects for ultrasound-mediated drug and gene delivery, we carried out an experimental study to quantitatively measure the effects of ultrasound and other physical parameters on the sequence of events leading from sonication to drug and gene delivery. Using a Coulter counter, the number and size distribution of bubbles was measured as a function of ultrasound pressure and time, and found to decrease as a function of acoustic energy exposure. The effects of these bubbles on cells were measured by electron and confocal microscopy, which indicated that cavitation created cell membrane defects that could be actively repaired and permitted the intracellular transport of molecules. Using flow cytometry, levels of molecular uptake and cell viability were quantified over a broad range of conditions and correlated with acoustic energy, the ratio of cells-to-cavitation nuclei and the size distribution of bubbles. Finally, levels of gene expression were quantified as a function of acoustic and other parameters and related to cavitation bubble dynamics. By examining each of the steps, rather than only measuring endpoints, we can understand and ultimately control the process by which ultrasound delivers drugs and genes into cells.

9:40–10:00 Break

Contributed Papers

10:00

4aBB6. Shock induced jetting of micron sized bubbles. Claus-Dieter Ohl, Roy Ikink, Detlef Lohse (Dept. of Appl. Phys., WSL, TU Twente, Postbus 217, 7500 AE Enschede, The Netherlands), and Andrea Prosperetti (Johns Hopkins Univ., Baltimore, MD 21218)

Gas bubbles having a radius between 10 μm and 100 μm and rising freely in water when being subjected to a shock front exhibit a liquid jetting phenomenon. The jet points in the direction of the propagating shock wave. A linear relationship between the jet length and the bubble radius is found and a lower bound of the averaged velocity of the liquid jet can be estimated to be between 50 m/s and 300 m/s increasing linearly for larger bubbles. In a later stage the jet breaks up and releases micron sized bubbles. In the course of shock wave mediated cell permeabilization this observation suggests a microinjection mechanism responsible for cell transfection when minute gas bubbles are present and exposed together with cells to shock waves.

10:15

4aBB7. Interaction of ultrasound driven microbubbles and lipid membranes. Philippe Marmottant and Sascha Hilgenfeldt (Appl. Phys., Univ. Twente, P.O. Box 217, 7500AE Enschede, The Netherlands, p.marmottant@tn.utwente.nl)

Micron-sized bubbles show pronounced oscillations when submitted to ultrasound, leading to increased scattering and improved echographic contrast. It has been reported that this excitation can also alter nearby cell membranes [M. Ward, J. Wu, and J. F. Chiu, Ultrasound Med. Biol. **26**, 1169–1175 (2000)], and increase the permeability for drug delivery. To elucidate the mechanisms at work in these sonoporation experiments, we developed a setup that allows for a controlled study of the interaction of single microbubbles with single lipid bilayer vesicles. Substituting vesicles for cell membranes is advantageous because the mechanical properties of vesicles are well-known. Microscopic observations reveal that vesicles near a bubble undergo vivid motion, being periodically accelerated toward and away from the bubble. This ''bouncing'' of vesicles is a vivid motion that we attribute to a streaming flow field set up by the bubble oscillation. Some vesicles undergo dramatic deformations as they follow the flow, reflecting the high shear rates attained. Break-up of vesicles could also be observed.

10:30

4aBB8. Evaluation of polymer solubility and cavitation on ultrasound-polymer synergy. Tyrone M. Porter, Lawrence A. Crum (Ctr. for Industrial and Medical Ultrasound (CIMU), Univ. of Washington, Seattle, WA 98105), Patrick S. Stayton, and Allan S. Hoffman (Univ. of Washington, Seattle, WA 98105)

Poly(propylacrylic acid) is a *p*H-sensitive membrane disruptive polymer designed to release therapeutic molecules from endosomes to the cell cytoplasm before degradation. In mildly acidic environments, the polymer becomes more hydrophobic and less soluble in aqueous medium. Previous research has demonstrated the capacity of polyelectrolytes and therapeutic ultrasound to synergistically disrupt red blood cell membranes, resulting in hemolysis [Mourad (2000)]. In this investigation, the ability of therapeutic ultrasound and PPAA to disrupt cell membranes and deliver fluorescently labeled polymer to the cell cytoplasm was demonstrated. The ultrasound/polymer synergy was dependent upon the solubility of PPAA and acoustic cavitation activity. The solubility of PPAA in phosphate buffered saline was quantified by measuring variations in liquid/gas interfacial tension with polymer concentration and acidity. Reductions of surface tension up to 28% were measured using the Wilhemy plate technique. Acoustic cavitation activity was measured using a passive cavitation detection scheme, quantified, and correlated with the number of fluorescent cells. This study demonstrates the potential of combining therapeutic ultrasound and membrane-disruptive polymers for drug delivery.

10:45

4aBB9. Ultrasound mediated gene transfection. Rene G. Williamson, Robert E. Apfel (Dept. of Mech. Eng., Yale Univ., New Haven, CT 06520-8286, rgw27@pantheon.yale.edu), and Janet L. Brandsma (Yale Univ., New Haven, CT 06520)

Gene therapy is a promising modality for the treatment of a variety of human diseases both inherited and acquired, such as cystic fibrosis and cancer. The lack of an effective, safe method for the delivery of foreign genes into the cells, a process known as transfection, limits this effort. Ultrasound mediated gene transfection is an attractive method for gene delivery since it is a noninvasive technique, does not introduce any viral

particles into the host and can offer very good temporal and spatial control. Previous investigators have shown that sonication increases transfection efficiency with and without ultrasound contrast agents. The mechanism is believed to be via a cavitation process where collapsing bubble nuclei permeabilize the cell membrane leading to increased DNA transfer. Our research is focused on the use of pulsed wave high frequency focused ultrasound to transfect DNA into mammalian cells *in vitro* and *in vivo*. Results from further *in vitro* experiments with a 2.5 MHz transducer and Optison as microbubble nuclei will be presented, as well as a discussion of the design of future *in vivo* experiments.

11:00

4aBB10. Quantification and optimization of transfection efficiency and cell viability by acoustic cavitation. Vladimir G. Zarnitsyn and Mark R. Prausnitz (School of Chemical Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0100, mark.prausnitz@che.gatech.edu)

Previous studies have shown that ultrasound-mediated cavitation can introduce large numbers of biological macromolecules, such as DNA, into cells. However, optimization of ultrasound conditions is complicated because elevated levels of cavitation increase uptake of molecules, but also decrease cell viability. The goal of this study was to optimize DNA transfection by identifying acoustic and physical conditions that promote DNA uptake without excessive viability loss. To achieve this, suspensions of DU145 prostate cancer cells were sonicated at 24 and 500 kHz in the presence of plasmid encoding green fluorescent protein (GFP) over a range of pressures, exposure times, cell concentrations, media composition, and nucleation site (Optison) concentration. Expression of GFP and cell viability were independently measured by flow cytometry one day after sonication. We found that above a threshold value, transfection increased, went through a maximum and then decreased with increasing pressure, exposure time, and energy exposure. Cell concentration, media composition, and nucleation site concentration also affected transfection levels and were adjusted to optimize DNA expression. Although exact conditions identified in this study may not be optimal in other biological scenarios, we believe the trends observed here can broadly guide optimization of transfection.

11:15

4aBB11. Characterization of ultrasound-mediated cellular effects. Robyn K. Schlicher, Monali Desai, Mark R. Prausnitz (Wallace H. Coulter Dept. of Biomed. Eng., Georgia Inst. of Technol., Atlanta, GA 30332), and Robert P. Apkarian (Emory Univ., Atlanta, GA)

Ultrasound can temporarily make cells and tissue more permeable, an effect that could be used for enhanced and targeted drug delivery. Increased permeability is believed to involve creation of transient disruptions in cell membranes. This study seeks to characterize these disruptions and the mechanisms by which they are created, reseal and permit intracellular transport. To achieve this, DU145 prostate cancer cells were exposed to 24 kHz ultrasound with 0.1 s pulse length and 10% duty cycle for 2 s total exposure at pressures from 0.36 to 0.89 MPa. Disruptions were estimated to be at least 50 nm in diameter with lifetimes of 1–2 min using a range of fluorescent molecules with known molecular radii studied using flow cytometry. Cell morphological effects were examined using scanning electron, transmission electron, and laser scanning confocal microscopies after rapid fixation (within seconds after exposure). Images indicate that cell death from ultrasound exposure occurs due to a combination of apoptosis, necrosis and mechanical fragmentation and uptake may occur through physical disruptions in cell membrane structure. Using red blood cell ghosts and ATP-depleted prostate cancer cells, it was found that molecular uptake into viable cells requires active cellular processes which infers that cell recovery is an energy-dependent process.

11:30

4aBB12. Measurement of cavitation bubble lifetime and size distribution and its correlation with bioeffects. Pavel Kamaev and Mark Prausnitz (Dept. of Chemical Eng., Georgia Inst. of Technol., 315 Ferst Dr., Atlanta, GA 30332)

Ultrasound-induced disruption of cell membranes can be used to deliver drugs and genes into cells. Because ultrasound bioeffects are governed by cavitation, measurement and control of cavitation bubble lifetime and size distribution should afford greater control over ultrasound's effects on cells. In this study, 500 kHz ultrasound was focused onto suspensions of albumin-stabilized gas bubbles (Optison) either with or without DU145 prostate cancer cells bathed in calcein. Bubble size and lifetime were determined by Coulter counting, while uptake of calcein and cell viability (i.e., bioeffects) were quantified by flow cytometry. We found that the lifetime and size of Optison bubbles decreased with increasing acoustic energy exposure. Moreover, bioeffects were shown to correlate well with disappearance of bubbles. For example, cell viability remained above 90% until approximately 75% of bubbles were destroyed; viability then dropped dramatically as more bubbles disappeared. Additionally, Optison solutions presonicated to destroy all detectable bubbles also caused significant bioeffects. These observations suggest that bioeffects were caused by the cavitation dynamics of free, short-lived, and/or very small daughter bubbles liberated from albumin-stabilized Optison parent bubbles. Regulation of bubble size distribution, possibly by presonication of Optison solutions, could provide a means to optimize high uptake and cell viability.

11:45

4aBB13. Nonlinear behavior of encapsulated bubbles relevant to ultrasound mediated gene transfection applications. Junru Wu, Jason Pepe, and William Dewitt III (Dept. of Phys., Univ. of Vermont, Burlington, VT 05405)

Ultrasound contrast agents (UCA) have been widely used therapeutic applications in medical research and clinical practice. Recently, UCA have been shown to be effective *in vitro* in ultrasound mediated sonoporation and gene transfection and treating vascular thrombosis. The possible primary reasons proposed for the therapeutic applications of UCA are acoustic cavitation [W. J. Greenleaf *et al.*, Ultrasound Med. Biol. **24**, 587–595 (1998)] and acoustic microstreaming surrounding acoustically driven UCA [J. Wu, Ultrasound Med. Biol. **28**, 125–129 (2002)]. Our study has shown that low frequency (0.6 MHz) is more effective than high frequency (2 MHz) in effects related to the acoustic cavitation, and the opposite is true for the microstreaming due to the nonlinear oscillation behavior of encapsulated bubbles of UCA.

Session 4aEA

Engineering Acoustics: Metrology Standards and Calibration on Acoustics, Ultrasound and Vibrations

George S. K. Wong, Cochair
Institute for National Measurement Standards, National Research Council, 1500 Montreal Road, Ottawa, Ontario K1A 0R6, Canada

Jose Salvador Echeverria, Cochair
Centro Nacional de Metrología, km 4.5, Carretera a los cues el Marques, Queretaro 76900, Mexico

Marc Antonio Nabucode Arauj, Cochair
Centro Nacional de Metrología, km 4.5, Carretera a los cues el Marques, Queretaro 76900, Mexico

Chair's Introduction—7:55

Invited Papers

8:00

4aEA1. Microphone interlaboratory comparison in the Americas. George S. K. Wong and Lixue Wu (Inst. for Natl. Measurement Standards, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

The final results of a Sistema Interamericano de Metrologia (SIM) interlaboratory comparison on microphone calibration are presented. Initially the comparison involved NORAMET countries: USA, Canada, and Mexico. Later, the comparison was extended to include Argentina and Brazil, resulting in a SIM AUV.A-K1 microphone interlaboratory comparison. The National Metrology Institutes (NMIs) of the five American countries that participated were the Institute for National Measurement Standards (INMS—Canada), National Institute of Standards and Technology (NIST—USA), Centro Nacional de Metrología (CENAM—Mexico), Instituto Nacional de Metrologia, Normalização e Qualidade Industrial (INMETRO—Brazil) and Unidad Técnica Acústica, (INTI—Argentina). INMS, Canada was the pilot laboratory that provided the data for the final report. The maximum rms deviation for the two LS1P laboratory standard microphones measured by the above participants is 0.037 dB that may be considered as the key comparison reference value.

8:15

4aEA2. Issues concerning international comparison of free-field calibrations of acoustical standards. Victor Nedzelnitsky (Natl. Inst. of Standards and Technol. [NIST], 100 Bureau Dr., Stop 8221, Sound Bldg. [233], Rm. A147, Gaithersburg, MD 20899-8221, Victor.Nedzelnitsky@nist.gov)

Primary free-field calibrations of laboratory standard microphones by the reciprocity method establish these microphones as reference standard devices for calibrating working standard microphones, other measuring microphones, and practical instruments such as sound level meters and personal sound exposure meters (noise dosimeters). These primary, secondary, and other calibrations are indispensable to the support of regulatory requirements, standards, and product characterization and quality control procedures important for industry, commerce, health, and safety. International Electrotechnical Commission (IEC) Technical Committee 29 Electroacoustics produces international documentary standards, including standards for primary and secondary free-field calibration and measurement procedures and their critically important application to practical instruments. This paper addresses some issues concerning calibrations, standards activities, and the international key comparison of primary free-field calibrations of IEC-type LS2 laboratory standard microphones that is being planned by the Consultative Committee for Acoustics, Ultrasound, and Vibration (CCAUV) of the International Committee for Weights and Measures (CIPM). This comparison will include free-field calibrations by the reciprocity method at participating major national metrology laboratories throughout the world.

8:30

4aEA3. A BIPM/CIPM key comparison on microphone calibration—defining the state of the art. Richard G. Barham (Acoust. Metrology Group, Natl. Physical Lab., Teddington, Middlesex TW11 0LW, UK, richard.barham@npl.co.uk)

The formation of the Consultative Committee on Acoustics, Ultrasound and Vibration (CCAUV) now gives acoustical quantities the same status as more established metrics like mass or voltage. The principle work of CCAUV is to establish degrees of equivalence between member states of the Metre Convention. This is achieved through key comparisons and subsequent regional comparison. For sound-in-air, the subject of this key comparison is most appropriately the calibration of laboratory standard microphones. The first such project considered the pressure calibration of IEC type LS1P microphones in the frequency range from 63 Hz to 8 kHz. Twelve national laboratories took part in the key comparison, piloted by the UK's National Physical Laboratory. The project has now been completed. The results for the measured pressure sensitivity level have a standard deviation of around 0.02 dB at frequencies below 1 kHz, rising to a maximum of 0.04 dB at 8 kHz. The mean of these results normalized to 0 dB at each frequency is considered as the key comparison reference value (KCRV) and the standard deviation provides an estimate of its standard uncertainty. The KCRV then defines the datum which enables the performance of all laboratories to be related.

8:45

4aEA4. Sound power measurement standards at laboratory and field environments. Lucìa Taibo, Jorge Riganti, and Waldemar Dittmar (U.T. Acùstica, Instituto Nacional de Tecnologìa Industrial, INTI, c.c. 157, 1650 Prov. Buenos Aires, Argentina)

Sound power level measurements, SWLs, of equipment and machinery through direct or comparison methods at laboratory or field conditions, are of increasing importance in terms of regulations and industrial competitiveness. SWLs constitute the input data when modeling the real world, in complex industrial premises for example, as a means for diagnosis and SPL predictions. SWLs have a major technical and economical impact, trying to keep the measurement uncertainties as low as possible, while balancing the testing complexity, time consumption, and costs. In the present work, a description about the calibration of a Reference Sound Source, RSS, B&K 4204 [ISO 6926] is made. The testing survey of normal equipment emitting steady noise and the RSS at different acoustical environments, that range from an hemi-anechoic room to a typical industrial facility is also described. Some measuring methods such as ISO 3745, ISO 3744, ISO 9614-2 are compared for the different sources and environments, yielding some discrepancies between standards as is extensively referred to in literature. A proposal based on a new multiple height microphone array is encouraged at laboratory conditions. Regarding field tests, sound intensity measurements by scanning is extensively adopted for practical reasons, with acceptable results according to our experience.

9:00

4aEA5. The current status of measurement standards for acoustics and vibration at INMETRO. Gustavo P. Ripper and Walter E. Hoffmann (INMETRO, DIAVI, Av. N. S. das Graças, 50, Xerém-D. Caxias, Rio de Janeiro 25250-020, Brazil, lavib@inmetro.gov.br)

The Division of Acoustics and Vibration of INMETRO (DIAVI) establishes, validates and maintains the Brazilian national measurement standards used for the realization of the units of physical quantities related to the field of acoustics and vibration. The basic vibration quantity realized by DIAVI is translational acceleration, from which the other motion quantities, i.e., velocity and displacement can be derived. Acoustics physical quantities include sound pressure and sound power. The national measurement standards comprise both absolute (primary) and comparison (secondary) calibration setups, which are used to cover the wide range of calibration and testing services demanded by the Brazilian society, trade, industry, and science. This paper outlines the current status of these measurement standards and presents the measurement capabilities of INMETRO in the field of acoustics and vibration.

9:15

4aEA6. The influence of standards on acoustic transducer design and calibration. Gunnar Rasmussen (G.R.A.S. Sound & Vibration ApS, Skelstedet 10B, 2950 Vedbaek, Denmark, gr@gras.dk)

The Western Electric 640AA microphone set the standard for measurement microphones and influenced their development for many decades. It has made its mark on $1-\frac{1}{2}-\frac{1}{4}-\frac{1}{8}$ in. microphones. It has also had an influence on calibration techniques and calibrators. Developments in microphones—calibration and standards will be discussed. The microphone is an important part of a sound level meter and of the complete measurement setup. In particular, measurements in the outdoor environment are greatly influenced by the requirements set by sound level meter standards.

9:30

4aEA7. Validity of acoustic reference standards. E. Frederiksen (Bruel & Kjaer, Skodsborgvej 307, 2850 Naerum, Denmark and Danish Primary Lab. of Acoust.)

Today, acoustic calibrations and associated uncertainties are hot topics. Countries all over the world establish new calibration facilities and coordinate their efforts to make sure that they can calibrate correctly and have their results internationally accepted. This work goes on within accreditation, between National Calibration Institutes, and within standardization. The working groups of the technical committee IEC/TC29, which today make essentially all acoustic instrument and calibration standards, thus expand existing standards with detailed test procedures and related uncertainty requirements. The new requirements are partly associated with the primary pressure and free-field calibration of Reference Standards performed by National Calibration Institutes, and partly with the testing of instruments in Calibration Service Centers, which are the main users of the references. This paper deals with calibration of Reference Standards, which generally needs to be analyzed and elaborated in connection with the ongoing international work. Thousands of microphone and pistonphone calibration results of the Danish Primary Laboratory of Acoustics have been analyzed. Resulting data including uncertainty, reproducibility, and repeatability will be presented. The data are based on calibrations performed over several decades and covering the range from 20 Hz to 25 kHz.

Contributed Papers

9:45

4aEA8. Uncertainties related to fast determination of free-field response for sound level meters. Johan Gramtorp (Bruel & Kjaer, Skodsborgvej 307, 2850 Naerum, Denmark)

As a part of the ongoing work on the new standard IEC 61672 Electroacoustics–Sound level meters, the TC29 WG4 working group is writing the specifications for the periodic verification of sound level meters. Recent investigations of historical data have shown that the frequency response for the microphone is the most likely cause for the rejection of a sound level meter under test. It is therefore very important to measure the frequency response of the sound level meter with acoustical input signals. In order to keep the cost of the test down, it is essential to be able to choose a fast and sufficiently accurate method. In this paper some of the possible measurement methods such as free-field comparison, electrostatic actuator, multifrequency calibrator, and comparison coupler are described. The focus is on their advantages, drawbacks, and their associated uncertainties.

10:00

4aEA9. Perspectives on anechoic chamber qualification. Kenneth A. Cunefare and Van B. Biesel (School of Mech. Eng., Georgia Tech, Atlanta, GA 30332-0405)

The qualification of a new anechoic chamber requires demonstration that the chamber produces a free-field environment within some tolerance bounds and over some acceptable volume. At the most basic level, qualification requires measurement of sound levels at increasing distances from a test source, and then comparing the levels to a theoretical free-field decay. While simple in concept, the actual performance of a qualification test is problematic in implementation, with troublesome issues relevant to the nature of the sound source, test signal (broadband or pure tone), spatial resolution of measurements (e.g., measurements at discrete locations or spatially continuous), and comparison of the data to a theoretical decay. This presentation will provide a brief historical perspective on chamber qualification and review current practice. It will demonstrate the inadequacy of broadband noise and widely spaced discrete measurements for qualification purposes. It will demonstrate that pure tone signals and spatially continuous measurements provide a rigorous test of a chambers performance.

Invited Papers

10:15

4aEA10. IEC measurement standards for ultrasonic hydrophones and radiation force balances. K. Beissner (Physikalisch-Technische Bundesanstalt, 38116 Braunschweig, Germany)

Acoustic output measurements of medical ultrasonic equipment are usually performed in water and using two widely established methods: ultrasonic power measurement with a radiation force balance and acoustic pressure measurement with an ultrasonic hydrophone. The International Electrotechnical Commission (IEC) has published a number of International Standards on both methods, prepared by its Technical Committee 87 ''Ultrasonics.'' These are discussed and future standardization trends are presented. The basic standard for ultrasonic power measurements is IEC 61161. It recommends using a radiation force balance and specifies the relevant technical properties and sources of uncertainty. It is currently envisaged to revise this standard and to include particular recommendations for the power range up to 20 W which have recently been worked out in an international cooperation of national metrology institutes. Hydrophone standards have continuously been developed over many years. The most recent one is IEC 62092 which extends the upper frequency limit of hydrophone standarization from 15 MHz to 40 MHz. The current trend is to reorganize all existing hydrophone standards and to find a new scheme of three parallel standards dealing with hydrophone properties, hydrophone use, and hydrophone calibration.

10:30

4aEA11. A BIPM/CIPM key comparison covering the calibration of ultrasonic hydrophones over the frequency range 1 MHz to 15 MHz. Bajram Zeqiri and Nigel D. Lee (Acoustical Metrology Group, Natl. Physical Lab., Teddington, Middlesex TW11 0LW, UK, bajram.zeqiri@npl.co.uk)

A central objective of the Mutual Recognition Arrangement (MRA), signed by national measurement institute (NMI) directors in 1999, is the establishment of the degrees of equivalence of national measurement standards held by each institute. International comparisons, known as key comparisons, represent the sole mechanism for establishing these degrees of equivalence. In this paper we describe a key comparison, undertaken under the auspices of the BIPM/CIPM Consultative Committee for Acoustics, Ultrasound, and Vibration, related to the realization of the acoustic pascal in water at ultrasonic frequencies. This is most appropriately achieved through a comparison of calibrations of stable transfer standard hydrophones; 1 mm active element bilaminar membrane hydrophones, being chosen for this purpose. With NPL acting as the pilot laboratory, two hydrophones were calibrated using the NPL primary standard laser interferometer and circulated sequentially to participant NMI laboratories in Germany, China, The Netherlands, and Denmark. Laboratories were asked to report values for the hydrophone open-circuit free-field sensitivity over the frequency range 1 MHz to 15 MHz. The principal calibration methods used by the NMIs were optical interferometry and/or two-transducer reciprocity. The key comparison process, its results, and the analysis used to derive the key comparison reference values, are all described in detail.

10:45

4aEA12. ISO standardization to ensure traceability of vibration and shock acceleration measurements. Hans-Jürgen von Martens (Physikalisch Technische Bundesanstalt PTB, Bundesallee 100, 38116 Braunschweig, Germany, hans-juergen.v.martens@ptb.de)

Traceability to the International System of Units (SI) is increasingly demanded for vibration and shock measurements as specified in international standards, recommendations, and regulations to ensure product quality, health, and safety. Hierarchies of measurement standards have been established and are operated at the National Metrology Institutes (NMIs) as well as accredited and nonaccredited calibration laboratories. To meet the increasing demands, a revision of the ISO 5347 series, renumbered to ISO 16063, started in 1995 (ISO TC 108). Improved and new standard methods for vibration and shock calibration have been specified at the different traceability levels. For primary vibration calibration by laser interferometry at the NMI level, ISO 16063-11:1999 has extended the frequency range (0,4 Hz–10 kHz) and included absolute-phase shift measurement. ISO 16063-13:2001 provides interferometric primary shock

calibration (100 m/s^2–100 000 m/s^2). For vibration and shock calibration at lower levels, parts 21 and 22 of ISO 16063 provide upgraded comparison methods. ISO standards for the calibration of angular transducers by laser interferometry (part 15) and by comparison (part 23) are under development. Using the methods specified, national standards and calibration equipment have been developed for calibration laboratories, offering many new and improved calibration and measurement capabilities and thus international traceability.

11:00

4aEA13. On the uncertainties of accelerometer calibration by laser interferometry. Guillermo Silva-Pineda (Centro Nacional de Metrologa, km 4,5 Carr. a los Cus, El Marques, Qro. CP 76241, Mexico)

Depending on both measurement standards and measurement methods, derived quantities such as acceleration are traceable to the basic quantities in the International System of Units, SI. The uncertainty of measurement is often related to the length of the traceability chain up to the standards of reference quantities and the accuracy of the methods used in every calibration step. The traceability chart of the experimental arrangements can be used to visualize the different uncertainty levels and their relationships from the application of measurement to the standards of basic quantities in the SI. For the case of accelerometer calibration, experimental results of the fringe-counting method using a Michelson interferometer, which has standards traceable to basic units of length and time, are shown. Also, the physical principles and the alignment of the laser interferometer are discussed, and how the measurement accuracy can be improved. The classical approach of the uncertainty budget, based on a first-order Taylor series approximation, for the sensitivity calibration of an accelerometer is compared with a model based on higher-order terms in the Taylor series approximation. Some concluding remarks obtained from the comparison of these two approaches are given.

11:15

4aEA14. International documentary standards and comparison of national physical measurement standards for the calibration of accelerometers. David J. Evans (Natl. Inst. of Standards and Technol., 100 Bureau Dr., M.S. 8221, Gaithersburg, MD 20899-8221)

The documentary standards defining internationally adopted methodologies and protocols for calibrating transducers used to measure vibration are currently developed under the International Organization for Standardization (ISO) Technical Committee 108 Sub Committee 3 (Use and calibration of vibration and shock measuring instruments). Recent revisions of the documentary standards on primary methods for the calibration of accelerometers used to measure rectilinear motion have been completed. These standards can be, and have been, used as references in the technical protocols of key international and regional comparisons between National Measurement Institutes (NMIs) on the calibration of accelerometers. These key comparisons are occurring in part as a result of the creation of the Mutual Recognition Arrangement between NMIs which has appendices that document the uncertainties, and the comparisons completed in support of the uncertainties, claimed by the National Laboratories that are signatories of the MRA. The measurements for the first international and the first Interamerican System of Metrology (SIM) regional key comparisons in vibration have been completed. These intercomparisons were promulgated via the relatively new Consultative Committee for Acoustics, Ultrasound and Vibration (CCAUV) of the International Committee for Weights and Measures (CIPM) and SIM Metrology Working Group (MWG) 9, respectively.

Contributed Papers

11:30

4aEA15. High-frequency acceleration comparison among NIST, CENAM, and INMETRO. Gustavo P. Ripper (INMETRO, DIAVI/LAVIB, Av. N. S. das Graças, 50, Xerém-D. Caxias, Rio de Janeiro 25250-020, Brazil, lavib@inmetro.gov.br), Beverly F. Payne (NIST, Gaithersburg, MD 20899-8221), and Guillermo Silva-Pineda (CENAM, C.P. 76900, Querétaro, Mexico)

This paper presents the results obtained in an international laboratory comparison among NIST/USA, CENAM/Mexico, and INMETRO/Brazil. This acceleration comparison was focused on the calibration of the charge sensitivity of a single-ended standard accelerometer in the frequency range from 3 kHz to 10 kHz. The measurements were carried out by laser interferometry in accordance with the standard ISO 16063-11:1999 and with a preagreed protocol. This comparison supplements the former comparison SIM.AUV.V-K1, which covered the frequency range from 50 Hz to 5 kHz.

11:45

4aEA16. Charge amplifier calibration comparison among NIST, CENAM, and INMETRO. Gustavo P. Ripper (INMETRO, DIAVI, Av. N. S. das Graças, 50, Xerém-D. Caxias, Rio de Janeiro 25250-020, Brazil, lavib@inmetro.gov.br), Beverly F. Payne (NIST, Gaithersburg, MD 20899-8221, payne@nist.gov), and Guillermo Silva-Pineda (CENAM, Querétaro, Mexico)

This paper presents the results obtained in an international laboratory comparison among NIST/USA, CENAM/Mexico and INMETRO/Brazil which was focused on the electric calibration of a charge amplifier B & K model 2626 from 50 Hz to 5 kHz. One of the main concerns of this comparison was to conclude all measurements within the period of 1 month, in order to minimize the influence of long-term temporal instabilities. Each participating laboratory carried out the measurements in accordance with a calibration protocol, which was established to avoid the effect of additional sources of uncertainty. The results are compared with the ones obtained during the former comparison SIM.AUV.V-K1 and the conclusions are presented.

Session 4aED

Education in Acoustics: Virtual Labs, Workshops and Multimedia in Acoustics Education

Ralph T. Muehleisen, Chair
Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, Colorado 80309-0428

Chair's Introduction—8:30

Invited Papers

8:35

4aED1. Concerthalls.org: A webpage for architectural acoustics education. Lily M. Wang and Jessica M. Hall (Architectural Eng. Prog., Univ. of Nebraska–Lincoln, 200B Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0681, lwang@unl.edu)

A website focusing on concert hall acoustics (www.concerthalls.org) has been developed under the Schultz Grant from the Newman Student Award Fund. The website includes historical information, discussions on a variety of architectural acoustical measures, links to other websites of interest including a collection of concert hall webpages, and a comprehensive reference list. Of particular interest are the many listening examples provided on the website, which help students and other visitors to understand each subjective quality audibly and give insight on how to measure and control the associated objective measure. Examples are provided for reverberation, clarity, intimacy, warmth, loudness, and spaciousness. Various samples will be played during this presentation. Instructors of architectural acoustics are encouraged to introduce the site to their students and incorporate it into their instructional materials. [Work supported by Schultz Grant from the Newman Student Award Fund.]

8:55

4aED2. Animations and auralizations for noise control education. Ralph T. Muehleisen (Univ. of Colorado, Boulder, CO 80309-0428)

For students just starting out in acoustics, some basic concepts can be terribly difficult to understand. Ideas like plane, spherical, travelling, standing, and evanescent waves are easily described mathematically but can be difficult to understand physically. The use of animation to show the particle motion, velocity, and/or pressure fields associated with these wave motions helps to solidify the understanding of the basic physics. In noise control, students often have problems conceiving the effects of reverberation on acoustic signals and the perceived noise reduction from various acoustic treatments. These acoustic effects can be simulated through the use of auralization. Animations and auralizations developed using MATLAB, CATT Acoustics, and Cooledit for a senior level architectural engineering undergraduate classes in building noise control and acoustical room design will be presented.

9:15

4aED3. Acoustics workshops for teachers in elementary schools and in secondary school science classes. Uwe J. Hansen (Dept. of Phys., Indiana State Univ., Terre Haute, IN 47809)

A number of acoustics workshops have been conducted throughout the US with support from ASA technical initiative and education funds. Emphasis in elementary schools is on using music to introduce science concepts. Concepts discussed include wave properties such as wavelength, frequency, propagation speed, and standing waves. The idea of resonance is experienced musically and physically, and the relationship between whole number frequency multiples and the musical harmonic series is developed. Computer programs such as COOL-EDIT are used extensively in secondary teacher workshops to experience frequency analysis, frequency synthesis, and resonances in open and closed tubes. Workshop teaching approaches for these topics will be illustrated in this presentation.

9:35

4aED4. Computer simulations enhance experimental demonstrations in the underwater acoustics and sonar course. Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402)

Underwater acoustics and sonar (SP411) is a 3 hour course that is offered to midshipmen in their senior year. Typically, general science and oceanography majors, totaling 110 students/yr, enroll. Since this course is offered without a lab, the ''in class'' experience has been enhanced with the development (over many years) of our demo carts and computer workstations which surround the classroom. In a studio classroom atmosphere, students perform a variety of experiments in small groups. How can scientific visualizations best develop learning of complex interactions? Two examples are presented. PC–IMAT (personal curriculum interactive multisensor analysis training) simulations of multielement array steering support the theory and enhance the experiments that are performed in class such as the two-element array. Mathematica simulations involving the programming and animation of a point source in a rigid–rigid infinite parallel wave guide are used to stress the method of images, superposition, group and phase velocity and far-field modal pattern that is observed as a function of depth and source frequency. Later, students have fun using a ripple tank with an eye dropper to generate a point source between two adjustable parallel boundaries, and their understanding of ''underwater sound'' is greatly enhanced.

Contributed Papers

10:05

4aED5. Experiences on Science, Technology and Society (ScTS) in teaching environmental noise. Pablo Roberto Lizana Paulin, Jose de Jesus Negrete Redondo, and Elena Romero Rizo (Acoust. Lab., ESIME, IPN, Mexico, plizana98@hotmail.com)

The ScTS (Science, Technology and Society) is a new work field for a critical and interdisciplinary way to evaluate the social dimension of science and technology with regard to the social and environmental effects, in this case of environmental noise contamination. This new way to see science promotes reflection and modification of personal and professional values. Teams were organized to research the environmental noise impact on society and the human being. Then discussions were carried out in order to prepare classroom presentations. After the presentations and proposals, further discussion and conclusions about the problem were made, getting the students involved in the possible solutions. (To be presented in Spanish.)

10:20

4aED6. Help! There are 60 screaming kids in my lab! Corinne Darvennes (Dept. of Mech. Eng., Tennessee Technolog. Univ., Box 5014, Cookeville, TN 38505, CDarvennes@tntech.edu)

What do you do when you have a large noise control facility available and you want to introduce local children to science and engineering? You invite them in! This presentation will describe a day of hands-on activities for 60 fifth graders in the Acoustics and Vibrations Laboratory at TTU. This includes the logistics of having 60 kids in the lab and keeping their attention, a description of each activity, and the equipment used. Activities included learning about hearing and hearing loss, propagating and standing waves, ultrasound, vibrations, and resonance. Students used slinkies, sound level meters, a strobe light, Vernier LabPro data collection systems and sensors (ultrasonic motion detectors and microphones). They popped balloons in the reverberation chamber to listen to the sound decay. Their most memorable experience, however, was the screaming contest in the anechoic room!

10:35

4aED7. Didactic toy for teaching the Fourier principle. Miguel Medina-Villanueva, Victor Medina-Tamez, Javier Medina-Tamez, and Adrian Garcia-Mederez (Acoust. Lab., FIME, Univ. A. de Nuevo Leon, P.O. Box 28 "*F*," Cd. Universitaria, San Nicolas, 66450 N.L., Mexico)

The mathematical tool of Fourier analysis is used in many areas like vibrations, communications, optics, electronics, etc. The understanding of this subject sometimes causes frustration in students. The main purpose of this presentation is to propose a didactic toy that calculates the harmonic magnitudes through the discrete values of analog periodic signals. This device shows the rotative vectors in a physical way that makes the principle of Fourier understandable.

10:50

4aED8. Using computers to overcome math-phobia in an introductory course in musical acoustics. Andrew A. Piacsek (Central Washington Univ., Ellensburg, WA 98926)

In recent years, the desktop computer has acquired the signal processing and visualization capabilities once obtained only with expensive specialized equipment. With the appropriate A/D card and software, a PC can behave like an oscilloscope, a real-time signal analyzer, a function generator, and a synthesizer, with both audio and visual outputs. In addition, the computer can be used to visualize specific wave behavior, such as superposition and standing waves, refraction, dispersion, etc. These capabilities make the computer an invaluable tool to teach basic acoustic principles to students with very poor math skills. In this paper I describe my approach to teaching the introductory-level Physics of Musical Sound at Central Washington University, in which very few science students enroll. Emphasis is placed on how vizualization with computers can help students appreciate and apply quantitative methods for analyzing sound.

11:05

4aED9. Lab experiment using physical models of the human vocal tract for high-school students. Eri Maeda, Takayuki Arai, Noriko Saika, and Yuji Murahara (Dept. of Elec. and Electron. Eng., Sophia Univ., 7-1 Kioi-cho, Chiyoda-ku, Tokyo 102-8554, Japan, m-eri@sophia.ac.jp)

Recently, the development of educational tools for acoustics has become popular in Japan. We believe that physical models of the human vocal tract are particularly useful for teaching acoustics. Formerly we proposed three models of the vocal tract corresponding to the Japanese vowels, /i/, /e/, /a/, /o/, and /u/. We presented cylindrical, nasalized, and plate type models. The models were made of transparent acrylic resin, enabling the configurations of the oral cavity to be seen from the outside of the model. In this presentation, we will discuss the results of a lab experiment in which we used these tools to teach the mechanism of vowel production to high-school students who had just finished studying basic acoustics. By manipulating the plates in the plate type model, students were able to simulate constrictions at nodes and antinodes, and they were able to hear the shift in formant frequencies. The exercise helped students to understand vowel production. We received positive feedback from those who participated in the experiment.

11:20

4aED10. Physical models of the human vocal tract as tools for education in acoustics. Takayuki Arai, Eri Maeda, Noriko Saika, and Yuji Murahara (Dept. of Elec. and Electron. Eng., Sophia Univ., 7-1 Kioi-cho, Chiyoda-ku, Tokyo 102-8554, Japan)

We recently replicated the Chiba and Kajiyama physical models of the human vocal tract and found that they are extremely effective as a manipulative tool for education in acoustics. This study is an overview of our previous development of the models and their application for education. The models we developed simulate the configuration of the vocal tract, and are of two major types: cylinder and plate. The interior, bottle shape of the cylindrical models was based on the original measurements of Chiba and Kajiyama. The diameters of the holes in the 10–17 movable plates of the plate-type model were also based on the Chiba and Kajiyama measurements, but the radius curves were approximated at 10 mm resolution in a stepwise fashion. We also developed a nasalized model, which consists of a side branch attaching to a cylindrical model. Our experience confirmed that when used in a classroom environment, the models increase student understanding of the acoustic theory of speech production.

11:35

4aED11. Musical appreciation. Maria del Consuelo Medina (Departamento de Musica, Centro Universitario de Arte Arquitectura y Diseo, Universidad de Guadalajara, Morelos 181, Zona Centro, Guadalajara, Jal. Mexico)

Pre-school listening to music is the principal way that leads to the appreciation of music that later facilitates knowledge and pleasure in the history of music. At the prescholastic age it is a very important aspect of education, and reasons and suggestions will be given. The activities must be brief, the teachers of music can at the most develop the activity every

five minutes, leaving time for rest or expansion. Another suitable way to bring the child to music is through stories, which please all children; let them go to an unreal and fantastic world and listen to a story or an exciting adventure. The story then, should be brief, simple, with action, with familiar characters, but with some mystery; some repetitive element; and an ending both surprising and happy. It is preferable to include small folkloric tales from the universal repertoire, with works of simple and clear structure.

THURSDAY MORNING, 5 DECEMBER 2002 CORAL GARDEN 2 AND 3, 7:55 A.M. TO 12:10 P.M.

Session 4aNS

Noise and Structural Acoustics and Vibration: Aircraft Interior Noise

David Reed, Cochair
The Boeing Company, P.O. Box 3707, Seattle, Washington 98124-2207

Kevin P. Shepherd, Cochair
Structural Acoustics Branch, NASA Langley Research Center, Hampton, Virginia 23681

Chair's Introduction—7:55

Invited Papers

8:00

4aNS1. Recent advances in active control of aircraft cabin noise. Gopal Mathur (The Boeing Co., 5301 Bolsa Ave., MC: H013-B308, Huntington Beach, CA 92647) and Christopher Fuller (Virginia Polytechnic Inst. & State Univ., Blacksburg, VA 24061)

Active noise control techniques can provide significant reductions in aircraft interior noise levels without the structural modifications or weight penalties usually associated with passive techniques, particularly for low frequency noise. Our main objective in this presentation is to give a review of active control methods and their applications to aircraft cabin noise reduction with an emphasis on recent advances and challenges facing the noise control engineer in the practical application of these techniques. The active noise control method using secondary acoustic sources, e.g., loudspeakers, as control sources for tonal noise reduction is first discussed with results from an active noise control flight test demonstration. An innovative approach of applying control forces directly to the fuselage structure using piezoelectric actuators, known as active structural acoustic control (ASAC), to control cabin noise is then presented. Experimental results from laboratory ASAC tests conducted on a full-scale fuselage and from flight tests on a helicopter will be discussed. Finally, a hybrid active/passive noise control approach for achieving significant broadband noise reduction will be discussed. Experimental results of control of broadband noise transmission through an aircraft structure will be presented.

8:20

4aNS2. Recent advances in active noise and vibration control at NASA Langley Research Center. Gary P. Gibbs, Randolph H. Cabell, Daniel L. Palumbo, Richard J. Silcox, and Travis L. Turner (Structural Acoust. Branch, NASA LaRC, Hampton, VA 23681)

Over the past 15 years NASA has investigated the use of active control technology for aircraft interior noise. More recently this work has been supported through the Advanced Subsonic Technology Noise Reduction Program (1994–2001), High Speed Research Program (1994–1999), and through the Quiet Aircraft Technology Program (2000–present). The interior environment is recognized as an important element in flight safety, crew communications and fatigue, as well as passenger comfort. This presentation will overview research in active noise and vibration control relating to interior noise being investigated by NASA. The research to be presented includes: active control of aircraft fuselage sidewall transmission due to turbulent boundary layer or jet noise excitation, active control of interior tones due to propeller excitation of aircraft structures, and adaptive stiffening of structures for noise, vibration, and fatigue control. Work on actuator technology ranging from piezoelectrics, shape memory actuators, and fluidic actuators will be described including applications. Control system technology will be included that is experimentally based, real-time, and adaptive.

8:40

4aNS3. Interior noise control through tuned loudspeaker absorption. Robert L. Clark and Mark A. McEver (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., 301 Hudson Hall, Durham, NC 27708-0300)

In this work we describe the development and application of an interior noise control approach based on active loudspeaker tuning for sound absorption. Loudspeakers are tuned to minimize global sound pressure at specific acoustic modes through diaphragm velocity feedback. An experimental model of the acoustic space used for controller design is obtained through system identification. H_2 optimal controllers are designed to minimize a performance microphone output by feeding back the tuned loudspeaker diaphragm velocity. By sensing speaker velocity rather than far-field sound pressure, the loudspeaker is actively tuned to minimize reflected sound at specific acoustic modes and therefore reduce the global pressure field. Experimental results, obtained from an acoustic enclosure modeled after a rocket fairing, are used to demonstrate the effectiveness of the tuned loudspeaker at global attenuation. Details of the design procedure and experimental applications are discussed.

9:00

4aNS4. Cabin noise and weight reduction program for the Gulfstream G200. C. Kearney Barton (Gulfstream Aerosp. Corp., P.O. Box 2206, Savannah, GA 31402)

This paper describes the approach and logic involved in a cabin noise and weight reduction program for an existing aircraft that was already in service with a pre-existing insulation package. The aircraft, a Gulfstream G200, was formally an IAI Galaxy, and the program was purchased from IAI in 2001. The approach was to investigate every aspect of the aircraft that could be a factor for cabin noise. This included such items as engine mounting and balancing criteria, the hydraulic system, the pressurization and air-conditioning system, the outflow valve, the interior shell and mounting system, antennae and other hull protuberances, as well as the insulation package. Each of these items was evaluated as potential candidates for noise and weight control modifications. Although the program is still ongoing, the results to date include a 175-lb weight savings and a 5-dB reduction in the cabin average Speech Interference Level (SIL).

9:20

4aNS5. An in-flight study of cabin buzz-saw noise. David Reed, Stefan Uellenberg, and Evan Davis (The Boeing Co., P.O. Box 3707-MC 67-ML, Seattle, WA 98124)

This paper examines the characteristics of multiple-pure-tone noise generated by high-speed turbofans under conditions of supersonic fan tip speeds, especially as it is observed in an aircraft passenger cabin. This phenomenon, also known as buzz-saw noise, is an important noise source in commercial airplane passenger cabins and has proved to be difficult to treat with sound-absorbing materials. Recent flight test experiments by The Boeing Company have demonstrated extraordinary success in suppressing cabin buzz-saw noise by strategic placement of engine inlet acoustic linings. The observed behavior is explained by a propagation and radiation model, which is validated by in-flight measurements made with a phased microphone array mounted on the fuselage skin of a Boeing 777. A structural acoustic model is also offered to explain the different transmission characteristics of buzz-saw noise and turbulent boundary layer excitation. Correlation length scales measured on the fuselage surface for these two noise sources are key inputs to the structural model.

9:40

4aNS6. Structural acoustic modeling of aircraft fuselage structures. Ralph Buehrle, Jacob Klos, Jay Robinson (NASA Langley Res. Ctr., Structural Acoust. Branch, M.S. 463, Hampton, VA 23681), and Ferdinand Grosveld (Lockheed Martin Eng. and Sci., Hampton, VA)

Recent studies at NASA Langley Research Center have examined the development and validation of finite element and boundary element modeling techniques for the prediction of structural acoustic response of aircraft fuselage structures. The goal of this work is to provide increased confidence in the modeling techniques so that interior noise criteria can be incorporated early in the design process. These efforts have focused on the development and validation of high-fidelity physics-based numerical models for structural acoustic predictions into the kilohertz region. Finite element models were developed based on the geometric and material properties of the aircraft fuselage structures. Experimental modal analysis and point force frequency response functions were used to validate and refine the finite element models. Once validated, the finite element predictions of the velocity response were used as boundary condition input for boundary element predictions of the radiated sound power. Experiments in the Structural Acoustic Loads and transmission (SALT) Facility at NASA Langley were used to validate the acoustic predictions. Numerical and experimental results will be presented for conventional aluminum rib and stringer-stiffened aircraft structures, a honey comb composite sidewall panel, and damped acrylic windows. Numerical predictions were in good agreement with the experimental data.

10:00–10:10 Break

10:10

4aNS7. Uncertain structural dynamics of aircraft panels and fuzzy structures analysis. Victor W. Sparrow (Grad. Prog. in Acoust., Penn State Univ., 316B Leonhard Bldg., University Park, PA 16802, vws1@psu.edu) and Ralph D. Buehrle (NASA Langley Res. Ctr., Hampton, VA 23681)

Aircraft fuselage panels, seemingly simple structures, are actually complex because of the uncertainty of the attachments of the frame stiffeners and longitudinal stringers. It is clearly important to understand the dynamics of these panels because of the subsequent radiation into the passenger cabin, even when complete information is not available for all portions of the finite-element model. Over the last few years a fuzzy structures analysis (FSA) approach has been undertaken at Penn State and NASA Langley to quantify the uncertainty in modeling aircraft panels. A new MSC.Nastran [MSC.Software Corp. (Santa Ana, CA)] Direct Matrix Abstraction Program (DMAP) code was written and tested [AIAA paper 2001-1320, 42nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conf., Seattle, WA, 16 April 2001] and was applied to simple fuselage panel models [J. Acoust. Soc. Am. **109**, 2410(A) (2001)]. Recently the work has focused on understanding the dynamics of a realistic aluminum fuselage panel, typical of today's aircraft construction. This presentation will provide an overview of the research and recent results will be given for the fuselage panel. Comparison between experiments and the FSA results will be shown for different fuzzy input parameters. [Work supported by NASA Research Cooperative Agreement NCC-1-382.]

10:30

4aNS8. New analysis methods for the calculation of interior noise. Donald B. Bliss and Linda P. Franzoni (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708, dbb@duke.edu)

Recently developed analysis methods applicable to the prediction of aircraft interior noise are described. Analytical-numerical matching (ANM) and Local-global homogenization (LGH) are methods for the low- to mid-frequency range. Energy-intensity boundary elements are applicable for high frequencies. ANM is a method that replaces the effect of structural discontinuities (ribs, stringers, attachments) with dynamically equivalent smooth force distributions, thereby removing the need for high-resolution numerics in the overall computational problem. LGH, which applies for periodic and quasiperiodic structures, provides a means to solve directly for the long wavelength spatial content of the structure. This part of the structural motion is strongly coupled most to the interior sound field. Energy-intensity boundary elements provide a very efficient means to predict high-frequency broadband sound fields in a fundamentally correct manner. Current research on this boundary element method involves its extension to handle coupling with structural surfaces vibrating at high frequency. The strengths and limitations of each of the above methods is discussed in the context of the interior noise problem. Innovative methods for the reduction of interior noise, as suggested by these approaches, are also mentioned.

10:50

4aNS9. A prototype spherical array for interior noise investigations. Earl G. Williams, Brian H. Houston (Code 7130, Naval Res. Lab., Washington, DC 20375), and Peter C. Herdic (SFA, Inc., Landover, MD)

A spherical array of radius 0.2 m with 18 uniformly distributed microphone elements on a spherical surface has been constructed. Using the theory of nearfield acoustical holography the instantaneous pressure, velocity, and intensity fields are reconstructed throughout a volume of 0.6 m radius centered at the array origin in the frequency range of 0–600 Hz. Since the measurements are instantaneous the array is intended to be used to measure distributed sources that need not be coherent or stationary, ideally suited for source identification in the interior of an in-flight aircraft, an automobile or in interior spaces in naval vessels. Preliminary experiments with a loudspeaker demonstrate the accuracy of the volume-field reconstruction of the instantaneous fields. This array is a prototype of an array with many more elements which will have a wider frequency range and finer spatial resolution. [Work supported by NASA and ONR.]

Contributed Papers

11:10

4aNS10. Recent developments in statistical energy analysis for aircraft interior noise. Paul Bremner, Michael Bloor (Vibro-Acoust. Sci., Inc., 12555 High Bluff Dr., Ste. 310, San Diego, CA 92130), and John F. Wilby (Wilby Assoc.)

Price and Crocker [J. Acoust. Soc. Am. **47** (1970)] have previously shown that Statistical Energy Analysis (SEA) is a robust tool for sound transmission loss problems in room acoustics. This paper presents recent developments in application of SEA to the structural-acoustic design of aircraft for control of interior noise. Developments for this application include the means to model boundary layer excitation of the fuselage; the means to model the dynamic response of both metallic-ribbed and composite sandwich pressurized shells; and the means to model the full noise and vibration reduction performance of multilayered interior trim and insulation treatments. The method has proven effective for acoustic design evaluation of new materials and constructions of aircraft structure or interiors but is equally attractive for optimization of the cost, weight, and packaging space of noise control treatments on aircraft of more conventional construction. Results of recent validation studies will be presented.

11:25

4aNS11. Application of statistical energy analysis to the design of crew rest compartments. Mark M. Gmerek (The Boeing Co., Seattle, WA 98124, mark.m.gmerek@boeing.com)

Longer flight times for modern commercial aircraft have led to the need for crew rest compartments. Noise levels in the crew rest compartments must be conducive to proper rest and recuperation. Statistical Energy Analysis (SEA) has been used to develop a 777 crew rest compartment design that achieves appropriate noise levels at acceptable weight and cost. In this paper the design of a 777 overhead crew rest compartment is outlined using SEA design tools and methods. Noise data were gathered in flight to distinguish airplane source components and develop model inputs. Crew rest panels, the airplane fuselage, and acoustic volumes were modeled as SEA subsystems by taking into account geometry, material properties, modulus, and damping. A model was built, excited with inputs, and analyzed to determine energy flow paths and acoustic pressure at receiver locations. Prospective add-on treatments were then assessed to engineer an effective noise control package. The model development was supplemented by laboratory sound transmission loss testing of individual components. The good agreement between the laboratory tests and individual SEA models of the components increased confidence in the approach. Once the crew rest was installed on the airplane, the measured in-flight noise levels closely matched the SEA estimates.

11:40

4aNS12. The analysis of a state-switched absorber design concept. Anne-Marie Albanese and Kenneth Cunefare (School of Mech. Eng., Georgia Tech, Atlanta, GA 30332-0405)

A tuned vibration absorber (TVA) is a spring-damper-mass system used in many industries for the suppression of a specific vibration frequency, and has application for the suppression of aircraft fuselage vibration. A state-switched absorber (SSA) is similar to a TVA, except that one or more components in the SSA is able to instantaneously and discretely change properties, thus increasing the effective bandwidth of vibration suppression. In order to design a replacement SSA for the classic TVA, the SSA must operate in the appropriate frequency range, be lightweight, and compact. An optimal SSA will also have a maximal frequency range between which it can switch. This paper discusses the development of a magnetorheological (MR) silicone gel used as the SSA switching element, the shape required to maintain a magnetic flux path, and the contribution of the magnet-mass to frequency shifting. The MR gel is iron-doped silicone, cured in the presence of a magnetic field. During operation, the applied magnetic flux is modified to change the natural frequency. The applied flux requirement forces the SSA to be a small ring. The SSA is designed to operate below 100 Hz.

11:55

4aNS13. Experimental evaluation of sound transmission through a thin-wall structure produced by turbulent pressure fluctuations. Boris M. Efimtsov (Acoust. Div., Central AeroHydrodynamics Inst. (TsAGI), 17 Radio St., 105005 Moscow, Russia, ayzverev@hotmail.com) and Sergey N. Baranov (RUSAVIA Res. Group., 125299 Moscow, Russia)

Problems of physical modeling of sound transmission through a thin-wall structure produced by turbulent pressure fluctuations under excitation are presented. A method of evaluating the sound transmission loss of a thin-wall structure from turbulent pressure fluctuations on the basis of experimental data is proposed. The experiment is associated with full scale tests of structure in two adjacent reverberation chambers and with model tests in two adjacent model reverberation chambers and in the wind tunnel equipped with a similar model sound chamber. An example of such method realization is presented and this indicates a possibility of its practical realization in predicting the noise inside the passenger aircraft, produced by the turbulent boundary layer on its exterior surface. [Work supported by INTAS Grant No. 99-0088.]

Session 4aPAa

Physical Acoustics: Laser Ultrasonics

Todd W. Murray, Chair
Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Invited Papers

9:00

4aPAa1. Ultrafast optoacoustic wave generation effects on picosecond-timescale reflectivity measurements in metal thin films. James B. Spicer (Johns Hopkins Univ., Rm. 102, Maryland Hall, 3400 N. Charles St., Baltimore, MD 21218, spicer@jhu.edu) and Christopher J. K. Richardson (Univ. of Maryland, College Park, MD 20740)

The increase in temperature at the surface of a metal during absorption of optical energy is a dynamic, multistep process. High-time-resolution, ultrafast measurements of thermal and elastic transients have proven to be useful for investigating the structure, thermophysical properties, and elastic properties of thin films [Paddock and Eesley, J. Appl. Phys. **60**, 285–290 (1986)]. During ultrafast optical excitation of metals, the rapid deposition of energy causes the absorbing electrons to enter a superheated state, dispersing beyond the excitation volume before energy is deposited to the lattice with the rate of energy transfer between electrons and phonons being described using the electron–phonon coupling parameter. Typically, ultrafast reflectivity measurements taken immediately after excitation can be modeled using a completely thermal analysis if reduced values for the electron–phonon coupling parameter are used. These reduced values often vary substantially [Gusev and Wright, Phys. Rev. B **57**, 2878–2888 (1998)] from accepted values that are experimentally determined from measurements using the cooling rate of superheated electrons. By including elastic contributions to the modeled signal at all times, agreement with measured signals can be obtained for short and long times using values for materials parameters that are in agreement with those obtained from other techniques.

9:25

4aPAa2. Dispersion and soliton formation from longitudinal acoustic phonons in crystalline solids. Humphrey J. Maris, Hsin-Yi Hao, and Wisit Singhsomroje (Dept. of Phys., Brown Univ., Providence, RI 02912)

Experiments are described in which the picosecond ultrasonic technique is used to study the effects of dispersion and elastic nonlinearity on the shape of longitudinal acoustic pulses propagating in crystalline solids. In these experiments, a subpicosecond light pulse (pump pulse) is used to generate an acoustic pulse at one surface of the sample. After propagation through the sample, the shape of this acoustic pulse is modified. When the amplitude of the pump pulse is sufficiently low, nonlinear effects can be neglected and a measurement of the acoustic pulse shape can be used to determine the phonon dispersion. This measured phonon dispersion is compared with various lattice dynamical models. When the pump pulse intensity is high, the nonlinear effects become strong enough to balance the dispersion, and acoustic solitons emerge from the original acoustic pulse. Measurements of the characteristics of these solitons are in agreement with the results of computer simulations based on the Korteweg-de Vries equation.

9:50

4aPAa3. Picosecond transient grating acoustic measurements of thin films, microfluidic networks, and microphononic bandgap structures. John Rogers (Bell Labs., 600 Mountain Ave., Rm. 1D-332, Murray Hill, NJ 07974)

We describe some new methods for transient grating (TG) photoacoustic measurements, and summarize their use in high frequency acoustic waveguide analysis of a variety of samples with dimensions smaller than 100 microns. This research is motivated in part by the growing interest in micro- and nanostructures whose mechanical and acoustic properties are important either for basic study or device design. Although the conventional TG technique provides a convenient method for exciting and detecting acoustic waves in bulk samples, it is not well suited to microsystems, where signal levels can be low and acoustic dispersion (i.e., variation in acoustic velocity with wavelength) can be pronounced. Here we illustrate how specially designed diffractive optics and imaging lenses remove these limitations by enabling convenient means for: (i) generating and overlapping many (up to 10 has been demonstrated)

excitation pulses for launching complex, user-definable acoustic waveforms that enable single-shot measurements of dispersion; and (ii) providing a phase-stabilized heterodyne detection scheme that enhances the sensitivity of the measurement and simplifies many of the alignment procedures. We describe our application of these methods to a range of samples, including thin films and multilayer assemblies, optical fibers, microfluidic networks, and microphononic bandgap structures.

10:15–10:30 Break

10:30

4aPAa4. Laser ultrasonic system for on-line steel tube gauging and process control. Daniel Levesque, Marc Choquet, Christian Padioleau, Christian Neron, Christian Corbeil, Richard Talbot, Abdelhakim Bendada, Jean-Pierre Monchalin (Industrial Mater. Inst., Natl. Res. Council of Canada, 75 de Mortagne Blvd., Boucherville, QC J4B 6Y4, Canada), Robert V. Kolarik II, Gerald V. Jeskey, Erich D. Dominik, Larry J. Duly, Kenneth J. Samblanet, Steven E. Agger, Kenneth J. Roush, and Michael L. Mester (The Timken Co., Canton, OH 44706)

A laser ultrasonic system has been installed on a seamless tubing production line at The Timken Company and is being used to measure on-line the wall thickness of tubes during processing. The seamless process consists essentially in forcing a mandrel through a hot cylindrical billet in rotation and results in wall thickness variations that should be minimized and controlled to respect specifications. The system includes a Q-switched Nd-YAG laser for the generation of ultrasound by ablation, a long pulse very stable Nd-YAG laser for detection coupled to a confocal Fabry–Perot interferometer. The lasers, data acquisition, and processing units are housed in a cabin off-line and connected to a front coupling head located over the passing tube by optical fibers. The system also includes a fiber-coupled pyrometer to measure tube temperature profile and two fiber-coupled optical velocimeters to measure the coordinates at the probing location on the surface of the passing, rotating hot tube. During the presentation further details of the system will be disclosed, as well as typical results and examples of its diagnostic capability. [Work partially supported by Department of Energy under Award No. DE-FC07-99ID 13651.]

Contributed Papers

10:55

4aPAa5. Rayleigh scattering and heterodyne detection as a nonintrusive microphone. Cesar Aguilar, Carlos Azpeitia, Andres Porta, and Catalina Stern (Lab. de Acustica, Depto Fisica, Fac. Ciencias, UNAM, Mexico)

Heterodyne detection of a monochromatic laser beam scattered by molecules in a transparent gas gives a signal that is proportional to the spatial Fourier transform of density fluctuations for a wave vector determined by the optics. In a turbulent air jet, density fluctuations are due to either acoustic waves or to entropy fluctuations. This technique captures both. The method can then be used as a nonintrusive microphone to study the propagation of acoustic waves inside the jet. At each point in the flow the direction of propagation of the acoustic wave is determined, and the acoustic field can be described. This will eventually help to localize the sources of the waves in the flow and hopefully determine the hydrodynamic event that generates these waves.

11:10

4aPAa6. Laser ultrasonic in-process inspection of paper for elastic properties. Paul Ridgway, Richard Russo (Lawrence Berkeley Natl. Lab., 1 Cyclotron Rd., Berkeley, CA 94720), Emmanuel Lafond, Charles Habeger, Jr., and Ted Jackson (Inst. of Paper Sci. and Technol., Atlanta, GA 30318)

A laser-ultrasonic (LUS) sensor has been developed that allows measurement of the bending stiffness (BS) and shear rigidity (SR) of paper and paperboard as it is being made on the papermaking machine. A prototype system was recently tested in a paper mill at web speeds up to 5000 ft/min with excellent precision and accuracy. The LUS technique performs well on paper and board with basis weights up to 130 g/m^2. Several laboratory methods exist for measuring the bending stiffness in small samples of paper and board. Currently, no commercial method exists for nondestructively measuring this property on the papermaking machine at production speeds. Commercial instruments using contact transducers measure ''tensile strength orientation'' (TSO) on heavier boards, where marking of the sheet by the contact transducers is not of concern. Unlike contact ultrasonic techniques, LUS does not visibly mark even the lightest grade papers. Contact ultrasonic measurements correlate approximately to the tensile strength of the sheet and can be used to calculate an approximate value for BS. LUS measurements are directly related to BS and should yield more accurate determinations. Optimum use of feed stock, reduced waste, and decreased energy consumption are potential benefits of the LUS technology.

11:25

4aPAa7. Characterization of graded coatings using laser generated acoustic waves. Oluwaseyi O. Balogun, Todd W. Murray (Aerosp. and Mech. Eng. Dept., Boston Univ., 110 Cummington St., Boston, MA 02215), and Soumendra N. Basu (Boston Univ., Brookline, MA 02446)

Theoretical and experimental results are presented for the pulsed laser generation of ultrasound in functionally graded coating materials. The case of 9–12 m compositionally graded mullite ($3Al_2O_3$:$2SiO_2$) environmental barrier coatings in which the Al composition varies as a function of depth is considered in detail. The laser source is seen to generate both the lowest order Rayleigh mode and Sezawa mode in the coatings. The waveforms are processed using the 2-D FFT technique to extract the dispersion curves. Theoretical dispersion curves and time domain signals generated in coatings with linear, sigmoidal, and exponential compositional gradations will be presented and the sensitivity of the wave modes to elastic property gradation explored. The graded coatings are treated as multilayer materials with a sufficient number of layers chosen to achieve convergence of the dispersion curves for a given elastic property gradation. The (depth dependent) elastic properties of the coatings are evaluated using the simplex optimization routine and the functional form of the property variation is correlated with the measured compositional profiles. This work demonstrates that laser based ultrasonic measurements of acoustic modes in functionally graded EBCs provides a useful tool for extracting the depth dependent elastic properties.

Session 4aPAb

Physical Acoustics: Global Infrasound Monitoring I

Henry E. Bass, Chair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Chair's Introduction—10:10

Invited Papers

10:15

4aPAb1. The verification system of the Comprehensive Nuclear-Test-Ban Treaty. Gerardo Suarez (Intl. Monitoring Systems, Vienna Intl. Ctr., P.O. Box 1200, A-1400 Vienna, Austria, gerardo.suarez@ctbto.org)

The Comprehensive Nuclear-Test-Ban Treaty was opened for signature in September 1996. To date, the treaty has been signed by 165 countries and ratified by 93; among the latter, 31 out of the 44 whose ratification is needed for the treaty to enter into force. The treaty calls for the installation and operation of a verification system to ensure compliance. The verification system is composed of the International Monitoring System (IMS), the International Data Centre (IDC), and the On Site Inspection Division (OSI). The IMS is a global network of 321 stations hosted by 90 countries. The primary network is composed of 50 seismic stations, 31 of which are seismic arrays and 19 three-component, broad-band stations, 11 hydroacoustic stations, 60 infrasound arrays, and 80 radionuclide monitoring stations measuring radioactive particulates and noble gases in the atmosphere. The radionuclide network is supported by 16 laboratories. The auxiliary network of 120 seismic stations is interrogated on request by the IDC to improve the accuracy of the locations. The data from the 321 stations and from the laboratories is transmitted to the IDC in Vienna via a dedicated Global Communication Infrastructure (GCI) based on VSAT antennas. The IDC collects and processes the data collected from the four technologies and produces bulletins of events. The raw data and bulletins are distributed to state signatories. Upon entry into force, an on-site inspection may be carried out if it is suspected that a nuclear explosion has taken place. Since mid-1997, when the Provisional Technical Secretariat responsible for the implementation of the verification system began its work in Vienna, over 86% of the sites have been surveyed and the final location of the stations selected. By the end of 2002 this number will reach about 90%, essentially completing this phase. To date, 131 stations have been built or upgraded, and 80 are now sending data to the IDC; 112 others are under construction or under negotiation. Over 392 authorized users from 53 state signatories are now receiving data and products from the IDC. In addition, under the auspices of the signatories of the treaty, several projects are being funded to improve the calibration of seismic stations, the software used routinely at the NDC, and the analysis of radionuclides, among others.

10:35

4aPAb2. Monitoring of atmospheric nuclear explosions with infrasonic microphone arrays. Charles R. Wilson (Geophysical Inst., Univ. of Alaska, Fairbanks, AK 99775)

A review is given of the various United States programs for the infrasonic monitoring of atmospheric nuclear explosions from their inception in 1946 to their termination in 1975. The US Atomic Energy Detection System (USAEDS) monitored all nuclear weapons tests that were conducted by the Soviet Union, France, China, and the US with arrays of sensitive microbarographs in a worldwide network of infrasonic stations. A discussion of the source mechanism for the creation and subsequent propagation around the globe of long wavelength infrasound from explosions (volcanic and nuclear) is given to show the efficacy of infrasonic monitoring for the detection of atmospheric nuclear weapons tests. The equipment that was used for infrasound detection, the design of the sensor arrays, and the data processing techniques that were used by USAEDS are all discussed.

10:55

4aPAb3. Processing and analysis of infrasound data at the Comprehensive Nuclear Test Ban Treaty International Data Center. Robert G. North and Robert L. Woodward (Sci. Appl. Intl. Corp., Ctr. for Monitoring Res., 1300 N. 17th St., Ste. 1450, Arlington, VA 22209)

The Center for Monitoring Research (CMR) in Arlington, VA served as a prototype of the International Data Center (IDC) that is specified in the Comprehensive Nuclear Test Ban Treaty (CTBT). The CMR developed the software that is now being used by the IDC that has been established in Vienna to acquire, process, and analyze data from the global networks of hydroacoustic, infrasonic, radionuclide, and seismic sensors that are being installed to assist verification of the CTBT. The prototype IDC tested this software and related procedures very extensively with seismic and radionuclide data, but the infrasonic aspects of the system are less advanced due to the limited number of sensors available and a less complete knowledge of sources and signal propagation. The infrasonic processing

system attempts to detect signals of interest by combining spatial coherence and power, and estimates signal arrival time, direction of approach, velocity, and other parameters. After a signal is detected, further processing is carried out to try to identify ''phases'' that correspond to reflections from different atmospheric layers. Finally, detections at multiple sensors of all types (seismic and hydroacoustic as well as infrasonic) are combined to estimate the time and location of potential sources.

Contributed Papers

11:15

4aPAb4. Status report on the establishment of the Comprehensive Nuclear-Test-Ban Treaty (CTBT) International Monitoring System (IMS) infrasound network. J. A. Vivas Veloso, D. R. Christie, P. Campus, M. Bell, T. L. Hoffmann, A. Langlois, P. Martysevich, E. Demirovik, J. Carvalho, A. Kramer, and A. Molero Alonso (CTBTO, VIC, Wagramer Str. 5, 1400 Vienna, Austria)

The infrasound component of the International Monitoring System (IMS) for Comprehensive Nuclear-Test-Ban Treaty verification aims for global detection and localization of low-frequency sound waves originating from atmospheric nuclear explosions. The infrasound network will consist of 60 array stations, distributed as evenly as possible over the globe to assure at least two-station detection capability for 1-kton explosions at any point on earth. This network will be larger and more sensitive than any other previously operated infrasound network. As of today, 85% of the site surveys for IMS infrasound stations have been completed, 25% of the stations have been installed, and 8% of the installations have been certified and are transmitting high-quality continuous data to the International Data Center in Vienna. By the end of 2002, 20% of the infrasound network is expected to be certified and operating in post-certification mode. This presentation will discuss the current status and progress made in the site survey, installation, and certification programs for IMS infrasound stations. A review will be presented of the challenges and difficulties encountered in these programs, together with practical solutions to these problems.

11:30

4aPAb5. Post-installation activities in the Comprehensive Nuclear Test Ban Treaty (CTBT) International Monitoring System (IMS) infrasound network. J. A. Vivas Veloso, D. R. Christie, T. L. Hoffmann, P. Campus, M. Bell, A. Langlois, P. Martysevich, E. Demirovik, J. Carvalho, A. Kramer, and A. Molero Alonso (CTBTO, VIC, Wagramer Str. 5, 1400 Vienna, Austria)

The provisional operation and maintenance of IMS infrasound stations after installation and subsequent certification has the objective to prepare the infrasound network for entry into force of the Comprehensive Nuclear-Test-Ban Treaty (CTBT). The goal is to maintain and fine tune the technical capabilities of the network, to repair faulty equipment, and to ensure that stations continue to meet the minimum specifications through evaluation of data quality and station recalibration. Due to the globally dispersed nature of the network, this program constitutes a significant undertaking that requires careful consideration of possible logistic approaches and their financial implications. Currently, 11 of the 60 IMS infrasound stations are transmitting data in the post-installation Testing & Evaluation mode. Another 5 stations are under provisional operation and are maintained in post-certification mode. It is expected that 20% of the infrasound network will be certified by the end of 2002. This presentation will focus on the different phases of post-installation activities of the IMS infrasound program and the logistical challenges to be tackled to ensure a cost-efficient management of the network. Specific topics will include Testing & Evaluation and Certification of Infrasound Stations, as well as Configuration Management and Network Sustainment.

THURSDAY MORNING, 5 DECEMBER 2002 CORAL GALLERY 1, 8:00 TO 11:20 A.M.

Session 4aSA

Structural Acoustics and Vibration: Reconstruction of Acoustic Radiation from Vibrating Structures I

Sean F. Wu, Chair
Department of Mechanical Engineering, Wayne State University, 505 Anthony Wayne Drive, Detroit, Michigan 48202

Chair's Introduction—8:00

Invited Papers

8:05

4aSA1. Patch nearfield acoustical holography. Earl G. Williams (Code 7137, Naval Res. Lab., Washington, DC 20375, williams@pa.nrl.navy.mil)

Nearfield acoustical holography (NAH) requires the measurement of the pressure field over a complete surface in order to recover the normal velocity on a nearby vibrating concentric surface. Patch NAH provides a major advancement by eliminating the need for complete surface pressure scans—only a small area needs to be scanned to determine the normal velocity on the corresponding (small area) concentric patch on the vibrator. The theory of patch NAH is based on: (1) an analytic continuation of the patch pressure which provides a spatially tapered aperture extension of the field; and (2) a decomposition of the transfer function (pressure to velocity and/or pressure to pressure) between the two surfaces using the singular value decomposition for general shapes and the fast Fourier transform for planar surfaces. Inversion of the transfer function is stabilized using Tikhonov regularization and the Morozov discrepancy principle. Experimental results show that the normal velocity reconstruction errors (rms) for a point-driven vibrator over 200–2700 Hz average less than 20% for two surfaces 0.4 cm apart. Reconstruction of the active normal acoustic intensity was also achieved, with less than 30% rms error over the frequency band. [Work supported by ONR.]

8:35

4aSA2. Reconstructing acoustic radiation using eigenfunction expansions. J. D. Maynard (Phys. Dept., Penn State Univ., University Park, PA 16802)

A powerful tool for reconstructing the sound radiation from a vibrating structure is near-field acoustic holography. In this method, the sound pressure is measured at points near the vibrating structure, and these data are analyzed by computer to determine any desired feature of the entire sound field, including acoustic energy sources at the vibrating structure. Some versions of this method have involved first solving the ''forward problem,'' which consists of finding the radiated sound field, assuming that the surface velocity of the vibrating structure is known. Usually such problems are solved using acoustic boundary elements. However, we have been studying a method for solving the reconstruction problem and by-passing the forward problem by using eigenfunction expansions. Furthermore, Sean Wu has developed a significant extension of the technique for solving the forward problem, and this method may challenge the boundary element method. In this talk, the fundamentals of the eigenfunction expansion method will be reviewed, and results of some practical tests will be presented.

9:05

4aSA3. Nearfield acoustic holography and its evolution. Sean Wu (Dept. of Mech. Eng., Wayne State Univ., 5050 Anthony Wayne Dr., Detroit, MI 48202)

This paper reviews the evolution of nearfield acoustic holography (NAH) over the past two decades. The original NAH (Williams and Maynard, 1980) could only reconstruct acoustic radiation from a surface containing a level of constant coordinate in an exterior region. Soon a generalized NAH based on the Helmholtz integral theory and boundary element method (BEM) was developed to reconstruct acoustic radiation from arbitrarily shaped surfaces. This BEM-based NAH is advantageous in many aspects, but has several inherent shortcomings. An alternative is to expand the acoustic field in terms of the spherical wave functions, the most prominent one being the Helmholtz equation–least squares (HELS) method. This HELS method offers a great flexibility and versatility in reconstructing the acoustic fields in both exterior and interior regions, and allows reconstruction of transient acoustic radiation from impulsively accelerated objects. Meanwhile, we have seen the extension of NAH to characterize acoustic sources of a moving object and variations of NAH based on measurements of acoustic intensity rather than pressures. Many techniques are developed along the way to tackle the ill-posedness difficulties inherent in this inverse acoustic problem. Progression of able researchers is transforming NAH into an ever more powerful diagnostic tool. [Work supported by NSF.]

Contributed Papers

9:35

4aSA4. On the validity and convergence of the Helmholtz equation least squares solutions. Tatiana Semenova and Sean F. Wu (Dept. of Mech. Eng., Wayne State Univ., Detroit, MI 48202)

The problem of reconstruction of acoustic radiation from arbitrary shaped vibrating bodies is often encountered in engineering applications. The Helmholtz equation least-squares (HELS) method proposed by Wu (2000) is an assumed-form solution method in which the reconstructed acoustic pressure is found as a superposition of spherical waves. The unknown coefficients are determined by matching the assumed-form solution to the measured acoustic pressures at a number of points around the source and their errors are minimized by the least-squares method. To further minimize the errors of reconstruction, constraints are used on the surface of the object. In this paper a two-dimensional, noncircular object is considered and the two-dimensional Helmholtz equation is solved. The locations of singularities in the analytic continuation of the solution across the source surface are found using the Schwarz function of the source surface. It will be shown that the HELS solution remains bounded and close to the true pressures inside the minimum circle enclosing the object as well as outside. Other types of modifications of the HELS method for reconstruction of acoustic radiation in an exterior two-dimensional region will be considered. [Work supported by NSF.]

9:50

4aSA5. Acoustic imaging of harmonic near-field sources from surface pressure measurements on a body using singular value decomposition. Peter R. Stepanishen (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882) and Irsan Brodjonegoro (Bandung Inst. of Technol., Indonesia)

An inverse or backward projection method based on a combined Greens function and singular value decomposition method is developed to locate and to determine the strength of near field harmonic sources from the acoustic field on the surface of a nearby rigid body. A resolution matrix, which is based on the free space Greens function, the geometry of the measurement surface of interest and the field locations of interest, is introduced to determine the resolution and accuracy of the backward projection method. Point source distributions located above a rigid planar surface are addressed both analytically and numerically. In addition, line source distributions located outside of infinite rigid circular and elliptical cylinders are also addressed. It is demonstrated that the method is able to identify the location and to determine the strength of harmonic near-field sources which are separated by less than 0.01 of a wavelength.

10:05–10:20 Break

10:20

4aSA6. The reconstruction of the acoustical field over nonconformal surfaces. Nicolas Valdivia and Earl G. Williams (Naval Res. Lab., Code 7130, 4555 Overlook Ave., SW, Washington, DC 20375-5320, valdivia@pa.nrl.navy.mil)

Consider the inverse problem of reconstructing the acoustical field over a surface S_v by measuring the acoustical pressure over an interior surface S_p. Ideally the measurement surface S_p is conformal to the reconstruction surface S_v, and close to S_v. In this work the reconstruction surface S_v will be a box, and the interior surface S_p will be a small sphere inside the box. We study the solution to this inverse problem with two different methods: Near-field Acoustical Holography (NAH) and Helmholtz Equation Least Squares (HELS). We will compare the reconstruction of the two methods using two test solutions. The first test solution will be a point source, and the second test solution will be a plate vibrator in one side of the box. In addition we study the inclusion of white noise to the test solution and regularization methods for its recovery. [Work supported by ONR.]

10:35

4aSA7. Health monitoring using Lamb waves. Application to the characterization of low impacts in aerospace structures. Sebastien Grondel, Jamal Assaad, and Christophe Delebarre (IEMN, DOAE, UMR CNRS 9929, Universite de Valenciennes et du Hainaut Cambresis, Le Mont Houy, 59313 Valenciennes Cedex 9, France)

This work was devoted to the development of a health monitoring system assigned to aerospace applications. The application concerned the detection of low damaging impacts on composite structures due to the extreme sensitivity of this material to this kind of solicitation. The chosen health monitoring was based first on the excitation and reception of Lamb waves along the structure by using thin piezoelectric transducers (active monitoring) and second on a continuous monitoring with the same transducers (passive monitoring). Preliminary tests were performed on a composite plate and the sensitivity of the $A0$ mode to the damage was demonstrated. Concerning the passive monitoring, the study showed the ability to use the acoustic signature of the impact to detect a possible damage. Further tests were performed on a wingbox composite structure. This wingbox structure consisted of composite skins with variable thickness mounted onto a metallic substructure. Moreover, four stringers were bonded on these skins. The aim was to demonstrate the ability of the system to perform rapid inspections of complex structures. After the application of a serial of impacts at diffent locations of the composite wingbox, the feasibility of the health monitoring was demonstrated.

10:50

4aSA8. Reconstruction of transient behavior of a nonuniformly vibrating transducer. Oleg A. Sapozhnikov (Dept. of Acoust., Faculty of Phys., M. V. Lomonosov Moscow State Univ., Leninskie Gory, Moscow 119992, Russia)

An accurate prediction of a transient acoustic field radiated from piezoceramic transducers is important for many applications. The field can be calculated if the normal velocity distribution along the transducer surface is known. However, up to now, there are no reliable direct methods of surface vibration measurement in liquids. The well-developed laser vibrometer method can be employed only when the transducer is in contact with vacuum or gas because of strong acousto-optic interaction in condense media. In this talk, a holographic method is proposed to reconstruct the radiator vibration. The method includes a measurement of the temporal waveform in different points of a plane grid perpendicular to the acoustic axis at some distance from the source, theoretical time reversal of the waveform in each grid-point, and back propagation of the field to the source using the transient form of the Rayleigh integral. The reconstruction procedure is described and accuracy of the method is studied depending on grid-step, distance to the source, and dimensions of the grid. [Work supported by CRDF RP2-2384-MO-02 and RFBR 02-02-17029.]

11:05

4aSA9. Ecological prognosis near intensive acoustic sources. Stanislav A. Kostarev (Lab. of Acoust. and Vib. Tunnel Assoc., 21 Sadovo-Spasskaya Str., Moscow 107217, Russia), Sergey A. Makhortykh (Russian Acad. of Sci., Pushchino, Moscow Region 142290, Russia), and Samuil A. Rybak (N. N. Andreev Acoust. Inst., Moscow 117036, Russia)

The problem of a wave-field excitation in a ground from a quasiperiodic source, placed on the ground surface or on some depth in soil is investigated. The ecological situation in this case in many respects is determined by quality of the raised vibrations and noise forecast. In the present work the distributed source is modeled by the set of statistically linked compact sources on the surface or in the ground. Changes of parameters of the media along an axis and horizontal heterogeneity of environment are taken into account. Both analytical and numerical approaches are developed. The latter are included in the software package VibraCalc, allowing to calculate distribution of the elastic waves field in a ground from quasilinear sources. Accurate evaluation of vibration levels in buildings from high-intensity underground sources is fulfilled by modeling of the wave propagation in dissipative inhomogeneous elastic media. The model takes into account both bulk (longitudinal and shear) and surface Rayleigh waves. For the verification of the used approach a series of measurements was carried out near the experimental part of monorail road designed in Moscow. Both calculation and measurement results are presented in the paper.

THURSDAY MORNING, 5 DECEMBER 2002 GRAND CORAL 3, 8:00 TO 11:30 A.M.

Session 4aSC

Speech Communication: Speech Perception in Normal and Impaired Hearing (Poster Session)

Catherine L. Rogers, Cochair
Communication Science and Disorders, University of South Florida, 4202 East Fowler Avenue, Tampa, Florida 33620-8150

Paolo Escudero, Cochair
School of Communication Science and Disorders, McGill University, 1266 Pine Avenue, West, Montreal, Quebec H3G 1A4, Canada

Contributed Papers

All posters will be on display from 8:00 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 9:45 a.m. and contributors of even-numbered papers will be at their posters from 9:45 a.m. to 11:30 a.m.

4aSC1. Consonant contrasts in a multiple, simultaneous word identification task. Mark Ericson (Air Force Res. Lab., 2610 Seventh St., Wright–Patterson AFB, OH 45433) and Pamela Mishler (Dept. of Veteran Affairs, Dayton, OH)

People often have difficulty communicating in the presence of concurrent conversations. This well-known ''cocktail party effect'' can be described as a combination of energetic and informational masking. The purpose of this study was to measure and model the effects of energetic and informational masking that occur when two people speak at the same time. The word identification test used in the experiments was the Modified Rhyme Test (MRT), which afforded a multitude of initial and final consonant pairs when phrases were played simultaneously. The MRT was used as both the stimulus and the masker, which were presented monaurally at 75 dB SPL over Sennheiser HD-520 headphones to four normal hearing listeners. The independent variables included 30 pairs of MRT word lists, and twelve pairs of same-sex talkers. The dependent variable

was speech (consonant) intelligibility as measured with the MRT. Results will be discussed in the context of place of articulation, manner of articulation, signal-to-noise-ratio of the phoneme pairs, and informational masking model predictions.

4aSC2. Realistic listening improved by adding fine structure. Ginger S. Stickney, Kaibao Nie, and Fan-Gang Zeng (Dept. of Otolaryngol., 364 Med. Surg. II, Irvine, CA 92697)

Speech communication often takes place in noisy and reverberant environments. Unfortunately, cochlear implant (CI) users perform significantly poorer than normal hearing listeners under adverse conditions. We hypothesize that the noise susceptibility of CI users is largely due to the lack of encoding the fine structure cue. Most CI speech processing algorithms extract and encode the slowly varying amplitude modulation in speech (the envelope), while the rapidly varying fine structure is replaced with a fixed rate carrier. We have developed a new speech processing algorithm, called Frequency-Amplitude-Modulation-Encoding (FAME), which codes both amplitude and frequency modulations. Normal hearing listeners were presented with IEEE sentences in the presence of an interfering voice as a function of the number of noise bands (information channels) and signal-to-noise level. Sentence recognition with a competing talker was generally improved with the addition of fine structure, and by as much as 60% with 4 frequency bands (i.e., from 20%–80% at a 20 dB signal-to-noise level). This result suggests that the fine structure is important for speech recognition under realistic situations and should be encoded in future cochlear implants.

4aSC3. Fricative spectral moments and the perception of anticipatory coarticulation. William F. Katz, Sneha V. Bharadwaj, Monica P. Stettler, and Peter F. Assmann (School of Human Development and the Callier Ctr. for Commun. Disord., Univ. of Texas, Dallas, P.O. Box 830688, Richardson, TX 75083, wkatz@utdallas.edu)

Spectral moments (mean, variance, skewness, kurtosis) have been used to classify fricative place of articulation in English and other languages. For example, studies have shown the first spectral moment (centroid) is consistently lower in /ʃ/ than in /s/. It has also been hypothesized that fricative spectral moments contribute to listeners' judgments of vowel identity in fricative-vowel (FV) syllables, providing acoustic cues for the perception of anticipatory coarticulation. To study the relationship between fricative spectral moments and traditional fricative spectral attributes as cues to the perception of coarticulation, we obtained a database of FV syllables (/si/, /su/, /ʃi/, /ʃu/) produced by adults and children (ages 5 and 7). Acoustic analyses examined spectral and temporal properties of the signal, while a perceptual experiment determined how adult listeners use gated "slices" ($\frac{1}{2}$ fricative, $\frac{3}{4}$ fricative, full fricative, fricative plus $\frac{1}{2}$ vowel) to identify the fricative and upcoming vowel in the FV syllables. The identification of full syllables showed effects of gate length and talker group, while the vowel identification data revealed little difference as a function of either variable. Logistic regression is used to test the independence of fricative and vowel cues in syllable perception, and to model developmental trends.

4aSC4. English-learning infants use juncture cues to segment speech. Elizabeth K. Johnson (Dept. of Psych., Johns Hopkins Univ., Baltimore, MD 21218)

Seven-month-olds perceive recurring sequences of strong–weak (SW) syllables as words, presumably due to their heavy reliance on the metrical segmentation strategy for English. For example, 7-month-olds will segment "taris" from a passage containing repeated tokens of "guitar is." However, by 10 months, infants no longer make this mistake [Jusczyk *et al.*, Cognit. Psychol. **39**, 630–645 (1999)]. There are many possible explanations for this result. Some have argued that 10-month-olds are sensitive to the subtle juncture cues suggesting a word boundary between syllables such as "tar" and "is." However, others have suggested that 10-month-olds have adopted a lexical segmentation strategy. In other words, infants know common function and auxiliaries such as "is." Therefore, they know that "taris" cannot be a word. The head-turn preference procedure was used to investigate whether English-learning infants would segment "toga" from "toe galore" just as readily as they would segment "toga" from "toga lore." Seven-month-olds segmented "toga" equally well in both cases, whereas 12-month-olds failed to segment "toga" from "toe galore." These results suggest that knowledge of frequently occurring function and auxiliaries cannot explain Jusczyk *et al.*'s results. Future work will investigate the information infants might use to distinguish "toe galore" from "toga lore."

4aSC5. Perception of different speech registers in noise. Jean Andruski and Eileen Bessega (Wayne State Univ., 581 Manoogian Hall, Detroit, MI 48202)

Infant-directed speech and adult-directed clear speech both exhibit acoustic characteristics (other than amplitude differences) which should make them easier to perceive against a background of noise than adult-directed conversational speech. For example, in adult-directed clear speech, vowels are typically acoustically more distinct. Infant-directed speech shows similar differences in vowel quality, and in addition typically shows larger pitch contours. This experiment examines listeners' perception of these three speech registers (adult-directed conversational speech, adult-directed clear speech, and infant-directed speech) in noise. Effect of the different speech registers will be compared for easy words (frequent words with few frequent lexical neighbors) and hard words (infrequent words with numerous, frequent neighbors).

4aSC6. Auditory and visual clear speech effects measured during a simulated conversational intercation. Jean-Pierre Gagne, Monique Charest, Ariane Laplante-Levesque, and Olivia Guilbert (Orthophonie et Audiologie, Univ. of Montreal, C.P. 6128, Succursale Centreville, Montreal, QC H3C 3J7, Canada)

Iterations of sentences were recorded audio-visually from talkers while they participated in a speech-tracking task. Six female talkers produced iterations of conversational and clear speech under two different experimental conditions: (a) while the talker was informed that only her visual–speech cues would be transmitted to the interlocutor and (b) while she was informed that only her auditory–speech cues would be transmitted to the interlocuter. In reality, both her auditory– and her visual–speech cues were recorded under each experimental condition. Target sentences were extracted for the recordings, edited, and presented in a random order to a group of 48 subjects. The subjects completed a speech-recognition task under two perceptual modalities: auditory-only and visual-only. The subjects' mean speech-recognition scores were used to determine the speech intelligibility scores of individual talkers for each experimental condition. The results failed to reveal any differences between the speech intelligibility scores obtained while a talker intended to produce iterations of visual-clear speech and those obtained while she intended to produce iterations of auditory-clear speech. Hence, the findings failed to demonstrate that talkers modify their articulation patterns in order to compensate for the perceptual modality under which the interlocutor receives the speech information. [Work supported by a NSERC grant awarded to J-PG.]

4aSC7. Visual enhancement for consonants, words, and sentences in normal-hearing young and older adults. Mitchell S. Sommers (Dept. of Psych., Campus Box 1125, Washington Univ., St. Louis, MO 63130), Nancy Tye-Murray, and Brent Spehar (Central Inst. for the Deaf, St. Louis, MO 63110)

Visual enhancement in speech perception refers to the benefit obtained from seeing and hearing a talker, compared with listening alone. Previous studies comparing visual enhancement in younger (under age 25) and older (over age 65) adults have been confounded by age-related hearing loss. Therefore, the present study was designed to compare visual enhancement for consonants, words, and sentences in normal-hearing older and younger adults. All participants had clinically normal audiograms for octave frequencies from 250 Hz to 4 kHz. Consonant identification was tested using the Iowa Consonant Confusion Test. Word identification was measured using the Children's Auditory–Visual Enhancement Test and sentence identification was measured using the Iowa Sentence Test. All testing was conducted in the presence of a multitalker background babble and baseline differences in auditory-only performance were minimized by setting signal-to-babble ratios individually to produce approximately 50% correct identification inthe auditory-only conditions. Visual enhancement, computed as the relative benefit obtained from adding visual speech information to the auditory signal, was significantly higher for younger than for older adults for consonants and words, but not for sentences. The results are discussed as supporting an age-related decline in intermodal sensory integration.

4aSC8. Examination of the relevance of cochlear dead regions to hearing aid fittings. Paul D. Dybala and Linda M. Thibodeau (Adv. Hearing Res. Ctr., Univ. of Texas at Dallas, 1966 Inwood Rd., Dallas, TX 75235, dybala@utdallas.edu)

Vickers *et al.* [J. Acoust. Soc. Am. **110**, 1164–1175 (2001)] suggested that identification of Cochlear Dead Regions (CoDR) may provide information relative to hearing aid fittings. The relationship between CoDR and hearing aid fittings was explored in the current study two ways: (1) by reviewing the hearing loss characteristics of Vickers *et al.* (2001) subjects with CoDR using the Articulation Index, and (2) examining the frequency of CoDR in a clinical population. Using amplified, low-pass filtered nonsense syllables Vickers *et al.* (2001) found a relationship between the edge of the identified high-frequency CoDR and the filter setting that provided optimal syllable recognition. Because of the severity of the high-frequency hearing loss for the CoDR group the Articulation Index results from the present study suggested that these listeners may have not benefited from high-frequency amplification because of inaudibility. The frequency of CoDR was determined in a clinical population using the Threshold Equalization Noise (TEN) test described in Moore *et al.* [Br. J. Audiol. **34**, 205–224 (2000)]. Results suggested that CoDR are not encountered as frequently as implied in the literature. These findings have important implications on how CoDR testing relates to hearing aid fittings.

4aSC9. Application of Plomp's model to recognition of the Hearing in Noise Test sentences. Linda M. Thibodeau and Jack M. Scott III (Adv. Hearing Res. Ctr., Univ. of Texas at Dallas, 1966 Inwood Rd., Dallas, TX 75235, thib@utdallas.edu)

Sensitive measures of speech recognition are needed for adequate characterization of the effects of various types and degrees of hearing loss. The Hearing in Noise Test (HINT), by Nilsson *et al.* [J. Acoust. Soc. Am. **95**, 1085–1099 (1994)], was developed to quantify the benefits of binaural amplification by measuring a reception threshold for sentences (RTS) in the soundfield in quiet and various noise arrangements. The purpose of the present study was to expand the use of the HINT as a tool to characterize effects of hearing loss in monotic listening conditions. Modeled after Plomp's [J. Acoust. Soc. Am. **63**, 533–549 (1978)] adaptive, speech-recognition paradigm with Dutch sentences, the HINT sentences were presented in quiet and in increasing noise levels via insert receivers to listeners with normal and mild-to-severe hearing impairments. Following Plomp's analysis, two variables were derived for each listener with hearing loss from the function relating RTS to noise level: A and D, which represent attenuation and distortion effects, respectively. As expected, the pure-tone average was more strongly correlated with A than with D. Derivation of these two variables was also explored using an abbreviated procedure of only two RTS values, one in quiet and one in noise.

4aSC10. Effects on performance of partial misalignments of spectral information in acoustic simulations of cochlear implants. Anthony Spahr, Michael Dorman (Arizona Biomed. Inst., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-0102), and Philipos Loizou (Univ. of Texas, Dallas, Richardson, TX 75083-0688)

Objective: To assess the effects on speech understanding of frequency misalignments in channel outputs for an acoustic model of a cochlear implant. Design: Consonants, vowels, and sentences were processed through a four-channel sine-wave simulation of a cochlear implant. In the five experimental conditions, channels 1 and 3 were always output at the correct frequency while channels 2 and 4 were output at frequencies varying from the correct frequency to frequencies 25%, 50%, and 75% lower than appropriate. In a fifth condition (2-of-4 condition), channels 2 and 4 were turned off. Results: Consonant recognition was reduced significantly with a 50% shift in channels 2 and 4. Vowel and sentence recognition were reduced significantly with a 75% shift in channels 2 and 4. For all material performance in the 75% shift condition was better than in the 2-of-4 condition. Conclusions: The perceiving system that underlies speech recognition is relatively tolerant of misplaced frequency information if other information is presented to the correct frequency location. This flexibility may account for some of the success of cochlear implants. [Research supported by a grant from NIDCD (No. RO1 00654-12) to the second author.]

4aSC11. The effects of completely-in-the-canal hearing aids, active noise reduction headsets, and ambient noise on speech intelligibility. John Hall, Mark Ericson (Air Force Res. Lab., Wright–Patterson AFB, OH 45433), and Pamela Mishler (Dept. of Veterans Affairs, Dayton, OH)

Many people with hearing loss need to communicate and work in high noise environments. Hearing aids and headsets are often worn to protect the ear from further damage and facilitate communication. The purpose of this study was to determine the effects of the active noise reduction (ANR) headsets and Completely-in-the-Canal (CIC) hearing aids on speech intelligibility in ambient noise. Six male listeners with noise-induced hearing loss were each fitted with two digital CIC hearing aids. The listeners wore the hearing aids successfully for one year. The speech stimuli were presented over the ANR headphones in quiet and in 95 dB and 105 dB SPL of ambient pink noise. The independent variables were the active noise reduction, hearing aids, and noise level. All MRT stimuli were presented at the most comfortable loudness level. The dependent variable was speech intelligibility as measured with the Modified Rhyme Test (MRT). A main effect was discovered for noise, however no effect was found for active noise reduction or hearing aids. Although no overall effect was found with ANR or hearing aids on speech intelligibility there were differences across phoneme identifications and initial and final consonants.

4aSC12. Perception of Mandarin Chinese tones by Mandarin versus French listeners. Pierre A. Hallé (CNRS-Paris V, 71 Av. Edouard Vaillant, 92774 Boulogne-Billancourt, France, halle@psycho.univ-paris5.fr), Catherine T. Best (Wesleyan Univ., Middletown, CT 06459 and Haskins Labs., New Haven, CT 06511), and Yueh-Chin Chang (Tsing-Hua Univ., Taiwan)

Few reports exist on perception of lexical tones by native listeners (e.g., Mandarin Chinese) versus non-native listeners whose language lacks tone contrasts. In those few, the non-native language was English, in which pitch differences signal stress alternations. No studies have examined perception of minimal tone contrasts, of central concern to theories of nonnative speech perception [e.g., PAM: Best (1995)]. We examined categorical perception of Mandarin tone contrasts, in order to detect fine-grained performance differences [see Hall *et al.* (1999)] between listeners of Mandarin and of French, a nonstress language. Three tone continua were derived from natural Mandarin target utterances within carrier sentences. We interpolated on both $F0$ and intensity contours. First, Mandarin listeners identified the tone of target syllables within carrier sentences. Next, both French and Mandarin listeners completed A×B identification and discrimination tests on isolated syllables. Mandarin, but not French listener perceived tones quasicategorically. French listeners showed substantial sensitivity to tone contour differences, although to a lesser extent than Mandarin listeners. Thus, despite the lack of linguistically relevant pitch contrasts in their language, French listeners are not ''deaf'' to tonal variations. They simply fail to perceive tones along the lines of a well-defined and finite set of linguistic categories.

4aSC13. Native-language influence on phonetic perception in Dutch–English bilinguals. Michele Mondini (Northeastern Univ., Boston, MA 02115), Petra M. van Alphen (Max Planck Inst. for Psycholinguist., Nijmegen, The Netherlands), and Joanne L. Miller (Northeastern Univ., Boston, MA 02115)

We examined how native-language experience influences processing a second language, focusing on how native Dutch listeners who learned English as a second language perceive the English voiceless consonant /p/. Previous research [J. E. Flege and W. Eefting, Speech Commun. **6**, 185–202 (1987)] shows that the voiced–voiceless boundary for an (English-based) voice-onset-time (VOT) series is located at a shorter VOT for such bilingual listeners than for native English listeners, consistent with the fact that voiceless stops are produced with shorter VOTs in Dutch than in English. We asked whether such bilinguals also differ from native English listeners in which stimuli throughout the series are perceived as reasonable exemplars of /p/. Native English listeners and native Dutch listeners were tested on a three-choice identification task with an (English-based) extended VOT series that ranged from /ba/ to /pa/ to an ''unnatural'' exaggerated /pa/, labeled */pa/. Both the /b/-/p/ and /p/-*/p/ boundaries were located at shorter VOTs for the native Dutch than the native English listeners, indicating that Dutch native-language experience influenced the entire range of VOTs perceived as reasonable exemplars of the /p/ category. Thus native-language experience has a comprehensive influence on the mapping from acoustic signal to phonetic category. [Work supported by NIH/NIDCD.]

4aSC14. An evaluation of the semantic categorization task in Japanese word recognition. Kiyoko Yoneyama (Natl. Inst. for Japanese Lang., 3-9-14, Nishigaoka, Kita-ku, Tokyo 115-8620, Japan, yoneyama@kokken.go.jp)

In this paper an evaluation of the semantic categorization task in Japanese word recognition is reported. Data used in this experiment were taken from Yoneyama (2002) in which fillers were the words for gyesh responses for semantic categories. In this experiment, participants had to deal with 28 multiple semantic categories. Two findings were concluded. First, although participants generally showed a good performance, the degree of task difficulty among 28 semantic categories was not equal. The average percent correct for gyesh filler-word responses by participants was 95%. The mean values of the words responded to correctly within each category and the number of words correctly classified by at least 93% of the listeners both indicate such tendencies. Second, although the degree of task difficulty among semantic categories within the experiment varied, the selection of semantic categories by itself seemed to be reasonable. The lowest number of words correctly classified by at least 93% of the participants was 17 out of 25 filler words for 4 categories, suggesting that at least 17 of the filler words were considered by general undergraduate students to be good instances of these semantic categories. A generalized methodology to select words in semantic categories will be also discussed. [Work supported by NIH (PI: Keith Johnson).]

4aSC15. Identification of English /r/ and /l/ in white noise by native and non-native listeners. Kazuo Ueda, Noriko Yamasaki (Unit of Auditory and Visual Commun. Sci., Kyushu Inst. of Design, 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan, ueda@kyushu-id.ac.jp), and Ryo Komaki (ATR Human Information Sci. Labs., Keihanna Science City, Kyoto 619-0288, Japan)

Fifty-three English word pairs, minimally contrasting in /r/–/l/, were identified by 14 native American English (AE) listeners and 42 native Japanese (J) listeners, under systematically varied signal-to-noise ratios (SNR) which ranged from no-noise to 21 dB. Word identification for filler contrasts (FL: /d/–/k/, /m/–/n/, etc.) was also tested under the same SNR conditions. Results show that: (a) identification accuracy by the AE listeners was nearly 100% in the no-noise condition compared to about 70% with a SNR of 21 dB; (b) the J listeners showed similar identification accuracy to the AE listeners for the FL words, whereas they performed much poorer on the /r/–/l/ contrasted words (70% accuracy without noise and 55% accuracy with a SNR of 21 dB); and (c) /r/–/l/ identification accuracy was affected by an interaction between the consonant position in a word and the listeners' native languages. The role of perceptual assimilation in the interaction will be discussed. [Work supported by TAO and JSPS, Japan.]

4aSC16. Influence which the sound as a noise has on maintenance of phoneme information. Takashi Yagyu (Dept. of Psych., Univ. of Tokyo, 7-3-1 Hongo, Bukyo-ku, Tokyo 113-0033, Japan, yagyu@l.u-tokyo.ac.jp)

This research is investigated from the viewpoint of cross-language processing about the influence which the noise generated by reading aloud has to the maintenance results of phoneme information. If a sentence is read aloud simultaneously when holding phoneme information, the quantity of the phoneme information which can be held and set will decrease. This phenomenon can check the rehearsal of the phoneme information which the noise produced by reading a sentence aloud holds, and it is also possible that maintenance results fall as a result. Then, in this research, the cognitive psychological experiment was conducted using the dual-task method of reading aloud of a sentence, and memory of a word (phoneme information). Two conditions of a native language and a foreign language were set up as a noise generated by reading aloud, and it examined to what noise memory of phoneme information would decline from the point of the individual difference of phonological short-term memory (phonological STM) capacity. Consequently, the more phonological STM capacity was large, the more it was suggested that the stoutness of the phoneme memory to a similar noise increases.

4aSC17. Auditory-visual integration: A comparison of neuropathy and developmental dyslexia. Joshua C. Ramirez and Virginia Mann (Dept. of Cognit. Sci., Univ. of California, Irvine, Irvine, CA 92612, ramirezj@uci.edu)

Auditory neuropathy and developmental dyslexia have been previously associated with inferior perception of acoustic speech in noise. To probe the underlying mechanisms of these deficits, an audio-visual integration experiment was conducted to compare the speech perception abilities of reading-impaired subjects and auditory neuropathy patients (relative to controls). Natural acoustic speech stimuli that were white-noise masked at various intensities and accompanied by a video of the talker were presented to subjects. Neuropathy patients were expected to benefit from the pairing of visually presented articulatory cues and auditory CV stimuli, owing to the fact that their speech perception impairment reflects a peripheral auditory disorder. To the extent that developmental dyslexia reflects a general impairment of speech processing rather than a peripheral disorder of audition, it was hypothesized that reading-impaired subjects would not benefit from an introduction of visual cues. Current experimental data yielded a significant correlation between reading ability and dependence upon visual cues. Furthermore, under conditions where articulatory cues are matched with heavily masked CV stimuli, neuropathy patients and controls made more use of the visual cues than reading-impaired subjects. This may suggest that some forms of reading impairment stem from deficits of general speeach processing.

4aSC18. Audio-visual integration of speech with time-varying sine wave speech replicas. Jyrki Tuomainen, Tobias Andersen, Kaisa Tiippana, and Mikko Sams (Cognit. Sci. and Technol., Lab. of Computational Eng., P.O. Box 9203, Helsinki Univ. of Technol., Finland)

We tested whether listener's knowledge about the nature of the auditory stimuli had an effect on audio-visual (AV) integration of speech. First, subjects were taught to categorize two sine-wave (sw) replicas of the real speech tokens /omso/ and /onso/ into two arbitrary nonspeech categories without knowledge of the speech-like nature of the sounds. A test with congruent and incongruent AV-stimulus condition (together with auditory-only presentations of the sw stimuli) demonstrated no AV integration, but instead close to perfect categorization of stimuli in the two arbitrary categories according to the auditory presentation channel. Then, the same subjects (of which most were still under the impression that the sw-stimuli were nonspeech sounds) were taught to categorize the sw stimuli as /omso/ and /onso/, and again tested with the same AV stimuli as used in the nonspeech sw condition. This time, subjects showed highly reliable AV integration similar to integration obtained with real speech stimuli in a separate test. We suggest that AV integration only occurs when subject are in a so-called ''speech mode.''

4aSC19. Effects of 3-D projection on audiovisual speech perception. D. H. Whalen (Haskins Labs, 270 Crown St., New Haven, CT 06511), Richard Gans (L.I.P.S., Woodbridge, CT 06525), Carol A. Fowler, and Julie M. Brown (Haskins Labs, New Haven, CT 06511)

Visual information about speech influences speech perception, leading to better perception in noise and to illusions such as the McGurk effect. Here, the question was addressed of whether visual influences would be greater with a three-dimensional visual speaker [the patented Life Imaging Projection System (L.I.P.S.)] than the two-dimensional one. Perception in noise was tested with 54 monosyllabic English words of moderate frequency. The McGurk effect was tested with acoustic words having bilabial onset consonants and visual nonwords with alveolar onsets. Visual information improved perception in noise, as found previously; the 3-D version (a videotape projected onto a life cast of the speaker's head) elicited better performance than the 2-D. However, the 3-D version elicited fewer McGurk responses. Since the presence of noise increases visual fixation on the lips [Vatikiotis-Bateson *et al.*, Percept. Psychophys. **60**, 926–940 (1998)], it is possible that more direct information about mouth shape was available in that condition. Further, the 3-D version may have been more likely than the 2-D to engage a direct eye gaze, perhaps reducing the McGurk effect. A test is planned in which noise is added to McGurk stimuli; this may restore the advantage for the 3-D information. [Work supported by NIH grant HD-01994.]

4aSC20. Multiple talker effects in auditory–visual speech perception among older adults. Kathleen Cienkowski (Dept. of COMS, Univ. of Connecticut, 850 Bolton Rd., Storrs, CT 06269)

Talker variability effects have been shown for speech stimuli presented auditorily to young adult and older adult participants. Recent investigations suggest that similar variability is seen among young adults for auditory–visual congruent and discrepant stimuli. Older adults are reported to be at a disadvantage for processing information spoken by multiple talkers. Some investigators have suggested that perceptually demanding conditions, such as listening in noise may be more difficult for older listeners. In the current investigation, talker variability effects for CV syllables were examined in multiple modalities visual-only, auditory-only, and auditory–visual for older listeners. Auditory–visual presentations were of two types: congruent [auditory and visual were of the same token —/visual bi and auditory bi/] or disparate [/visual gi and auditory bi/]. The disparate condition has elicited the McGurk effect, in which listeners may report a fused response that is neither the auditory or the visual component of the stimulus. The talkers were 11 individuals with known equivalent auditory intelligibility. Consistent with previous work, individual older listeners varied considerably in their overall perceptions, from complete fusion for all talkers to almost no fusion for any talker. Further, older listeners demonstrated poorer performance in comparison to younger adults in the multitalker condition.

4aSC21. Audio-visual integration in listeners with normal hearing and hearing aid users. Nicole L. Marrone (Dept. of Commun. Disord., Boston Univ., 635 Commonwealth Ave., Boston, MA 02215), Ellen S. Stockmann, Frank H. Guenther, Jennell C. Vick, Joseph S. Perkell, and Harlan Lane (MIT, Cambridge, MA)

In listeners with normal hearing, the sight of a speaker's face articulating a syllable can influence the auditory percept, most observably when the auditory and visual stimuli are different from one another. This study investigates differences in audio-visual (AV) integration (''the McGurk effect'') between adults with hearing loss who wear hearing aids (HA) and their normal-hearing (NH) counterparts. The following hypothesis is being tested: HA users will rely more on visual input and thus be biased more toward the visual stimulus in the mismatch condition. Audio-visual stimuli from three speakers are presented, pairing the consonants /b/, /d/, and /g/ with the vowels /a/, /i/, and /u/, in three conditions (auditory-only, visual-only, and AV) to the two subject groups, NH and HA. Participants label each stimulus according to the consonant perceived. Responses are coded into four categories: fusion, combination, auditory, or visual. Data analysis examines the relative strength of visual influences in the two groups. Pilot data show fusion and visual bias in an HA user. Further results will be presented. [Work supported by NIH.]

4aSC22. Intelligibility testing of dysarthria in native Spanish speaking adults. Michael Fraas (Dept. of Commun. Sci. and Disord., Univ. of Cincinnati, P.O. Box 670394, Cincinnati, OH 45267-0394)

Intelligibility testing in dysarthria has attempted to assess, not only the severity of the motor speech deficit but also the nature of the disorder. Kent, Weismer, Kent, & Rosenbek (1989) developed a single word test of intelligibility consisting of acoustic/phonetic contrasts sensitive to both dysarthric impairment and its contribution to speech intelligibility. Although effective for assessing dysarthria in neurodegenerative diagnoses such as Parkinson's disease (PD) and amyotrophic lateral sclerosis (ALS), it was constructed to serve an English speaking population. The current investigation will use the Kent *et al.* (1989) study as a model to develop a single word test of intelligibility in native Spanish speakers. Here 30 subjects, 15 with dysarthria due to neurodegenerative disease and 15 normal controls will be used. Subjects will be recorded reading a list of Spanish minimal pair contrasts. In the study we will attempt to describe the speech characteristics associated with dysarthria in Spanish speakers using acoustic-instrumental analysis and comparatively evaluate these characteristics against those of Spanish-speaking normals. Recordings of the responses will be examined to determine the effectivess of the Spanish paired word test of intelligibility, and its sensitivity to error patterns of Spanish-speaking dysarthrics.

4aSC23. Representational specificity of lexical form in the perception of spoken words. Conor T. McLennan, Paul A. Luce (Language Percept. Lab., Univ. at Buffalo, 245 Park Hall, Buffalo, NY 14260, mclennan@acsu.buffalo.edu), and Jan Charles-Luce (Univ. at Buffalo, Buffalo, NY 14214)

In previous research using the long-term repetition-priming paradigm, we demonstrated that flapped intervocalic alveolar stops in American English may be mapped onto underlying representations of /t/ and /d/ during spoken word processing. More specifically, we found that casually articulated, flapped stimuli primed their carefully articulated counterparts (e.g., casually articulated atom primed carefully articulated atom and Adam) and *vice versa*. We argued that form-based or phonological ambiguity (present in our flapped stimuli) is one of the conditions leading to the activation of underlying representations. However, our flapped items were also lexically ambiguous (e.g., our casually articulated atom was ambiguous between Adam and atom). Is the long-term priming effect we observed for flapped and carefully articulated words lexically or sublexically mediated? In the present research we attempt to answer this question by examining form-based ambiguity in the absence of lexical ambiguity. Long-term repetition priming experiments were conducted with nonword stimuli that preserve form-based ambiguity (e.g., oytem, oydem) and flapped stimuli without lexical pairs (e.g., pretty). The results provide further insights into the conditions under which underlying representations are activated during spoken word processing.

4aSC24. Nasal consonant speech errors: Implications for "similarity" and nasal harmony at a distance. Rachel Walker, Narineh Hacopian, and Mariko Taki (Dept. of Linguist., USC, GFS 301, Los Angeles, CA 90089-1693, rwalker@usc.edu)

It has been observed that nasals interact with ''similar'' consonants in phonological long distance nasal harmony [Walker, BLS **26**, 321–332 (2000)]. In such patterns, nasals cause other consonants in a word to become nasal. Voiced stops and liquids are preferentially affected; in some cases obligatorily homorganic with the nasal. In this study we ask whether a parallel occurs in consonants with greater likelihood to participate in speech errors with nasals; specifically, is preferred interaction of voiced and homorganic stops with nasals also evident in the error pattern of a language without nasal harmony? Experiments were conducted on English speakers using the SLIPS technique [Baars and Motley, Catalog of Selected Documents in Psychology, 1974]. CVC word pairs were presented for reading. Critical pairs, primed to encourage initial consonant errors, were cued to be spoken. Experiment 1 (35 subjects) investigated whether more errors involve nasals and voiced stops than nasals and voiceless stops (partially replicates Stemberger's 1991 study). Experiment 2 (37 subjects) examined whether more errors involve nasals and homorganic voiced stops than nasals and heterorganic voiced stops. Results indicate that consonants participating in more errors with nasals parallel those preferentially affected in long distance nasal harmony. [Work supported by a Zumberge grant.]

4aSC25. Accessing coarticulatory information. Suzanne Curtin (Dept. of Linguist., 2806 Cathedral of Learning, Univ. of Pittsburgh, Pittsburgh, PA 15260, scurtin@pitt.edu), Neelam Ladhar, and Janet Werker (Univ. of British Columbia, Vancouver, BC V6T 1Z4, Canada)

The speech signal is comprised of coarticulatory cues. Here, we explore whether adults' access to coarticulatory information depends upon the particular language task. Specifically, we tested adults' sensitivity to coarticulation when segmenting new sequences versus remembering presegmented items. Forty undergraduates were familiarized to either a continuous string of appropriately coarticulated nonsense CV syllables where every third syllable was stressed to facilitate segmentation, or to the same CV syllables presegmented into trisyllabic units without any stress contour. During the test phase, adults were presented with familiar and novel sequences appropriately coarticulated, and to both sequence types with inappropriate coarticulation. Subjects rated the familiarity of items using a seven point scale. Both groups rated familiar test sequences as more familiar than novel sequences. However, only the group presented with presegmented items demonstrated sensitivity to coarticulation. These subjects, compared to the segmentation group, significantly preferred the coarticulated items ($p<0.05$). These results suggest that access to information depends on the task. In the case of word segmentation, adults play attenation to the most useful properties (in this case, stress). Whereas adults in the presegmented group were able to pick up all the details of the sequences since their task only required word recognition.

4aSC26. The perception of gated Dutch diphones. James M. McQueen, Roel Smits, Anne Cutler (Max Planck Inst. for Psycholinguist., Postbus 310, 6500 AH Nijmegen, The Netherlands, james.mcqueen@mpi.nl), and Natasha Warner (Univ. of Arizona, Tucson, AZ 85721-0028)

The results of a large-scale speech perception study are reported. Eighteen Dutch listeners identified gated fragments of 1179 Dutch diphones. Diphones were presented six times (for some diphones beginning with voiceless stops, four times) to each listener, in fragments ranging from the shortest gate (the first sixth of the diphone), through to the longest gate (the complete diphone). Order of presentation of all gates of all diphones was fully randomized. Listeners were asked to identify the complete diphone on every trial. The results showed that listeners based their decisions on the acoustic information available in the stimuli, not on higher-order factors such as phoneme occurrence frequencies or transitional probabilities. Perceptual confusions reflected the temporal evolution of phoneme similarities (for example, in their shorter gates, long vowels were confused with their short counterparts). This database provides detailed information about the perceptual hypotheses which listeners entertain about phoneme sequences, as those sequences unfold over time. It can also be used to generate predictions about the temporal pattern of activation of spoken words. Such data are necessary for the evaluation and further

development of current models of spoken word perception, in which fine-grained acoustic information is mapped continuously onto the mental lexicon.

4aSC27. Temporal integration of acoustic cues in fricative perception. Santiago Fernández and Sergio Feijóo (Dept. de Fisica Aplicada, Fac. de Fisica, Univ. de Santiago, 15782 Santiago de Compostela, Spain)

An important issue in speech perception is to determine how the components of a syllable interact to enhance perception of both consonant and vowel. To date the mechanism underlying that integration has not yet been discovered. Different approaches to the temporal integration between fricative and vowel in a set of natural syllables were explored. Two hypotheses were considered. (a) The two segments are evaluated separately and then combined into a single percept; (b) both cues are evaluated jointly. To test those hypotheses several computational models were considered. If the F and V segments are evaluated separately, two statistical functions are available: An ''OR'' function corresponding to the perceptual hypothesis predicting that only one of the segments determines the identity of the fricative; an ''AND'' function corresponding to the perceptual hypothesis predicting the use of both cues. The hypothesis of the joint evaluation of both cues was tested using the whole FV segment. Their performances were compared with the perceptual performance of a group of listeners in a fricative identification task. Although the joint evaluation model was superior to the other models, it was unable to extract the same benefits as listeners from the fricative–vowel interaction. [Work supported by Xunta de Galicia.]

4aSC28. Learning-induced neural plasticity associated with acquisition of a difficult second-language phonetic contrast. Daniel Callan, Rieko Kubo (Human Information Sci. Labs., ATR Intl., Japan), Akiko Callan, and Shinobu Masaki (Brain Activity Imaging Ctr., ATR Intl., Japan)

Adult native Japanese speakers have difficulty perceiving the English /rl/ phonetic contrast even after years of exposure. However, after extensive perceptual identification training long lasting improvement in identification performance that generalizes to novel stimuli can be attained. In this fMRI study we investigate localized changes in brain activity associated with one-month of extensive feedback-based perceptual identification training by native Japanese speakers learning the English /rl/ phonetic contrast. Before and after training separate functional brain imaging sessions were conducted for identification of the English /rl/ contrast (difficult for Japanese speakers), /bg/ contrast (easy), and /bv/ contrast (difficult). Neural plasticity, denoted by exclusive enhancement in brain activity for the /rl/ contrast (not present for the /bv/ and /bg/ conditions), does not only occur in brain regions involved with acoustic–phonetic processing (superior temporal areas, supramarginal gyrus, planum temporale) but also in additional bilateral cortical (Broca's area, premotor cortex, orosensory cortex) and subcortical regions (cerebellum, basal ganglia, substantia nigra) involved with speech production as well as with formation of perceptual–motor mappings. The results support the hypothesis that learning of an auditory–articulatory mapping improves identification performance by allowing a perception to be made in reference to potential action. [Work supported by TAO, Japan.]

THURSDAY MORNING, 5 DECEMBER 2002 GRAND CORAL 1, 7:50 A.M. TO 12:00 NOON

Session 4aUW

Underwater Acoustics and Acoustical Oceanography: Littoral Environmental Variability and Its Acoustic Effects II

Peter H. Dahl, Cochair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105-6698

James F. Lynch, Cochair
Woods Hole Oceanographic Institution, 203 Bigleow Building, Woods Hole, Massachusetts 02543

Chair's Introduction—7:50

Invited Papers

7:55

4aUW1. Overview of ASIAEX field experiments in the South and East China Seas. James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543) and Peter H. Dahl (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

In the spring of 2001 the Asian Seas International Acoustics Experiment (ASIAEX) was performed in the South and East China Seas. The ASIAEX program originated from the Office of Naval Research's initiative to develop a Sino-American cooperation in the field of ocean acoustics, and expanded to involve Taiwan, Korea, Singapore, and Japan. In the South China Sea, the emphasis was on lower frequency (50–600 Hz) acoustic propagation through complex oceanography along and across the shelfbreak. The oceanographic variability was driven at the mesoscale by the monsoonal wind stress, buoyancy fluxes from the Chinese coast, and by occasional Kuroshio intrusions through the Luzon Strait. Large, nonlinear internal waves also significantly affected the acoustics. In the East China Sea the emphasis was on low- to- midfrequency acoustic interaction with, and reverberation from, the sea floor and sea surface. The experimental site was chosen for flatness, to minimize the influence of bathymetric trends, and emphasize bottom

roughness and subbottom structures, in the measurements. This talk will describe the scope of the ASIAEX experimental program, including the ocean acoustic and environmental characterization of the seafloor, sea surface, and water column. Results from various measurement programs will be described in separate papers.

8:15

4aUW2. Long range sediment tomography. James H. Miller (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882)

The development of long range sediment tomography using shot data and its application in various field experiments is presented. Whereas traditional ocean acoustic tomography has mainly relied on linear inverse theory using the travel time of rays in deep water, our technique relies on nonlinear optimization techniques using the frequency-dependent travel time of normal modes in shallow water. Our experience with broadband sediment tomography using explosive sources started with the Barents Sea Polar Front Experiment where sediment properties were extracted using ground wave analysis. The techniques were further refined during the Shelf Break Primer Experiment in the New England Bight where a nonlinear inversion scheme using the Genetic Algorithm was successfully applied. The inversion scheme was based on modal travel time dispersion characteristics, which in shallow water conditions, is highly sensitive to the bottom properties. Advanced signal processing tools for the space–time–frequency processing and hybrid optimizaton techniques for global search were developed and implemented in parallel. In addition to the individual modal arrival times, emphasis has now been placed on the arrival corresponding to the Airy phase. These inversion techniques are currently being applied to the data from ASIAEX-East China Sea Experiment. Characteristics of this tomography approach such as spatial coverage, resolution, error bars, robustness, and simple instrumentation requirements will be presented. The usefulness of this inversion approach as a rapid environmental assessment tool will be highlighted. [Work supported by ONR.]

8:35

4aUW3. The effects of the ocean environmental variability on sound propagation and reverberation from ASIAEX (ECS). Fenghua Li, Renhe Zhang, Zhenglin Li, and Jianjun Liu (Natl. Lab. of Acoust., Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, lfh@farad.ioa.ac.cn)

During the duration of the Asian Seas International Acoustics East China Sea Experiment, the ocean environment (e.g., wind speed, sea state, internal wave) varied significantly. Propagation and reverberation data from explosive charges recorded by a vertical line array and horizontal line array at different times during the experiment show that the acoustic data have a strong fluctuation. In this paper, the recorded data are analyzed to estimate the relationship between the acoustic signal fluctuation and the environmental variability. A PE code is used to simulate the effect of the recorded ocean environmental variability on the sound signal fluctuation. The experimental data and numerical results indicate the strong dependence between the acoustic fluctuation and ocean environmental variability. A coupled normal-mode-based model is also used to explain some of the acoustic fluctuation. [Work supported by the National Natural Science Foundation of China and ONR.]

8:55

4aUW4. Environmental measurements and mid-frequency backscatter in shallow water. Dajun Tang (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

To understand bottom backscattering mechanism at the mid-frequency range (nominally 3.5 kHz), direct measurements of bottom roughness and sub-bottom heterogeneity were made along with concurrent backscatter measurements in a shallow water site with 105 m water depth. The backscatter was recorded on a vertical line array with 31 elements with a 3.5 kHz source attached at the bottom of the array. Bottom roughness and sub-bottom heterogeneity were measured using an *in situ* conductivity probe. The roughness measurements cover a one-dimensional profile of approximately 4 m in length with vertical resolution of 4 mm and horizontal resolution of 1.5 to 2.5 cm. Heterogeneity measurements cover a depth of 5 to 10 cm. Ambient noise received by the vertical line array was used to estimate the sound speed, density, and attenuation coefficient of the surficial sediments. A Monte Carlo modeling capability was developed to extensively simulate the backscatter with different environmental inputs. The results of this effort covering all three areas of acoustics, environments, and modeling will be reported.

Contributed Papers

9:15

4aUW5. Analysis of time series data in the East China Sea generated from explosive sources. D. P. Knobles, Thomas W. Yudichak (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029), Peter Cable, Y. Dorfman (BBN Technologies), Peter H. Dahl (Univ. of Washington, Seattle, Seattle, WA), James H. Miller, Gopu R. Potty (Univ. of Rhode Island, Providence, RI 02803), Renhe Zhang, Zhaohui Peng, Fenghua Li, and Zhenglin Li (Inst. of Ocean Acoust., Beijing, PROC)

Time series data collected on the APL–UW/URI VLA in the East China Sea as part of the Asian Sea International Acoustics Experiment are analyzed for the information they contain on the characteristics of the seabed. Sound generated by explosive sources deployed by the IOA propagates in a shallow water wave guide under downward refracting conditions, making the received field at the VLA sensitive to the structure of the seabed. A broadband normal mode approach is used to model the measured time series in the 10–500 Hz band. The complex multipath arrival pattern as a function of source–receiver range and source depth allows one to infer certain characteristics of the seabed without the aid of an inversion approach. A finer specification of the seabed, including the determination of the statistics of the geoacoustic parameters is achieved by a simulated annealing inversion methodology. The sensitivity of the acoustic propagation to the elastic properties of the seabed is also investigated along with the nature of the attenuation. The estimated properties of the seabed are compared with independent geophysical measurements and those that can be obtained from reverberation data. [Work supported by ONR.]

9:30

4aUW6. Analysis of sound propagation data taken in the East China Sea. Zhaohui Peng, Ji-xun Zhou (Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC and School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332, zpeng@sununo.me.gatech.edu), Renhe Zhang (Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC), and Peter H. Dahl (Univ. of Washington, Seattle, WA 98105)

As a part of the Asian Seas International Acoustic Experiment (ASIAEX2001), sound propagation data from wideband (explosive) sources were recorded in the East China Sea by using a 32-element suspended array. The propagation was measured as a function of range in two perpendicular tracks (one up to 60 km) and as a function of azimuth for a fixed range of 30 km. Supporting environmental data, obtained from a 17-element thermister chain, XBT, and CTD, showed very complex variation in the water column. In this paper, transmission loss (TL) as a function of range, frequency and azimuth is briefly introduced. Seabottom acoustic parameters such as density, velocity, and attenuation are inverted from the sound propagation data. Then, these parameters plus internal wave data are used as inputs to PE and normal-mode codes qualitatively to explain observed strong fluctuations in sound propagation. The inversion techniques used for estimating seabottom parameters, including spatial mode filtering and dispersion analysis, are also discussed. [Work supported by the National Natural Science Foundation of China and ONR.]

9:45

4aUW7. A comparison of East China Sea low frequency bottom scattering strength determinations. P. G. Cable, R. Gibson, (BBN Technologies, 1300 N. 17th St., Arlington, VA 22209), Y. Dorfman (BBN Technologies, Cambridge, MA 02138) D. P. Knobles, T. W. Yudichak Univ. of Texas, Austin, TX 78713-8029), R. Zhang, Z. Peng, F. Li, Z. Li (Chinese Acad. of Sci., Beijing, PROC), P. H. Dahl (Univ of Washington, Seattle, WA 98105) J. H. Miller, and G. R. Potty (Univ. of Rhode Island, Kingston, RI 02881)

Low frequency shallow water bottom scattering strength determinations conventionally involve accounting for the two-way transmission from source to scattering region to receiver and correcting for the size of the contributing bottom scattering area. Experiment uncertainties in transmission loss, bottom homogeneity and isotropy, contributions from volume, and ocean surface backscatter, can all contaminate the accuracy and robustness of such determinations. Bottom scattering strengths were determined in octave bands from 50–800 Hz reverberation data using Institute of Acoustics sources and the receiving array on Shi-Yan III and from transmission data obtained on the APL-UW/URI receiving array on R/V Melville during East China Sea ASIAEX in 2001. Seabed geoacoustic parameters were inferred from the forward data, which were then used to model transmission and extract scattering strength from the reverberation data. The ASIAEX scattering strengths were compared with other measurements of bottom scattering strength derived from previously obtained East China Sea broadband data sets. The separate bottom scattering strength determinations, though obtained at different times and East China Sea sites, and by different methods, are very close. The experimental sensitivity of these measurements will be described and implications regarding the robustness of the results to seasonal variability and geographic change discussed. [Work supported by ONR.]

10:00–10:15 Break

10:15

4aUW8. Spatial coherence of signals forward scattered from the sea surface in the East China Sea. Peter H. Dahl, Christian J. Eggen, and Russell D. Light (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

Measurements of sea surface forward scattering, wind speed, and directional wave spectra made in 100 m of water in the East China Sea are discussed. The experiment was part of the Asian Seas International Acoustics Experiment (ASIAEX) conducted in the spring of 2001. Signals were received at ranges near 500 m on 2 vertical line arrays that were colocated but separated in depth by 25 m. Estimates of the vertical spatial coherence along these arrays as a function of frequency, path geometry, and sea surface environmental conditions are compared with a model for spatial coherence. The model is based on identifying the probability density function that describes vertical angular spread at the receiver position. An alternative approach utilizing the van Cittert–Zernike theorem from statistical optics is shown to give equivalent results. Both approaches require computation of the sea surface bistatic cross section, done here with the small slope approximation. Forward scattering from the sea surface represents an important channel through which sound energy is transmitted, and spatial coherence determines in part the performance of imaging and communication systems that utilize the sea surface bounce path.

10:30

4aUW9. Time- and space-varying interference patterns of broadband acoustic field sampled by drifting buoys during ENVERSE 97 experiments. Jean-Pierre Hermand (Royal Netherlands Naval College, Postbus 10000, 1780 CA Den Helder, The Netherlands, jhermand@ulb.ac.be), Serge Scevenels (Université Libre de Bruxelles, B-1050 Brussels, Belgium), and Frans G. J. Absil (Royal Netherlands Naval College, 1780 CA Den Helder, The Netherlands)

During the winter of 1997 SACLANTCEN deployed a fixed controlled sound source, a vertical receive array, and drifting hydrophone buoys in a complex coastal environment on the western Sicilian shelf (ENVERSE 97). The acoustic impulse response of the medium was measured in a broad frequency band as a function of range and azimuth from the source, using repeated, large time-bandwidth-product FM transmissions and DGPS positioning. This paper investigates the combined effects of water column and bottom variability upon the space-frequency distribution of the sound-field intensity. The time and space dependence of extracted features such as patterns of field extrema are analyzed and related to the observed environmental conditions. In particular, perturbations of acoustic field invariants are detected and shown to be well correlated with the range-dependent bottom properties and the time-varying ocean sound speed and current fields. Preliminary modeling (C-SNAP) results obtained from concurrent oceanographic and geophysical ground-truth data (ENVERSE 98) are compared to the acoustic measurements to determine their sensitivity to various environmental parameters. [Work supported by the Royal Netherlands Navy.]

10:45

4aUW10. Temporal resolutions of time-reversal and passive-phase conjugation for underwater acoustic communications. I. Stationary source and receivers in a dynamic shallow-water environment. T. C. Yang (Naval Res. Lab., Code 7120, Washington, DC 20375)

Multipath arrivals cause interference between transmitted symbols and hence, symbol errors in underwater acoustic communications. A time-reversal mirror uses the ocean to combine the multipath arrivals and can be used for underwater acoustic communications. Passive-phase conjugation uses the received data of a probe signal to deconvolve the channel transfer function and thereby removes the multipath effects. In this paper, we study the following temporal resolutions of a time-reversal or passive-phase conjugation process as applied to underwater acoustic communications: (1) the time resolution or the pulse width of a backpropagated, time-compressed pulse as compared with the original transmitted pulse; (2) the effectiveness of temporal focusing as measured by the peak-to-sidelobe ratio of the backpropagated or phase-conjugated pulse; both pulse elongation and sidelobe leakages are causes of intersymbol interference and bit errors for communications; (3) the duration of temporal focusing or the temporal coherence time of the underwater acoustic channel; and

(4) the stability of temporal focusing as measured by the phase fluctuations of successive pulses (symbols). BPSK signals collected at sea were used to extract the above four parameters. The bit error rates are modeled with simulated data. [Work supported by ONR.]

11:00

4aUW11. Temporal resolutions of time-reversal and passive-phase conjugation for underwater acoustic communications. II. The effect of source motion. T. C. Yang (Naval Res. Lab., Code 7120, Washington, DC 20375)

Multipath arrivals cause interference between transmitted symbols and hence, symbol errors in underwater acoustic communications. A time-reversal mirror uses the ocean to combine the multipath arrivals and can be used for underwater acoustic communications. Passive-phase conjugation uses the received data of a probe signal to deconvolve the channel transfer function and thereby removes the multipath effects. In this paper, we study the following temporal resolutions of a time-reversal or passive-phase conjugation process as applied to underwater acoustic communications: (1) the time resolution or the pulse width of a backpropagated, time-compressed pulse as compared with the original transmitted pulse; (2) the effectiveness of temporal focusing as measured by the peak-to-sidelobe ratio of the backpropagated or phase conjugated pulse; both pulse elongation and sidelobe leakages are causes of intersymbol interference and bit errors for communications; (3) the duration of temporal focusing or the temporal coherence time of the underwater acoustic channel; and (4) the stability of temporal focusing as measured by the phase fluctuations of successive pulses (symbols). We analyze BPSK signals collected at sea from a moving source, extract the above four parameters, and evaluate the bit error rate. [Work supported by ONR.]

11:15

4aUW12. Direct path fluctuations due to shallow-water variability. Stephen Karpi, Kevin B. Smith, Steven Ramp (Naval Postgrad. School, Monterey, CA 93943), and Peter H. Dahl (Univ. of Washington, Seattle, WA 98105)

The current interest in enhancing the forecasting capabilities of both active and passive sonar systems employed in littoral regions has greatly escalated. This requires improvements in the general understanding of the influence of shallow-water internal waves on acoustic propagation. This work will contribute to a more fundamental understanding of ocean acoustic propagation and fluctuations in shallow-water regions and examine the influence of shallow-water variability on the relatively short-range waterborne propagation paths. Specifically, internal wave fluctuations will be considered and the influence on the acoustic propagation will be quantified in terms of spatial (vertical) coherence functions. The data to be examined will be generated numerically based on an acoustic propagation model employing environmental data taken from the East China Sea as part of the ONR-sponsored ASIAEX experiments. The results of this analysis will be compared with the measured data currently being analyzed at the Applied Physics Laboratory at the University of Washington.

11:30

4aUW13. The measurement of acoustic dispersion in loosely consolidated, saturated sediments using a water-filled impedance tube. Preston S. Wilson, Eun-Joo Park, Ronald A. Roy, and William M. Carey (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

At frequencies of several kilohertz and below, the measurement of sound propagation in marine sediments is difficult due to the larger wavelengths. Laboratory experiments can be limited by the size of the facility required for propagation studies but are amenable to material property measurements. Impedance tube techniques can be used to measure the complex interfacial properties of small samples over a broad and continuous range of frequencies. From this, frequency-dependent sound speed and attenuation is obtained. Results from compressional wave speed and attenuation measurements made with a laboratory impedance tube using artificial and natural water-saturated sediments will be presented and compared to existing propagation models. Variation of attenuation with frequency will be discussed. [Work supported by the U.S. Navy Office of Naval Research and the Coastal Systems Station.]

11:45

4aUW14. Study on ambient noise generated from breaking waves simulated by a wave maker in a tank. Ruey-Chang Wei and Hsiang-Chih Chan (Inst. of Undersea Technol., Natl. Sun Yat-Sen Univ., Kaohsiung City, Taiwan)

This paper studies ambient noise in the surf zone that was simulated by a piston-type wave maker in a tank. The experiment analyzed the bubbles of a breaking wave by using a hydrophone to receive the acoustic signal, and the images of bubbles were recorded by a digital video camera to observe the distribution of the bubbles. The slope of the simulated seabed is 1:5, and the dimensions of the water tank are 35 m$\times$1 m$\times$1.2 m. The studied parameters of ambient noise generated by breaking wave bubbles were wave height, period, and water depth. Short-time Fourier transform was applied to obtain the acoustic spectrum of bubbles, MATLAB programs were used to calculate mean sound pressure level, and determine the number of bubbles. Bubbles with resonant frequency from 0.5 to 10 kHz were studied, counted from peaks in the spectrum. The number of bubbles generated by breaking waves could be estimated by the bubbles energy distributions. The sound pressure level of ambient noise was highly related to the wave height and period, with correlation coefficient 0.7.

Meeting of the Standards Committee Plenary Group

to be held jointly with the

ANSI-Accredited U.S. Technical Advisory Group (TAG) Meeting for: ISO/TC 43 Acoustics and IEC/TC 29 Electroacoustics

P. D. Schomer, Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43 Acoustics
2117 Robert Drive, Champaign, Illinois 61821

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43
1325 Meadow Lane, Yellow Springs, Ohio 45387

V. Nedzelnitsky, U.S. Technical Advisor (TA) for IEC/TC 29, Electroacoustics
National Institute of Standards and Technology (NIST), Sound Building, Room A147, 100 Bureau Drive, Stop 8221, Gaithersburg, Maryland 20899-8221

The meeting of the Standards Committee Plenary Group will precede the meetings of the Accredited Standards Committees S1, S3, and S12, which are scheduled to take place in the following sequence on the same day:

S12	9:45 a.m. to 12:00 noon
S3	1:45 p.m. to 3:15 p.m.
S1	3:30 p.m. to 5:00 p.m.

Discussion at the Standards Committee Plenary Group meeting will consist of national items relevant to all 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 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 five of its Subcommittees, SC1, SC2, SC3, SC5, and SC6) are also encouraged to attend the Standards Committee Plenary Group meeting even though the S2 meeting will take place one day earlier, on Wednesday, 4 December 2002, at 9:00 a.m.

The U.S. Technical Advisory Group (TAG) Chairs for the various international Technical Committees and Subcommittees under ISO and IEC, which are parallel to S1, S2, S3, and S12 are as follows:

U.S. TAG Chair/Vice Chair	TC or SC	U.S. TAG
ISO		
P. D. Schomer, Chair H. E. von Gierke, Vice Chair	**ISO/TC 43** Acoustics	S1 and S3
P. D. Schomer, Chair H. E. von Gierke, Vice Chair	**ISO/TC 43/SC1** Noise	S12
D. J. Evans, Chair	**ISO/TC 108** Mechanical Vibration and Shock	S2
R. Eshleman, Acting Chair	**ISO/TC 108/SC1** Balancing, including Balancing Machines	S2
A. F. Kilcullen, Chair R. Eshleman, Vice Chair	**ISO/TC 108/SC2** Measurement and Evaluation of Mechanical Vibration and Shock as Applied to Machines, Vehicles and Structures	S2
D. J. Evans, Chair	**ISO/TC 108/SC3** Use and Calibration of Vibration and Shock Measuring Instruments	S2
D. D. Reynolds, Chair	**ISO/TC 108/SC4** Human Exposure to Mechanical Vibration and Shock	S3
R. L. Eshleman, Chair R. F. Taddeo, Vice Chair	**ISO/TC 108/SC5** Condition Monitoring and Diagnostics of Machines	S2
G. Booth, Chair	**ISO/TC 108/SC6** Vibration and Shock Generating Systems	S2
IEC		
V. Nedzelnitsky, U.S. TA	**IEC/TC 29** Electroacoustics	S1 and S3

Meeting of Accredited Standards Committee (ASC) S12 Noise

to be held jointly with the

ANSI-Accredited U.S. Technical Advisory Group (TAG) Meetings for: ISO/TC 43/SC 1, Noise

P. D. Schomer, Chair S12, and Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC 1, Noise
2117 Robert Drive, Champaign, Illinois 61821

R. D. Hellweg, Vice Chair S12
Compaq Computer Corporation, Acoustics Lab, Mechanical Engineering Group, MR01-3/03, 200 Forest Street, Marlborough, Massachusetts 01752

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC 1, Noise
1325 Meadow Lane, Yellow Springs, Ohio 45387

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. There will be a report on the interface of S12 activities with those of ISO/TC 43/SC 1 Noise, including plans for future meetings of ISO/TC 43/SC 1. The Technical Advisory Group for ISO/TC 43/SC 1 consists of members of S12 and other persons not necessarily members of the Committee. 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.

THURSDAY AFTERNOON, 5 DECEMBER 2002 GRAND CORAL 2, 1:30 TO 2:30 P.M.

Session 4pAAa

Architectural Acoustics, Musical Acoustics and Physical Acoustics: Ancient Acoustics III Sounding Instruments

James P. Cottingham, Cochair
Department of Physics, Coe College, Cedar Rapids, Iowa 52402

Roberto Velasquez, Cochair
Trujillo 726, Col. Lindavista, C.P. 07300, D. F. Mexico

Invited Papers

1:30

4pAAa1. The Asian free-reed mouth organs. James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402, jcotting@coe.edu)

Mouth-blown instruments employing a free reed coupled to a pipe resonator have long been known and used throughout East and Southeast Asia. Details of the origin and development of these instruments are not known, but are closely connected with the history and prehistory of a multitude of ethnic groups. These instruments exemplify a pipe–resonator coupling significantly different from the standard wind instruments of European origin. The free reed used is approximately symmetric, often operating on both directions of air flow. In some cases the reed is at or near one end of a closed pipe, but in other examples the reed is mounted in the side of a resonator open at both ends. The instruments can be either multiple pipe instruments with one pipe per note, or a pipe with a single reed and tone holes. A number of experimental studies have been conducted on examples of Asian free-reed instruments, primarily the khaen, bawu, and sheng. These include studies of reed vibration, measurements of sound spectra, and impedance measurements of the pipes. Comparisons have been made between experimental results and theoretical work on the coupling of reed vibration with the pipe resonator.

1:50

4pAAa2. Ancient celtic horns. Murray Campbell (Univ. of Edinburgh, Edinburgh EH9 3JZ, UK)

There is considerable evidence from iconographic and documentary sources that musical lip-reed instruments were important in the early celtic communities of Scotland and Ireland. In recent years several studies have been undertaken with the aim of gaining a better understanding of the musical nature of these ancient horns, and of their place in the life and culture of the time. A valuable source of tangible evidence is to be found in the archaeological remains deposited across Scotland and the whole of Ireland. A project is now under way, under the auspices of the Kilmartin House Trust and the general direction of John Purser, which has brought together an international team of musicians, craftsmen, archaeologists, musicologists and physicists with the aim of analyzing ancient musical artifacts, reconstructing some of the original instruments, and analyzing the sounds they produce. This paper describes acoustical studies carried out on a number of recent reconstructions of wooden and bronze instruments, and discusses the role of acoustics in this type of investigation. [Work supported by Sciart and EPSRC.]

2:10

4pAAa3. Characteristics and speculations on the uses of Strombus trumpets found at the ancient Peruvian center Chavín de Huántar. John Rick (Dept. of Anthropological Sci., Stanford Univ., Bldg. 360, Stanford, CA 94305-2117) and David Lubman (David Lubman & Assoc., Westminster, CA 92683-4514)

A cache of 20 *Strombus* shell trumpets was excavated in 2001 from an underground gallery at Chavín de Huántar, the type site of the Peruvian Early Horizon period (ca. 1200 to 400 B.C.). *Strombus* shell usage stretches from antiquity to present day Peru, with the trumpet function showing remarkable continuity. Soon after their discovery, a dozen of the ancient shells were played to an enthusiastic audience at Lima's National Museum of Archaeology, Anthropology and History. Evidence suggests that *Strombus* trumpets are, and were, used as ritual instruments for the legitimization of political and religious power. Those uses may have developed from earlier practical uses for ritual and communication. This paper describes measured acoustical properties of a *Strombus* horn. *Strombus* blasts are rich in overtones. Source strengths as high as 111 dBA @ 1 M were observed. These high source levels support speculation that shell trumpets could have been used for signaling over great distances. The trumpets can produce strong combination tones, which, accompanied by likely ingestion of local hallucinogenic substances, the use of reflected light, and other sound manipulation, suggests that early leaders in Chavín were using a range of methods to help establish early religious and political authority.

THURSDAY AFTERNOON, 5 DECEMBER 2002 GRAND CORAL 2, 2:45 TO 5:05 P.M.

Session 4pAAb

Architectural Acoustics, Musical Acoustics and Physical Acoustics: Ancient Acoustics IV Pre-Columbian Sounding Instruments

Roberto Velasquez, Cochair
Trujillo 726, Col. Lindavista, C.P. 07300, D. F. Mexico

James P. Cottingham, Cochair
Department of Physics, Coe College, Cedar Rapids, Iowa 52402

Invited Papers

2:45

4pAAb1. Construction practices in pre-Hispanic flutes. Susan Rawcliffe (P.O. Box 924, San Pedro, CA 90733)

The ancient cultures of the Americas were separated by thousands of miles and thousand of years. There was a long history of trade over the miles and the years with many shared cultural ideals and artifacts, as evidenced by their musical instruments. Flutes were the most prevalent instruments found throughout the ancient Americas. Some types are unique to the pre-Hispanic world. Although flutes were constructed from a variety of materials, including bone, cane, seed pods, skulls, it is primarily the ceramic ones that survived, and it is ceramic flutes which form the bulk of this writer's work and research. This paper includes musical demonstrations to show how ancient flutemakers could have manipulated timbre during construction. Clay's plasticity enabled the construction of some instruments, and limited the development of others. Pitch jump flutes, certain Veracruzano whistles, and chamberduct flutes and whistles all share the addition of clay flaps or chambers around the aperture, as do hooded pipes. Some instruments exhibit a seemingly cultural predilection for complex tones that are windy, raspy, or animalistic. Simple adjustments of the airduct promote these timbres. Also included will be samples of the original sounds of ancient flutes.

3:05

4pAAb2. Zapotec whistles. Gonzalo Sanchez (Virgen del Carmen 158, Col. Virgencitas, Nezahualcoyotl, Edo. Mexico, C.P. 57300, Mexico)

Sonorous artifacts or ''pre-Hispanic musical instruments'' have attracted the attention of ethnomusicologists, acousticians and archaeologists. The Mexican cultures that have had most studies on the matter are the Maya, Mexica, and Totonaca ones. In the case of the Oaxaca culture, until recently, investigations of this type had not been made. Thus, the task was to look for and to analyze material that belonged to the cultures that inhabited the present Oaxacan territory from a perspective acoustics-organlogic to be able to understand what type of sonorities that were used. The first step was to consider the project of investigation within the field of ethnomusiclogy. In the explorations of the archaeological sites of the Oaxaca Valley it is common to find whistles, which are complete or in fragments, that show certain organologic characteristics that until now have only detected for the Zapotec culture. In this study the morphology, the chronological location, the contexts, the characteristics of the sounds, the symbolism, and the techniques of elaboration of Zapotec whistles are presented. (To be presented in Spanish.)

3:25

4pAAb3. Yaxchilan's whistles. Roberto Velazquez (Virtual Res. Inst. Tlapitzcalzin, Trujillo 726, Col. Lindavista, Mexico D.F., C.P. 07300, Mexico, rvelaz.geo@yahoo.com)

Eight clay whistles found buried in Small Acropolis temples (650–800 BC) of Yaxchilan were analyzed. The study was centered on three dual whistles with body shapes resembling those of frogs. A method developed in previous studies and evidence from several disciplines were applied. Experimental replicas were made to find possible ancient ways of construction and to test hypotheses. The Helmholtz equation for globular resonators was calculated, sound signals were analyzed with spectrograms, and radiated acoustic power in different modes was estimated (0.0005–0.003W). The power level indicates that whistles' sounds could not be heard well if they were played along with louder Mayan aerophones like those of Bonampak's mural band nor in noisy celebrations in big plazas. Clay frogs were adequate to be used in the Small Acropolis, in the Labyrinth and by big groups, surely related to religion. The whistles can sing like natural frogs and produce beats, and they might have been used to produce a chorus in ceremonies to the god Chaac to call for rain, in H-men rites, or in the infraworld where they were discovered.

3:45

4pAAb4. The songs of Tlaloc: Interference of ten ceramical duct flutes, Offering 89 of the Aztec Templo Mayor. M. A. Adje Both (Inst. for Latin-American Studies, Ruedesheimer Str. 52-54, 14197 Berlin, Germany, adje@antamauna.de)

Among the many preserved sound artefacts deposited in the offerings of the Aztec Templo Mayor are a set of ten tubular duct flutes made from clay, dating Late Postclassic Mesoamerica, 1350–1521 AD. The aerophones are completely painted in blue, and characterized by: (1) a short mouthpiece; (2) a framed aperture; (3) a tube with four fingerholes; and (4) an applicated mask with features of the Aztec rain god Tlaloc, basically three rings and a standardized relief structure of two clouds. While all measurements follow the same pattern, one particular organological distinction was made, as five flutes show an exit hole in the middle ring of the mask and five flutes are stopped. Thus, five instruments sound considerably higher, apart from the minimal pitch deviation of each specimen. Both the tonal capacity of each flute and the acoustics of several flutes played simultaneously were recorded and measured. A series of remarkable interference effects could be produced, which were strongly related to the ritual complex reflected in the offering. Taking in consideration the Aztec concept of music, it could be supposed that they were perceived as a principle of the song, or proper voice of Tlaloc.

Contributed Papers

4:05

4pAAb5. Acoustical differences between pre-Columbian Mexican aerophones. J. Guillermo Contreras Arias (Dept. of Ethnomusicology, Natl. School of Music, UNAM, 126 Xicotencatl Coyoacan, D.F. 04100, Mexico and CENIDIM, INBA, Centro Nacional de las Artes)

This paper discusses Mesoamerican musical instruments that produce special sounds and noises. The types of musical instruments to be discussed and demonstrated include those that (a) produce animal sounds; (b) allow performers to exploit multiphonic and beat effects; and (c) allow performers to produce microtonal and polyphonic effects. Examples are shown of simple flutes, vessel flutes, and whistles of Aztec, Mixtec, Mayan, Olmec, and other pre-Colombian cultures. (To be presented in Spanish.)

4:20

4pAAb6. Musical instruments from the Mayan culture. Isabel Elena Romero Rizo, Jose de Jesus Negrete Redondo, and Pablo Roberto Lizana Paulin (Acoust. Lab. ESIME, IPN, Mexico, plizana98@hotmail.com)

This research work presents different kinds of sound generation instruments from the Mayan culture together with the Pablo Castellanos and Arturo Chamorro classification. It has to be noted that most of these mu-

sical instruments, whistles and noise generators are lost and some of them can only be seen on codex and museums. (To be presented in Spanish.)

4:35

4pAAb7. Ancient noise generators. Sergio Beristain, Rolando Menchaca, and Roberto Velazquez (Acoust. Lab., ESIME, IPN, Mexico, sberista@hotmail.com)

There has been found a whole family of similar artifacts which produce singing noises. These noises sometimes resemble those sounds generated by some animals and/or naturally produced by strong winds passing through holes and edges. It means that these sounds have wide frequency spectrums and very often some clear tones are identified. The original purpose of these artifacts is unknown, but some researchers think that some were used in mortuary ceremonies and employed by H-men. The Olmecan whistle previously presented belongs to this family, and now it is compared with a bone or wooden instrument shown in the Florentine codex.

4:50

4pAAb8. The musical instruments in the codexes mixtecos. Luis A. Gomez (Coord. de Documentacion, CENIDIM, CNA, Rio Churubusco y Tlalpan s/n, Country Club, Coyoacan, Mexico)

The objective of this work is to describe the inventory of prehispanic musical instruments in the Codexes Mixtec: Vindobonensis, Nuttall, Selden, Colombino, Becker I and II. For this, relative data are gathered as to their quantity, localization and classification; their function in the Meso-American colored writing; and the meaning of the image of which they are part. This inventory will allow academics of different disciplines to select and to locate the image of the musical instrument of interest quickly and to understand in a general way, its function in the image that contains it. (To be presented in Spanish.)

THURSDAY AFTERNOON, 5 DECEMBER 2002 CORAL KINGDOM 2 AND 3, 1:00 TO 2:05 P.M.

Session 4pAB

Animal Bioacoustics: History of Animal Bioacoustics

Mardi C. Hastings, Chair
Biomedical Engineering Center, The Ohio State University, 270 Bevis Hall, 1080 Carmack Road, Columbus, Ohio 43210-1002

Chair's Introduction—1:00

Invited Paper

1:05

4pAB1. History of animal bioacoustics. Arthur N. Popper (Dept. of Biol. and Neurosci. & Cognit. Sci. Prog., Univ. of Maryland, College Park, MD 20742, apopper@umd.edu) and Robert J. Dooling (Univ. of Maryland, College Park, MD 20742)

The earliest studies on animal bioacoustics dealt largely with descriptions of sounds. Only later did they address issues of detection, discrimination, and categorization of complex communication sounds. This literature grew substantially over the last century. Using the *Journal of the Acoustical Society of America* as an example, the number of papers that fall broadly within the realm of animal sound production, communication, and hearing rose from two in the partial first decade of the journal in the 1930's, to 20 in the 1970's, to 92 in the first 2 years of this millennium. During this time there has been a great increase in the diversity of species studied, the sophistication of the methods used, and the complexity of the questions addressed. As an example, the first papers in JASA focused on a guinea pig and a bird. In contrast, since the year 2000 studies are often highly comparative and include fish, birds, dolphins, dogs, ants, crickets, and snapping shrimp. This paper on the history of animal bioacoustics will consider trends in work over the decades and discuss the formative work of a number of investigators who have spurred the field by making critical theoretical and experimental observations.

Session 4pBB

Biomedical Ultrasound/Bioresponse to Vibration: Ultrasound Applications

Ronald A. Roy, Cochair
Department of Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Salvador Echeverria, Cochair
CENAM, Los Cues, Queretaro, Mexico

Contributed Papers

1:00

4pBB1. Ultrasound detection and computer segmentation of high intensity focused ultrasound (HIFU) lesions. Neil R. Owen, Michael R. Bailey (Ctr. for Indust. and Med. Ultrasound, Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698), Yongmin Kim (Univ. of Washington, 406 EE/CSE, Seattle, WA 98195), and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105-6698)

High intensity focused ultrasound (HIFU) can necrose tumors or cauterize tissue bleeds at intensities on the order of 1000 w/cm^2. A synchronized HIFU and *B*-mode ultrasound system reveals a hyperechoic region at the treatment site, which grows with treatment duration and intensity. Our goal was to segment the hyperechoic region representing the lesion via image analysis and measure the ratio of its major and minor axes. Our algorithm uses the RF data as input, processes it, and outputs a binary image that represents the lesions cross-sectional profile. With depth settings from the clinical ultrasound imager, it is possible to calculate lesion dimensions from the binary image. The algorithm was tested with lesions made in a transparent polyacrylamide tissue phantom that became opaque in response to focal heating during HIFU exposure. Lesion size was recorded simultaneously with ultrasound and a CCD camera, and both measurements agreed well. Additionally, computer segmentation agreed well with segmentation by HIFU users blinded to the experimental conditions. The average difference of the determined ratio was 13% for lesions less than 0.5 cm in length. Thus, it is possible to localize precisely the treated tissue region. [Work supported by NSBRI, NSF, and NIH-SBIR.]

1:15

4pBB2. A quantitative comparison of theory and experiments on high intensity focused ultrasound (HIFU) induced heating in a vascularized tissue phantom. Ronald A. Roy, Jinlan Huang, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, ronroy@bu.edu)

High intensity focused ultrasound (HIFU) can be used to help control bleeding, both from individual blood vessels as well as from gross damage within the capillary bed. The dominant mechanism is rapid localized heating, and vascularity limits ones ability to elevate the temperature of blood vessels owing to convective heat transport. To better understand HIFU-induced heating in tissues with vascular structure, we employ a numerical scheme that couples models for ultrasound propagation, acoustic streaming, ultrasound heating, and blood cooling in Newtonian viscous media. We coordinate the theoretical effort with a series of *in vitro* experiments employing nonuniform flow-through tissue phantoms and designed to provide a ground truth verification of the model predictions. Calculated and measured results are compared over a range of parameter values (insonation pressure, insonation time, and flow rate). We demonstrate excellent agreement between predictions and measurements, and the simulations are extended to a study of the efficacy of HIFU in producing temperature elevations in large and small blood vessels. [Work supported by DARPA and the U.S. Army.]

1:30

4pBB3. High intensity focused ultrasound for treatment of bleeding and air leaks due to lung trauma. Shahram Vaezy, Roy Martin, Larry Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98195), Carol Cornejo, Jerry Jurkovich, Sam Sharar (Harborview Medical Ctr., Univ. of Washington, Seattle, WA 98195), and Savvas Nicolaou (Univ. of British Columbia)

In thoracic surgery, bleeding and air leaks from the lungs can be difficult to control. We have investigated the use of high intensity focused ultrasound (HIFU) for control of lung bleeding and air leaks in operative situations. An intraoperative HIFU device, equipped with a Titanium coupler, was used. The HIFU transducer was a PZT-8 concave element, with a focal length of 5 cm, and a diameter of 2.5 cm. The transducer was operated at 5.7 MHz and intensity of 5000 W/cm^2. The coupler length was 4 cm, placing the focal volume, defined by full-width half-maximum at approximately 1 cm from the tip of the coupler. A pig animal model was used. Incisions in the lung were made, having lengths of 2–5 cm, and depths of 3–10 mm which created both parenchymal hemorrhage and air leakage from the lung. HIFU was applied within 10 seconds of inducing the injury. The average hemostasis time was approximately 60 seconds. All incisions were completely sealed, and no blood or air leaked from the incisions. Intraoperative HIFU may provide an effective method in various pulmonary surgery indications, and hemostasis and control of air leaks from lacerations due to trauma. [Work supported by NIH and NSF.]

1:45

4pBB4. A needle-size hydrophone and thermocouple for high-intensity focused ultrasound. Shahram Vaezy (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98195), Mark Wyzgala, and George Keilman (Sonic Concepts, Inc., Woodenville, WA 98072)

Thermal and acoustic measurements at high acoustic intensities are desirable in studies involving high-intensity focused ultrasound (HIFU). We are developing a needle-size probe comprised of a hydrophone and a thermocouple, that can withstand high-intensity fields without damage. The probe can simultaneously map the acoustic pressure and temperature field in a test medium. The hydrophone will be nominally 25-m PVDF or PVDF-TrFE with thin-file electrodes, with a flat frequency response from 1 to 20 MHz, isolated from the metallic needle by an ultrasound-absorbing lossy polymer material. The thermocouple (accurate within 1C) signals will be rf-isolated from the hydrophone signals and electromagnetic interference originating from the HIFU source transducer, allowing rapid temperature rise measurements (250 readings per second) to be made with minimal noise and interference. A conformational coating will protect the hydrophone and the thermocouple from cavitation at high acoustic intensities of 4000 W/cm^2. The complete assembly will be integrated into a 13-gauge (0.24-mm o. d.) needle (150 mm long). The results of the manufacturing, characterization, and testing in novel, protein-based tissue-mimicking phantoms will be presented. [Work supported by NIH, NSF.]

2:00

4pBB5. Thrombolytic effects of 120-kHz and 1-MHz ultrasound and tissue plasminogen activator on porcine whole blood clots. Christy K. Holland, Sampada S. Vaidya, Constantin-C. Coussios, and George J. Shaw (Dept. of Biomed. Eng., Univ. of Cincinnati, 234 Goodman St., Cincinnati, OH 45267-0761)

Stroke is the third leading cause of death and the leading cause of disability in the United States. For patients with ischemic stroke, the thrombolytic drug tissue plasminogen activator (tPA) is the only FDA-approved treatment. To aid in the development of a stroke therapy, the synergistic thrombolytic effect of tissue plasminogen activator (t-PA) and ultrasound was assessed *in vitro* in a porcine clot model. Whole blood clots were prepared from fresh porcine blood by aliquoting 1.5 ml into 8-mm-i.d. glass tubes, immersing the tubes in a 37 °C water bath for 3 h and storing the clots at 5 °C for at least 3 days prior to use in comparative ultrasound and t-PA studies, which ensured complete clot retraction. The 120-kHz or 1-MHz ultrasound peak-to-peak pressure amplitude used for exposures was 0.35, 0.70, or 1.00 MPa. The range of duty cycles varied from 10% to 100% (continuous wave) and the pulse repetition frequency was 1.7 kHz. Clot mass loss was measured as a function of t-PA concentration (without ultrasound) from 0.25 to 8.5 times the human clinical dose. The mass loss increased monotonically as a function of [t-PA] and saturated at 4 times the clinical dose. The degree of the ultrasound enhancement of t-PA was also explored at 120 kHz and 1 MHz. With ultrasound exposure, clot mass loss increased by as much as 250% over sham (t-PA alone). A weak dependence of clot mass loss on duty cycle was noted. We conclude that 120-kHz and 1-MHz ultrasound enhances thrombolysis. [Work supported by Senmed Medical Ventures.]

2:15

4pBB6. Influence of vasoactive drugs on the induction of microvascular leakage in rat hearts by contrast echocardiography. Douglas Miller (Dept. of Radiol., 3315 Kresge III, Univ. of Michigan, 200 Zina Pitcher Pl., Ann Arbor, MI 48109-0553, douglm@umich.edu), Peng Li, and William Armstrong (Univ. of Michigan, Ann Arbor, MI)

Diagnostic ultrasound can induce microvascular leakage in rat hearts when a gas-body based contrast agent is present in the circulation, a response which might be partly physiological in origin. Anesthetized rats with tail vein cannulas were mounted in a 37 °C water bath. Rats were pretreated with saline, or with isoproterenol, propranolol or dobutamine, drugs which can modulate microvascular leakage. Evans blue dye and a 0.5 ml/kg bolus of Optison (Mallinckrodt) were injected before 5 min exposures. A diagnostic ultrasound scanner (GE Vingmed System Five) operated at 1.7 MHz and a measured Mechanical Index of 1.5 was used to obtain cardiac images with end-systole triggering each fourth beat. For shams, only a few petechia were occasionally observed on the excised hearts. For exposure (five rats in each treatment), an average of 208 (86 standard deviation, SD), 200 (72 SD), 236 (52 SD), and 159 (57 SD) petechiae were counted for saline, isoproterenol, propranolol, and dobutamine, respectively (no significant differences). Evans blue dye leakage also appeared to be approximately the same for all treatment conditions. These results suggest that the microvascular leakage induced by contrast echocardiography in rats in primarily a mechanical effect of the ultrasound-induced gas body activity.

2:30

4pBB7. Effect of absorption on nonlinear propagation of short ultrasound pulses generated by rectangular transducers. Vera A. Khokhlova, Anatoly E. Ponomaryov (Dept. of Acoust., Faculty of Phys., M. V. Lomonosov Moscow State Univ., Leninskie Gory, Moscow 119992, Russia), Michalakis A. Averkiou (Philips Ultrasound, P.O. Box 3003, Bothell, WA 98041), and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

A numerical solution of the KZK-type parabolic nonlinear evolution equation is presented for finite-amplitude sound beams radiated by rectangular sources. The initial acoustic waveform is a short tone burst, similar to those used in diagnostic ultrasound. The generation of higher harmonic components and their spatial structure are investigated for media similar to tissue with various frequency dependent absorption properties. Nonlinear propagation in a thermoviscous fluid with a quadratic frequency law of absorption is compared to that in tissue with a nearly linear frequency law of absorption. The algorithm is based on that originally developed by Lee and Hamilton [J. Acoust. Soc. Am. **97**, 906–917 (1995)] to model circular sources. The algorithm is generalized for two-dimensional sources without axial symmetry. The diffraction integral is adapted in the time–domain for two dimensions with the implicit backward finite difference (IBFD) scheme in the nearfield and with the alternate direction implicit (ADI) method at longer distances. Arbitrary frequency dependence of absorption is included in this model and solved in the frequency–domain using the FFT technique. The results of simulation may be used to better understand the nonlinear beam structure for tissue harmonic imaging in modern medical diagnostic scanners. [Work supported by CRDF and RFBR.]

2:45–3:00 Break

3:00

4pBB8. Constrast enhancement in harmonic imaging using time reversal processing. Mickael Tanter, Thomas Gauthier, Christophe Barriere, and Mathias Fink (Laboratoire Ondes et Acoustique, E.S.P.C.I, C.N.R.S. UMR 7587, Universit Paris VII, 75005 Paris, France)

Harmonic imaging coupled to the use of contrast agents has become a topic of wide interest in ultrasonic medical imaging. Instead of achieving images at the fundamental frequency of the emission signals, images are built at twice the excitation frequency that corresponds to the resonance frequency of the bubbles embedded in the medium. Consequently, the contrast becomes important between areas of low and high concentration of bubbles. However, at a high mechanical index, the harmonic components of backscattered echoes depends on the intrinsic nonlinear properties of tissues as well as the bubble's resonance. The focused beam is generating harmonics during its nonlinear propagation. It results in a degradation of the harmonic image contrast. Time reversal is an elegant way to find the emission codes allowing us to cancel the harmonic components due to nonlinear propagation of the ultrasonic beam and so to recover an optimal image contrast. Indeed, the wave equation in a nonlinear regime remains time reversal invariant: in the absence of bubbles, if the backscattered echoes are time-reversed and reemitted by the array, the harmonic components of the resulting wavefront are transferred back to the fundamental frequency during propagation. Experiments conducted with 1-D-linear arrays illustrate this cancellation technique.

3:15

4pBB9. Optical monitoring of ultrasound-induced damage in glass catfish after injection with contrast agents. Subha Maruvada and Kullervo Hynynen (Harvard Med. School, Cambridge, MA 02139 and Brigham and Women's Hospital)

This work is an investigation of the ultrasound interaction with contrast agent bubbles in the blood vessels of fish *in vivo*. Glass catfish were used for these experiments because they are optically transparent. Anesthetized fish were injected with the contrast agent, Optison, using microinjection techniques. Shortly after injection, the fish were insonified with one of two single-element focused transducers (1.091 and 0.747 MHz). An inverted microscope combined with a digital camera was used to optically monitor ultrasound interaction with the blood vessels in the tail of the fish at 200× magnification. The entire interaction was videotaped and digitized. The fish were insonified at power levels between 1–80 W, which translated to acoustic pressures from 0.45–3.55 MPa. Sonifications were pulsed with burst lengths of 1 ms, 10 ms, 100 ms, and 100 s, and a

repetition frequency of 1 Hz. The entire length of one sonification at a specific power was 20 s. The effects of the sonification were analyzed at each power level. Ultrasound-induced damage to the fish included structural damage to the cartilage rods in the tail, formation of large gas bubbles within the blood vessels, and hemorrhaging.

3:30

4pBB10. Application of pulse laser induced vapor bubbles and shock waves to neurosurgery. Atsuhiro Nakagawa, Takayuki Hirano (Dept. of Neurosurgery, Tohoku Univ. Grad. School of Medicine, 1-1, Seiryou-machi, Aoba, Sendai 980-8574, Japan), Kazuyoshi Takayama, and Tsutomu Saitoh (Tohoku Univ., 2-1-1, Katahira, Aoba, Sendai 980-8577, Japan)

As a part of the basic research of tissue damages which occurred during ESWL treatments, shock/bubble interactions are found to be responsible to tissue damage. Having this background, we have been applying pulse laser induced vapor bubbles and shock waves to various neurosurgical therapies. Pulsed HO;YAG laser beams were transmitted via 0.6 mm diam optical fiber, whose edge was formed in a convex lens shape so as to enhance the laser beam focusing. Laser beams were focused inside a 2 mm diam water filled tube. Then, vapor bubbles and shock waves were generated, which successively drove a microwater jet. The mechanism of the initiation of bubbles and shock waves and consequential jet formation has been intensively investigated. It was found in *in vitro* experiments that microjets wonderfully pierced cerebral thrombi. Combining this system with a very small amount of fibrinolytic agents, cerebral thrombi were removed four times as efficiently as simple irradiation of pulsed laser beams.

3:45

4pBB11. Mie scattering off coated microbubbles. Radboud Nelissen, Elmer Koene, Sascha Hilgenfeldt, and Michel Versluis (Appl. Phys., Univ. of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, m.versluis@tn.utwente.nl)

The acoustic behavior of coated microbubbles depends on parameters of the shell coating, which are in turn dependent on bubble size. More intimate knowledge of this size dependence is required for an improved modeling of a distribution of coated microbubbles such as found in ultrasound contrast agents (UCA). Here a setup is designed to simultaneously measure the optical and acoustic response of an ultrasound-driven single bubble contained in a capillary or levitated by the pressure field of a focused transducer. Optical detection is done by Mie scattering through an inverted microscope. Acoustical detection of the single bubble by a receiving transducer is made possible because of the large working distance of the microscope. For Mie scattering investigation of excited bubbles, two regimes can be distinguished, which require different detection techniques: Conventional wide-angle detection through the microscope objective is sufficient for bubbles of radius exceeding 10 μm. For smaller bubbles, two narrow-aperture detectors are used to reconstruct the bubble dynamics from the complex angle-dependence of the scattered light.

4:00

4pBB12. High-speed optical imaging of microbubble dynamics at 40 ns time scale. Michel Versluis (Phys. of Fluids, Appl. Phys., Univ. of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, m.versluis@tn.utwente.nl), Chien Ting Chin, and Nico de Jong (Erasmus Univ., Rotterdam, The Netherlands)

The interactions between ultrasound and microbubble contrast agents, such as nonlinear oscillations, acoustic disruptions, and destruction of microbubbles, are exploited in diagnostic methods such as harmonic imaging and triggered imaging, as well as emerging therapeutic techniques such as controlled drug delivery. Optical visualization at the microsecond to nanosecond time scale has the potential to elucidate such interactions. However, these recordings are limited by image resolution, frame rate, total number of frames, and light sensitivity. We have constructed a digital ultra-high-speed camera system, named Brandaris 128, by combining the flexibility of highly sensitive CCDs with the high number of frames available in rotating mirror cameras. In a standard run our camera records 6 full sequences of 128 digital images at a frame rate of 25 million frames per second. The high number of frames allows for the monitoring of acoustic disruptions and nonlinear oscillations of a single contrast agent microbubble under consecutive diagnostic ultrasound pulses at various amplitudes applied to the same bubble. In this way the contribution of each individual bubble to the acoustic response can be measured, which is essential for further optimization of contrast agents and ultrasound detection methods.

4:15

4pBB13. Enhancement of the biocidal efficacy of a mild disinfectant through enhanced transient cavitation. Kenneth A. Cunefare (School of Mech. Eng., Georgia Tech., Atlanta, GA 30332-0405), Stephen D. Carter (Snelleville, GA 30078), and Don Ahern (Georgia State Univ., Atlanta, GA 30302)

Cavitation is known to have biocidal effect upon micro-organisms. The work presented here considered the impact of transient cavitation at elevated hydrostatic pressures as a means to improve the biocidal efficacy of otherwise mild disinfectant solutions. Cultures of Bacillus stearothermophilus and Bacillus subtilis were exposed to a mild isopropyl alcohol-based disinfectant for 15 to 18 minutes, which resulted in negligible cell reduction. Spores from the same cultures were exposed to the disinfectant for 15 to 18 minutes while being cavitated for one minute at a frequency of 24 kHz, and at ambient pressure, with an estimated energy density of 70 watts per liter. This achieved cell reduction of approximately 2%. Spores from the same cultures were exposed to the same cavitation field but at a hydrostatic pressure of approximately 210 kPa. The pressure was selected to exploit ''the anomalous depth effect'' for enhanced transient cavitation. These conditions achieved cell reduction of approximately 90% after 1 minute of cavitation exposure. The results are suggestive of more efficient means of sterilizing the burgeoning numbers of heat sensitive surgical instruments.

Session 4pEA

Engineering Acoustics: Noise Identification and System Elements

Federico Miyara, Cochair
Acoustics and Electroacoustics Laboratory, National University of Rosario, Riobamba 245 BIS, Santa Fe 2000 Rosario, Argentina

Brandon Tinianov, Cochair
Acoustical Laboratory, Johns Manville, 10100 West Ute Avenue, P.O. Box 625005, Littleton, Colorado 80162

Chair's Introduction—1:30

Contributed Papers

1:35

4pEA1. Noise source identification: A comparison of seven different techniques. Marc Marroquin (34465 Hainesville Rd., Round Lake, IL 60073)

One of the most powerful tools that can be used in modern acoustics labs today is noise source identification. By identifying where the noise source(s) may be located, it becomes much easier then to try and reduce/remove/modify the overall dB level of a product. This is typically the end result of a noise reduction program, which is the typical goal of all NVH programs. Over the years many different techniques have become available as technology moves forward. In the beginning it was simple sound-pressure mapping techniques, then sound intensity added a new dimension to the art and science of NSI (noise source identification); last, multichannel applications such as acoustic holography and beamforming have given us tools that far exceed our capabilities of even 10 years ago. As with any advance in technology, each technique has strengths and weaknesses which must be understood to ensure proper measurement and analysis. This paper will offer an overview of seven commonly used NSI techniques along with a comparison of the strengths and weaknesses associated with each technique.

1:50

4pEA2. The effect of envelope pattern on the impression of sound quality. Sonoko Kuwano (Osaka Univ., 1-2 Yamadaoka, Suita, Osaka 565-0871, Japan) and Seiichiro Namba (Takarazuka Univ. of Art and Design, Hanayashiki, Takarazuka, Hyogo 665-0803, Japan)

Since the 1970's the present authors have conducted a series of experiments on the effect of envelope patterns on a subjective impression of sounds and a model of dynamic hearing characteristics was proposed on the basis of the results of these experiments. The present investigation was designed to examine the effect of envelope pattern on the impression of sound quality. Three kinds of envelope pattern were used; they were sounds with short rise time and long decay, sounds with long rise time and short decay, and steady state sounds. The duration of each sound was 125, 250, and 500 ms. In total, nine kinds of sound were prepared. The sound was repeated with silent intervals and the total duration was 5 sec. The sound quality was judged using a semantic differential and a paired comparison. It was found that the envelope pattern has a significant effect on the impression of pleasantness and the duration of the sound on the impression of sharpness and articulation.

2:05

4pEA3. Evaluation of the noise emitted by the condenser of a household refrigerator. Juan Llado Paris and Beatriz Sanchez Tabuenca (Dept. of Mech. Eng., Zaragoza Univ., 50018 Zaragoza, Spain, juan.llado@posta.unizar.es)

It has been determined that the noise emitted by a household refrigerator is the addition of two main sources: the direct noise emitted by both compressor shells, and the structural noise caused by the condenser vibration, that depends on the coolant excitation, on its own structural rigidity, and on the rigidity of the joints that fix it to the back panel of the refrigerator and to the compressors. The decrease of the sound power due to the condenser implies the modification of their mechanical and geometric characteristics: shape, thickness, type of material, redesign of the joints to the back panel of the refrigerator, and optimization of their number and location, etc., to achieve a less rigid structure. From these proposals, a new more flexible design of the joints was tested because it was the faster one to implement due to the refrigerator efficiency is not altered. A sound power reduction of about 2 dBA, was obtained. Finally, the location and number of the joints were optimized, getting the same sound power level with three than with four joints. (To be presented in Spanish.)

2:20

4pEA4. Characterization of the noise emitted by a washing-machine due to the pump. Beatriz Sanchez Tabuenca and Juan Llado Paris (Dept. of Mech. Eng., Zaragoza Univ., 50018 Zaragoza, Spain, bstb@posta.unizar.es)

The noise sources during the pump operation of a washing machine are: the electric motor, the water–air impulsion, and the structural radiation of the pump-housing, hose, and cabinet. The determination of the sound power level under different working conditions let us identify that the noise emitted by the cabinet vibration was the more noticeable, being irrelevant the noise emitted by the other elements. In order to know which part of the cabinet is the noisiest, the standard ISO-TR 7849 was applied to calculate the contribution of the noise emitted by the structural radiation of each part of the cabinet to the total sound power level. This experimental procedure relates the noise radiated by a structure with its vibration velocity, and it was found that the kick plate was the more relevant element. Once the noise transmission paths have been characterized, the proposals to reduce noise are focused on a new design of the kick plate to make it more flexible and the modification of the join zone of the pump to the kick plate to reduce the force transmitted between both elements. (To be presented in Spanish.)

2:35

4pEA5. A study of standing pressure waves within open and closed acoustic resonators. C. Daniels (OAI, 22800 Cedar Point Rd., Cleveland, OH 44142, Christopher.Daniels@grc.nasa.gov), B. Steinetz (NASA Glenn Res. Ctr., 21000 Brookpark Rd., MS 23-3, Cleveland, OH 44135), J. Finkbeiner, G. Raman, and X. Li (Illinois Inst. of Technol., 10 W. 32nd St., E1, Chicago, IL 60616)

The first section of the results presented herein was conducted on an axisymmetric resonator configured with open ventilation ports on either end of the resonator, but otherwise closed and free from obstruction. In the remaining section we present the results of a similar resonator shape that was closed, but contained an axisymmetric blockage centrally located through the axis of the resonator. Ambient air was used as the working fluid. In each of the studies, the resonator was oscillated at the resonant frequency of the fluid contained within the cavity while the dynamic pressure, static pressure, and temperature of the fluid were recorded at both ends of the resonator. The baseline results showed a marked reduction in the amplitude of the dynamic pressure waveforms over previous studies due to the use of air instead of refrigerant as the working fluid. A sharp reduction in the amplitude of the acoustic pressure waves was expected and recorded when the configuration of the resonators was modified from closed to open. A change in the resonant frequency was recorded when blockages of differing geometries were used in the closed resonator, while acoustic pressure amplitudes varied little from baseline measurements.

2:50

4pEA6. Acoustic characteristics of a three-chamber hybrid silencer. Iljae Lee, Ahmet Selamet (Ctr. for Automotive Res., The Ohio State Univ., 930 Kinnear Rd., Columbus, OH 43212, selamet.1@osu.edu), and Norman T. Huff (Owens Corning, Inc., Novi, MI 48377)

The acoustic characteristics of a hybrid silencer consisting of two identical single-pass, concentric, perforated dissipative chambers combined with a reactive Helmholtz resonator in between are investigated computationally and experimentally. Transmission loss predictions from a three-dimensional boundary element method are compared with experimental results obtained from an impedance tube setup in the absence of mean flow. In addition to the overall design, the effects of filling material density in dissipative chambers and the neck geometry of the reactive resonator are examined. The dissipative chambers are found to be effective at high frequencies, while the reactive resonator is shown, in general, to improve the acoustic attenuation at low frequencies typical of airborne noise in internal combustion engines running at low- to mid-speed range. The acoustic behavior near the resonance frequency is determined, however, to be sensitive to the duct length connecting dissipative and reactive chambers. Potential merits of the hybrid concept are assessed.

3:05

4pEA7. Frequency-domain methods for modeling a nonlinear acoustic orifice. David P. Egolf[a] (Dept. of Elec. and Computer Eng., Univ. of Idaho, Moscow, ID 83844, degolf@ece.uidaho.edu), William J. Murphy, John R. Franks (National Inst. for Occupational Safety and Health, Cincinnati, OH 45226), and R. Lynn Kirlin (Univ. of Victoria, BC, Canada)

This presentation describes frequency-domain methods for simulating transmission loss across a single orifice mounted in an acoustic waveguide. The work was a preamble to research involving earplugs containing one or more orifices. Simulation methods included direct Fourier transformation, linearization about an operating point, and Volterra series. They were applied to an electric-circuit analog of the acoustic system containing the orifice. The orifice itself was characterized by an empirical expression for nonlinear impedance obtained by fitting curves to experimental resistance and reactance data reported by other authors. Their data-collection procedures required the impedance expression presented herein to be properly labeled as a describing function, a quantity well known in the nonlinear control systems literature. Results of the computer simulations were compared to experimental transmission-loss data. For a single-tone input sound pressure, the computer code accurately predicted the output fundamental (i.e., without harmonics). For a broadband input, the simulated output was less accurate, but acceptable. Levels of the sound-pressure input ranged from 60 to 160 dB. [Work supported by the National Institute for Occupational Safety and Health, Cincinnati, OH, through a research associateship granted the first author by the National Research Council.] [a]Currently on leave at National Institute for Occupational Safety and Health, Cincinnati, OH.

3:20

4pEA8. Experimental investigation of sound absorbers based on microperforated panels. Nilda S. Vechiatti, Antonio M. Mendez (Laboratorio de Acustica y Luminotecnia, C.I.C. Pcia. de Buenos Aires, Cno. Centenario y 506, (1897) M. Gonnet, Argentina, ciclal@gba.gov.ar), and Juan C. Gimenez de Paz (Decibel Sudamericana, Las Bases 165 (1706), Haedo, Argentina)

Microperforated panels have been studied as a good, interesting absorbing element. In previous papers the properties of high absorption obtained in a wide frequency band was demonstrated, based on the impedance of the very small perforations. The perforation area ratio, the diameter of the holes, the thickness of the panel, and the density and viscosity of the air are the terms that define the sound absorption provided by the element. The microperforated panels have a simple structure and it is possible to build single or double resonators, in order to obtain a wide band response. An orifice may be considered as a short tube. Many years ago, Rayleigh and Crandall studied the propagation of sound in small tubes, of a very short length compared to wavelength. They found a high acoustics resistance and a very small reactance. So, the microperforated panel can be used as a dissipative element. An experimental investigation was carried out on different samples of microperforated panels, in order to obtain their sound absorption coefficient and so verify the validity of the mathematical models. Microperforated panels have been developed to cover a welding cabin internally, where classical absorbents are useless. (To be presented in Spanish.)

3:35

4pEA9. Interaction of a slot-tone with a pipe. J. Alexis Billon, Vincent Valeau, and Anas Sakout (L.E.P.T.A.B., Univ. La Rochelle, Av. M. Crepeau, 17042 La Rochelle Cedex 1, France)

The self-sustained tone generated by a low Mach number free flow impinging on a slotted plate, usually referred to slot-tone, is experimentaly studied. In this work we focus on its coupling with the pipe from which the jet flows out. The jet nozzle and the beveled slot are aligned and have the same large aspect ratio. The tones generated have a single well defined frequency (around 1 kHz), a high amplitude (typically, 105 to 115 dB) and exhibit typical characteristics of self-sustained tones. Their frequencies are about the same order of magnitude than the natural frequency of the shear layer calculated from the linear stability theory. For some Reynolds numbers and some plate distances, a standing wave pattern is observed in the pipe. The outer and inner acoustic fields show high coherence and the slot-tone locks on a pipe eigenmode, which indicates an acoustic feedback. However, especially for lower Reynolds numbers, this coherence vanishes but the self-sustained tones are still present. It suggests that the hydrodynamic feedback becomes preponderant.

3:50

4pEA10. Turbulent airflow noise production and propagation patterns of a subsonic jet impinging on a flat plate. Stephen Martin (Wyle Labs., 128 Maryland St., El Segundo, CA 90245) and William Meecham (UCLA, Los Angeles, CA 90024)

Turbulent airflow noise production is examined utilizing a four-inch diameter jet impinging on a flat plate. The analysis utilizes cross-correlation techniques to investigate the far field sound pressure relationship to turbulent sources produced near the plate surface. The variables examined include the effects of jet speed, the angle of the plate with respect to the jet axis, the location of the turbulent source from the plate center, the location of the far-field microphone, and the plate-jet distance.

The cross-correlation techniques allow discrimination of the far-field sound pressure related to the turbulent noise source located near the surface of the plate. By rotating the far-field microphones, directivity patterns are obtained for a source located a particular distance from the plate center. The resulting cloverleaf shaped directivity patterns indicate a quadrupole source, verifying Powells developments on Lighthills theory on jet noise. The variations of the parameters of the experiment allow a comparison of the directivity patterns for each condition. [Work was performed through the UCLA School of Engineering with the experimental data collected at the NASA-Ames research facility.]

4:05

4pEA11. The development of polymer layers for vibration damping steel sheet. T. S. Oh and J. H. Ryou (Dept. of Metal and Coating, RIST (Res. Inst. of Industrial Sci. and Technol.), Pohang, P.O. Box 135, 790-600, South Korea, ohts@rist.re.kr)

The characterization of damping resins, and the forming characteristics of vibration damping steel sheets were carried out in order to evaluate polymer layers for vibration damping steel sheets. The bonding strength and damping properties depend on the chemical nature of individual polymers on the vibration damping steel sheets. Polyester-based thermoplastic polymer shows the best performance. Damping characteristics of vibration damping steel sheet can be well predicted from the dynamic loss factor of resins via the WLF equation. The vibration damping properties of polymers can be changed by a plasticizer. It was found that the service temperature of polymer generally decreased using an increasing amount of plasticizer. It is concluded that the vibration damping steel sheet made by the loss factor greater than 0.1 in the range of normal service temperature as well as the tensile shear strength greater than 80 kg_f/cm^2 at room temperature and the addition of Ni powder have been found to improve weldability the same an ordinary steel sheet. [Work supported by POSCO.]

4:20

4pEA12. Development of beam volume displacement sensors using genetic algorithm. Saiprasad Jangiti, Marcellin Zahui, and George Bibel (Dept. of Mech. Eng., Univ. of North Dakota, Box 8359, Grand Forks, ND 58202-8359, Marcellin.Zahui@und.edu)

Further development of total (local) volume displacement sensors is presented. This development supports the implementation of noise control techniques that are based on the minimization of local volume displacement of vibrating structures. In this work, a genetic algorithm is used for the design of volume displacement sensors for vibrating beams using PolyVinyliDene Fluoride (PVDF). First the sensor is assumed to cover the entire beam surface. Then, the covered area is discretized into small patches. The algorithm selects the required patches to yield the PVDF shape necessary to measure the beam volume displacement. The sensor is numerically verified for various beam boundary conditions. The results show close agreement between the calculated beam volume displacement of the beam and the simulated output charge of the sensor. The extension of the genetic algorithm methodology to volume displacement sensors for 2-D vibrating structures is briefly discussed.

THURSDAY AFTERNOON, 5 DECEMBER 2002 CORAL GALLERY FOYER, 1:00 TO 2:30 P.M.

Session 4pED

Education in Acoustics: Take Fives: Sharing Ideas for Teaching Acoustics

Uwe J. Hansen, Chair
Department of Physics, Indiana State University, Terre Haute, Indiana 47809

Chair's Introduction—1:00

Do you have a novel demonstration, a new laboratory experiment, a favorite video, a recorded sound example, or a new idea for teaching acoustics which you are willing to share with your colleagues? At this session a sign-up board will be provided for scheduling presentations. No abstracts are printed. Presenters are encouraged to have handouts to distribute. Multiple presentations are acceptable (not consecutively). Presentations are limited to 5 minutes. Keep them short! Keep them fun!

Session 4pNS

Noise and Engineering Acoustics: Noise in Urban Communities

Martha Orozco, Chair
Div. de Ciencias Biologicas y Ambientales, CUCBA, Universidad de Guadalajara Km 5 1/2, Carr. a Nogales, Las Agujas Zapopan, Jalisco, Mexico

Chair's Introduction—1:00

Invited Papers

1:05

4pNS1. Characteristics of successful local noise enforcement programs. Eric Zwerling (Rutgers Univ. Noise Tech. Assistance Ctr., 14 College Farm Rd., New Brunswick, NJ 08901)

A well crafted local noise ordinance, customized to the needs of the jurisdiction, enforced by motivated and well-trained officers, can resolve long-standing problems, and prevent new ones from arising. The ordinance must be simple enough for all parties to understand: enforcement, prosecution, adjudication, residents and importantly, the regulated community. Uncomplicated performance standards lead to reliably frequent enforcement, the deterrence from which engenders self-policing by source facilities, after initial unsuccessful court challenges to the new ordinance (e.g., St. Augustine, FL). New facilities will be designed for compliance with a well established ordinance (e.g., Denville, NJ). Innovative approaches, such as an interior C-scale standard, successfully curb amplified low-frequency rhythmic emissions from bars without the need for 1/1 or 1/3 octave band analysis (e.g., Anchorage, AK). A ''plainly audible'' standard for vehicular sound systems is effective, easily enforced (e.g., Lafayette, LA), court tested, and enhanced penalties such as towing can reduce complaints and recidivism to negligible levels (e.g., Rochester, NY). Objective, content-neutral standards successfully control amplified speech (e.g, Long Beach, NY) and street music (New York City). Local communities are increasingly developing noise enforcement programs, many of which are highly effective.

1:25

4pNS2. Urban noise pollution in the city of Curitiba, Brazil. Paulo H. T. Zannin and Fabiano B. Diniz (Environ. Acoust. Lab., Federal Univ. of Parana, Parana, Brazil)

This paper presents the results obtained from a study on environmental noise pollution in the city of Curitiba, Brazil. The equivalent sound-level values LA_{eq}, 2 h were measured and tabulated for 1000 locations spread over the urban zones of the city of Curitiba in several urban zones: residential, mixed, services, downtown, and industrial. It was found that 93.3% out of the surveyed locations display during the day equivalent sound levels over 65 dB(A), and 40.3% out of the total number of locations measured during the day display extremely high values of equivalent sound levels: over 75 dB(A). The average noise levels per zone has ranged from 73.4 to 78.1 dB(A). (To be presented in Portuguese.)

1:45

4pNS3. Interior noise profile of buses in Curitiba. Paulo H. T. Zannin, Clifton R. Giovanini, Fabiano B. Diniz, and Jose C. Ferreira (Environ. Acoust. Lab., Federal Univ. of Parana, Parana, Brazil)

Evaluating the noise levels to which the bus drivers of Curitiba are exposed to during their working days is the main scope of this study. The city is served by an internationally known public transportation system featuring 1902 buses, which attend 1.9 million people per day. Two measurements have been taken inside each one of the 60 buses surveyed, one close to the driver and another one at the back of the bus. The results have showed that the dose levels the drivers are exposed to were below 50% in 92% out of the buses, but the normalized exposure levels were over 65 dB(A) in all cases. This level is considered as the threshold of comfort according to the Brazilian legislation on occupancy health NR-17—Ergonomics. The surveyed buses have been divided into three categories, according to their characteristics: feeder, rapid, and bi-articulated. A total of 20 buses within each category have been surveyed. Among the different categories, it has been found that the feeders have presented the highest noise levels. (To be presented in Portuguese.)

2:05

4pNS4. Analysis of the levels of ambient noise present in the Colonia Auditorium (October 2001). Zapopan Jalisco, Lourdes Palafox, Martha Orozco, and Erika Rodriguez (Div. de Ciencias Biologicas y Ambientales, CUCBA, Universidad de Guadalajara Km 15 1/2 Carr. a Nogales, Las Agujas Zapopan, Jalisco, Mexico)

The study area is located in the city of Guadalajara, defined by: delimited in the north by the Gonzlez Gallo street; in the south by the Anillo Perifrico street; in the east by Avenida Mezquitan; and in the west by Avenida Alcalde. There were 38 measurement points outside the auditorium and 25 inside. A poll was conducted among the neighbors and visitors in both areas, getting 70 replies. Noise levels were between 41–101 dB(A) on the outside, average Leq on the outside was 64.5 dB(A), and inside was 80.15 dB(A). The conclusion is that the whole area is affected by traffic noise, as well by the huge number of visitors, the electric games, and the electric generators that supply energy to them. On the outside 21 points were above 65 dB(A), which means that 55% were above the average; on the inside, 100% were above 65 dB(A). Thus the neighbors are exposed to high noise levels due to the festivities during October. The biggest problem, however, is the noise exposure of the workers inside the auditorium, due to music, games, people, and electric generators.

2:20

4pNS5. Noise levels at critical points in the municipality of Guadalajara, Jalisco, Mexico. Arturo Figueroa, Jesus Garcia, Jorge Macias (Depto. de Estudios y Proyectos Amb. Direccin General de Medio Ambiente y Ecologia, H. Ayuntamiento de Guadalajara, Marsella 49, 5to. piso, Mexico), Martha Orozco, Javier Garcia (Las Agujas, Zapopan Jalisco, Mexico), and Alan Delgadillo (Secretaria de Educacion, Guadalajara, Jalisco, Mexico)

Studies of acoustic conditions are planning tools on which we can diagnose the problem of noise pollution in the cities. The first study on noise pollution made in the city was made by the University of Guadalajara in 1995 and updated in 1998 covering with measuring points the city center. This paper discusses the problem of noise pollution by motor vehicles at critical points and covers a total of 105 points. The study also analyzes the problem of noise pollution base on the community annoyance from which a regulation policy should derive. Results of the study show that the most critical points are located within zone 1 (center) where Leq levels within the range of 70–85 dB were found. Such levels exceed by far the international standard of 65 dB as recommended for ambient noise by the World Health Organization.

2:35

4pNS6. Acoustic performance of roadside barriers in urban environments. Kai Ming Li and Siu Hong Tang (Dept. of Mech. Eng., The Hong Kong Polytechnic Univ., Hung Hom, Kowloon, Hong Kong)

Extensive research has been conducted for assessing the acoustic performance of a barrier in a sub-urban area where the receivers are normally situated at a considerable distance from the barrier. However, in a typical urban community, a noise barrier is frequently erected in front of a tall building. It provides a shadow zone to listeners on the other side of the barrier. The size of this shadow zone depends on the height and position of the barrier away from the building facade. Sound waves are diffracted at the top edge of the barrier before they reach the receivers either directly, or on reflection from the ground and/or from the facade surface. Diffraction and multiple scattering play important roles in determining the sound pressure levels between the facade and barrier. Furthermore, the possibility of a direct wave between the source and the receivers higher up the building requires further evaluations when considering the total sound fields. In this paper, the theory behind the facade–barrier system is outlined. Indoor scale model experiments are conducted to validate theoretical predictions. [Work supported by Research Grants Council of the Hong Kong Special Administrative Region and The Hong Kong Polytechnic University.]

2:50

4pNS7. Noise barrier for the Temuco ByPass Highway. Christian Gerard, Aldo Campos (Control Acustico Ltda., Rogelio Ugarte 1817, Santiago, Chile), and Sergio Belfiore (Tubosider America Latina Las Condes, Santiago, Chile)

Increases in traffic and the emergent Highway Concession Program created by the Public Works Ministry of the Chilean government, have led to the necessity of providing acoustic protection in inhabited areas in which the level of noise exceeds that allowed by the standard. Around the world, the installation of acoustic barriers has been one of the main measures of the controlling traffic noise. This is the case with the Temuco ByPass Highway, an extension of 21 km, located in a rural area typical of the environment and landscape of the South of Chile. A total of 6000 lineal meters of transparent acoustic barrier (Methacrylate) was installed. This barrier is the first one of its type to be installed in South America, and theoretical calculations and practical results are analyzed in this work. (To be presented in Spanish.)

3:05–3:20 Break

3:20

4pNS8. Measurement of urban traffic noise due to heavy cargo transport services. Francisco Ruffa, Daniel S. Gavinowich, Erica A. Barkasz, and Pablo R. Ciccarella (LACEAC, School of Eng., Paseo Colon 850, 1piso Dpto. Electr., 1063 Buenos Aires, Argentina)

The Acoustic and Electroacoustic Laboratory of the Buenos Aires University is working on a Traffic Noise Recommendation to be used in Argentina. To that end, measurements of traffic noise due to heavy cargo transport services were taken in a Buenos Aires neighbourhood, especially asked for by a neighborhood association. The goal of this measurement was to try to define a correct standard set for the annoyance level in order to pass that information to the city government and work together to improve the living standard of the inhabitants. The measured indexes were Leq, L90 and L10, as well as octave measurement focused on frequencies under 250 Hz. This paper will show the results of the measurements and the conclusions that were made.

3:35

4pNS9. Absorption characteristics of porous media: A model for traffic-noise reducing asphalt. Hasson M. Tavossi (Dept. of Physical and Environ. Sci., School of Natural Sci. & Mathematics, Mesa State College, Grand Junction, CO 81501)

This paper presents preliminary investigation into the feasibility of the use of granular materials, with special grain size distribution as the asphalt constituents, for selective absorption of traffic-noise—generated by the engine and tires-asphalt-interactions—of the roads and highways passing near residential areas. The near-term goal of this research is laboratory data collection and modeling on the sound absorption properties of asphalt model-media, selecting media with required absorption properties, and analysis of their traffic-noise absorption characteristics. Long-term efforts are aimed at the general reduction of traffic-noise reflected from the surfaces of these highways by selective absorption of noise of different frequency bands in the granular constituents of asphalt. Reducing noise from the road surfaces by selective sound absorption at the surface results in lowering of the noise level pollution of the streets in the residential areas and the communities situated close to highways.

3:50

4pNS10. Analysis of a pollution problem of ambient noise and bases for a municipality inspection in the center of Talquepaque, Jalisco, Mexico, 2001. Erika Rodriguez, Martha Orozco, and Lourdes Palafox (Div. de Ciencias Biologicas y Ambientales, CUCBA, Universidad de GuadalajaraKm 15 1/2 Carr. a Nogales, Las Agujas Zapopan, Jalisco, Mexico)

The city of Tlaquepaque is delimited at the north by Guadalajara city, at the west side by Zapopan, at the south side by Tlajomulco and El Salto, and in its east side by Tonala. Our purpose in this study is to create the first noise map of Tlaquepaque, being the study area between the streets: Avenida Ninos Heroes, Revolucion, Florida, and Marcos Montero. We take 53 points of which 77.3% are above 65 dBA, the Leq average of the points was 69.7 dBA, having as a minium 44.8 dBA and as a maximum 108.6 dBA. Nowadays Tlaquepaque does not have any noise map, so we are trying to identify its more complex areas. The high dB levels found in the study show us how high the noise in the street is, and this is reflected by several citizen complaints about noise received by the municipality. Because of this it is necessary to have an ambient noise diagnostic program, and do the necessary modifications to the current inspection and monitoring system. The preliminary results of this first noise map show a really serious problem with ambient noise caused by vehicular traffic, bad urbanization, and the existence of commercial centers that raise the ambient noise.

4:05

4pNS11. Sound pollution in schools nearby highways crossing Buenos Aires City. Daniel S. Gavinowich, Francisco Ruffa, Erica A. Barkasz, and Pablo R. Ciccarella (Acoust. Lab., Electron. Dept., School of Eng., Buenos Aires Univ., Paseo Colon 850, 1 Piso, 1063 Buenos Aires, Argentina)

According to an agreement between the concessionary company of the highways crossing Buenos Aires City (AUSA) and the School of Engineering of The Buenos Aires University, measurements of the sound level were made in the neighborhood of primary schools and high schools near the roads. The purpose of this work was to supply registered values to be able to obtain a general knowledge of sound levels ''in front of'' and ''inside'' schools, as reliably as possible. Relying on the previous measurements of some points, the registered values can be compared with the previous ones and possible changes of the sound level produced in the last times can also be estimated. Finally, the work was completed with the study of possibility of acoustic attenuation using noise barriers in critical areas because of high levels of pollution.

4:20

4pNS12. Balanced noise control design: A case study for co-generation power plant. Yong Ma and Salem Hertil (ATCO Noise Management, 1243 McKnight Blvd. NE, Calgary, AB T2E 5T1, Canada)

Power generation plant generally requires noise mitigation treatment to achieve the specified noise regulations. In this paper, a case study of the noise control design for a cogeneration power plant was presented. Major noise sources included two GE gas combustion turbines, two generators, two heat recovery steam generators (HRSGs), one steam turbine and generator, one 12-cell cooling tower, and other accessory equipment. The acoustic modeling software Cadna/A was used to predict the noise contributions from sources. During the acoustic modeling, alternative noise mitigation measures underwent two specific investigations before they were chosen as a noise solution recommendation. The first was to determine the technical feasibility of attenuating the source equipment. The second was to perform a cost benefit analysis, necessary to find the most cost-effective solution. For example, several acoustic wall and roof assemblies were entered into the acoustic model and the acoustic performance of the ventilation system was varied until we were able to arrive at the most economical acoustic solution. This is the premise on which so called balanced design is based.

4:35

4pNS13. Declaration and verification of noise emission values of machinery and equipment in Russia. Ilya Evseevich Tsukernikov (SC ''NIIpolugraphmash,'' 57 Profsoyusnaya St., 117420 Moscow, Russia, odpm@rmail.ru) and Igor Alekseevich Nekrasov (JSC ''Algorithm-Acoust.,'' 107370 Moscow, Russia)

Original positions of the interstate standard GOST 30691-2001 are considered, in which the requirements are established to declaration and verification of noise emission values of machinery and equipment, being manufactured in Russia and imported there from different countries. The comparison of the standard with the International standard ISO 4871:1996 is conducted. The main differences are marked. The suggestions for standards revision are made.

4:50

4pNS14. Study on the correlation between field experiments and laboratory experiments for audible environmental noise. Takeshi Tokashiki (Faculty of Eng., Univ. of the Ryukyus, 1 Senbaru Nishihara cho Nakagami gun Okinawa Pref., Japan), Yasuhiro Yamashita, and Naoki Takagi (Shinshu Univ., 500 Wakasato Nagano shi, 380-0922, Japan)

Our objective in this study is to clarify, by some experiments, that there would be a different response and that subjective appraisal differs when audible sound types change even under the same sound level. So far various methods that we, appraise sound environment by our subjectivity, have been proposed and utilized. But we have seen a few study cases about the correlation between sound types, subjective appraisal, and, moreover, sound level. In this research, we have surveyed environmental noise in terms of data obtained by physical quantity (LAeq), and subjective appraisal and audible noise types, then examined the correlation between them. Consequently we could confirm their audible sound types might influence noise level and subjective appraisal. Moreover, as a result of the surveillance of noise environment along boulevards, commercial, and residential areas, the sound heard in each area constitutes the same ratios. The data obtained by subjective appraisal on each area is almost the same. This fact indicates we might infer an annoying response through its noise types.

5:05

4pNS15. Analysis of Spanish adjectives to develop a Likert scale useful for studies of urban noise annoyance. Rosendo Vílchez-Gómez, Juan Miguel Barrigón-Morillas, Valentín Gómez-Escobar, Juan Antonio Méndez-Sierra (Universidad de Extremadura, Escuela Politécnica, Avda. de la Universidad, s/n-10071 Cáceres, Spain, vilchez@unex.es), and José Manuel Vaquero-Martínez (Universidad de Extremadura, Escuela Politécnica, 10071 Cáceres, Spain)

The present work describes the development and analysis of a Spanish annoyance scale for use in community noise assessments. The use of different descriptors and intervals for different authors prevents comparing studies of assessment of noise reactions. It is not clear how many semantic distinctions can be made to describe subjective reactions to noise. There is also great variability in language use reflecting educational, social class, regional, and subcultural differences. Finally, scaling annoyance responses, when the considered intervals are not equally distributed, may seriously distort results when analyzed with parametric statistics. Using three different questions, almost 400 subjects were asked to rate 25 different Spanish descriptors of annoyance. Our attempt to produce a Spanish standardized annoyance scale with descriptors marking clear semantic distinctions, roughly equidistant from each other, and having wide acceptability is presented here. Taking into account that Spanish is the official language in 21 countries and it is spoken by more than 300 million people, one of our goals in this meeting is to discuss our method and try to generalize it in order to obtain a real, universal Spanish annoyance scale.

Session 4pPA

Physical Acoustics: Global Infrasound Monitoring II

Milton A. Garces, Chair
Infrasound Laboratory, University of Hawaii, Manoa, 73 4460 Queen Kaahumanu Highway, #119, Kailua Kona, Hawaii 96740-2632

Invited Papers

1:15

4pPA1. Wind noise reduction research: Possible lessons for infrasound? Richard Raspet (Dept. of Phys. and Astron., Univ. of Mississippi, P.O. Box 1848, University, MS 38677) and Michael A. H. Hedlin (Inst. of Geophys. and Planetary Phys., Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0225)

Morgan [Ph.D. dissertation, University of Mississippi, 1992] extended the theory of wind noise reduction by spherical screens proposed by Phelps [RCA Rev. **3**, 203–212 (1938)] to modern porous foam windscreens and achieved good predictions of the large ($>$12 dB) low frequency reduction afforded by spherical screens even when the turbulence scale is much greater than the wind screen size. In this paper the findings of Morgan are reviewed as background for recent measurements of the infrasonic wind noise reduction of a wind barrier design suggested by Ludwik Liska. The data suggest that significant reductions in infrasonic wind noise can be achieved with devices smaller than a turbulence scale size.

1:35

4pPA2. Large meteoroid detection using the global International Monitoring System infrasound system. Douglas O. ReVelle (Los Alamos Natl. Lab., P.O. Box 1663, MS J577, Los Alamos, NM 87545, revelle@lanl.gov)

We will review the subject of infrasound from large bolides (large meteor-fireballs) entering the atmosphere at hypersonic speeds and their expected rate of detection by the 60 infrasonic arrays of the global IMS network (International Monitoring System). This will include the details of the generation of a quasiline source blast wave and its subsequent decay for near-continuum flow conditions. We will also discuss new highly refined models of bolide ablation and fragmentation and of known compositional types and their effect on sound and light production. In addition, we will consider the effects of refraction of the waves by the middle atmospheric and tropospheric thermodynamic sound speed and horizontal wind profiles in a range-independent atmosphere so that the characteristic velocity and wave normal directions radiated at the source are conserved during the propagation. Next, we will discuss the detection of the signals and their interpretation in terms of plane wave arrivals regarding the 3-D source location (latitude, longitude, height), the source energy level, etc. Finally, we will use the infrasound data from bolides to estimate the expected steady-state global influx rate, including formal errors, as a function of their observed source energy. Infrasound from recent large events will also be examined.

1:55

4pPA3. Global infrasound monitoring—Research issues. Henry E. Bass, Kenneth Gilbert (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, P.O. Box 1848, University, MS 38677), Milton Garces, Claus Hetzer (Univ. of Hawaii, Kailua-Kona, HI 96740), Gene Herrin, Paul Golden (Southern Methodist Univ., Dallas, TX 75275), John Olson, Daniel Osborne, Curt A. L. Szuberla, Charles Wilson (Univ. of Alaska, Fairbanks, AK 99775), Jon Berger, Michael Hedlin (Scripps Inst. of Oceanogr., La Jolla, CA 92093), Rod Whitaker, Doug Revelle (Los Alamos Natl. Lab., Los Alamos, NM 87545), Bob Woodward, Bob North (Ctr. for Monitoring Res., Arlington, VA 22209), and Richard Kromer (Sandia Natl. Lab., Albuquerque, NM 87185)

The International Monitoring System being installed to support monitoring compliance with the Comprehensive Nuclear Test Ban Treaty provides scientists with a unique opportunity for research. There are still a number of problems which limit the full exploitation of the system. These include limitations on signal-to-noise imposed by wind noise and the absence of well defined, internationally accepted calibration standards for sensors. But perhaps the major research challenges lie in the area of source characterization and definition. Most of the signals recorded at the few sites now operating come from unidentified sources. There has been some effort devoted to identifying local and regional sources but the unidentified category still exceeds 50% of all distinct events. There are a number of infrasound sources that occur naturally. These include volcanic eruptions, bolides, microbaroms, mountaintop/wind interactions, severe storms, and earthquakes. Manmade sources include most any energetic activity. After sorting out all these sources, there may remain more exotic sources of infrasound not yet identified. Therein lies a major source of excitement.

2:15

4pPA4. Evaluation of rosette infrasonic noise-reducing spatial filters. Michael A. H. Hedlin (Inst. of Geophys. and Planetary Phys., Univ. of California, San Diego, San Diego, CA), Benoit Alcoverro (Commissariat a l'Energie Atomique, Departement Analyse et Surveillance de l'Environement), and Gerald D'Spain (Scripps Inst. of Oceanogr., Univ. of California)

The spatial noise filter currently preferred for use at the new International Monitoring System (IMS) infrasound array sites consists of an array of low-impedance inlets connected by solid tubes to a microbarometer. We present results from recent tests of ''rosette'' infrasonic noise-reducing spatial filters at the Pinon Flat Infrasound test-bed in southern California. At wind speeds up to 5.5 m/s, the 96 inlet 18-m rosette filter reduces wind noise levels above 0.2 Hz by 15–20 dB. Under the same conditions, the 144 inlet 70-m rosette filter provides noise reduction of up to 15–20 dB between 0.02 and 0.7 Hz. Standing wave resonance inside the 70-m filter degrades the reception of acoustic signals above 0.7 Hz. Synthetics accurately reproduce the noise reduction and resonance observed in the 70-m filter at all wind speeds above 1.25 m/s. Experiments with impedance matching capillaries indicate that internal resonance in the rosette filters can be removed. Rosette filters are tuned to vertically incident energy. Attenuation of signals by the 70-m rosette filter at frequencies above 3.5 Hz arriving at grazing angles of <15 deg from the horizontal are predicted to range upward from 10 dB to total cancellation at 5 Hz.

2:30

4pPA5. An optical fiber infrasound sensor. Mark A. Zumberge, Jonathan Berger, and Michael A. H. Hedlin (Univ. of California, San Diego, La Jolla, CA 92093-0225)

A new sensor for detecting infrasonic signals from distant sources has been designed, tested, and deployed. The instrument consists of a long (of order 100 m), compliant, tubular diaphragm tightly wrapped with optical fibers. The air-filled diaphragm is sealed from the ambient atmosphere and therefore expands and contracts radially along its length as the pressure changes. An optical system monitors minute variations in the optical path length of the fibers, which change as the tube expands and contracts. The new instrument responds to the spatially integrated pressure average along its length. As it relies on optical interferometry, it does not suffer from the propagation delays inherent to mechanical noise filters and as a result can be made very long. Above 1 Hz, the optical fiber infrasound sensor (OFIS) is less noisy than sensors relying on mechanical filters. The records collected from an 89 m long sensor indicate a new low noise limit in the band from 1 Hz to 10 Hz. Current research focuses on the relationship between sensor length and wind-induced noise, on the utility of the sensor as a directional infrasound antenna, and on the sensor's performance at audio frequencies.

2:45–3:00 Break

3:00

4pPA6. A study of microbarom wave packets. John V. Olson, Curt A. L. Szuberla, and Charles R. Wilson (Geophysical Inst., Univ. of Alaska, Fairbanks, AK 99775)

Microbaroms are a class of atmospheric infrasound that is characterized by narrow-band, nearly sinusoidal waveforms with periods near 5 s. These waves are generated by marine storms and are observed at infrasound sites across the globe. We report here on a study of microbarom wave packets. The data analyzed were obtained using the University of Alaska infrasound array comprised of 4 microphones located in central Alaska. Exploiting the narrow-band feature of the microbarom signals we are able to apply the Hilbert transform as a method for finding phase breaks in the signal. The phase breaks are interpreted to be the demarcations of the boundaries of individual wave packets. When applied to long sequences of microbaroms we find a broad distribution of packet lengths that diminishes monotonically with length, with a mean near 10 cycles and a variance nearly as large. Once the packets are identified, high-resolution estimators have been used to obtain the distribution of arrival directions at the array site. The ensemble of arrival directions is a reflection of the possibility of multiple sources and/or multiple propagation paths between the sources and the array.

3:15

4pPA7. Observations of infrasonic signals from natural and manmade sources. John M. Noble, Stephen M. Tenney (Army Res. Lab., 2800 Powder Mill Rd., Adelphi, MD 20783, jnoble@arl.army.mil), R. W. Whitaker, and D. O. Revelle (Earth and Environ. Sci. Div., Los Alamos Natl. Lab., Los Alamos, NM 87545)

We report on data recorded from small aperture (20-m) infrasonic microphone arrays. These data will include the well-observed Pennsylvania bolide of 23 July 2001, space shuttle launches, sounding rocket launches, and local severe storm passages. Data were obtained from two infrasonic arrays in southern Maryland. Analysis results from two independent packages will be presented, including bearing estimates and detector performance. Bolide detection results will be correlated with eyewitness reports of the flight trajectory. Estimates will be made of the bolide source energy from the application of the historical explosion scaling law relating the period at maximum acoustic amplitude. Detection of sounding rockets from over 150 km will be discussed. Finally, data from the passage of several severe storms will be presented showing the detection and tracking of a low-frequency acoustic signature of the storm.

3:30

4pPA8. Detection and interpretation of infrasonic signals observed in Hawaii. Milton Garces and Claus Hetzer (Infrasound Lab., Univ. of Hawaii, Manoa, 73-4460 Queen Kaahumanu Hwy. #119, Kailua-Kona, HI 96740-2632)

Various array processing and signal detection algorithms have been tested and evaluated at International Monitoring System infrasound array I59US, also known as KONA, located on the western side of the Big Island of Hawaii. The array consists of four elements deployed as a triangle with a central element, with a baseline of 2 km. Initial tests were made by defining multiple beams and running STA/LTA, coherence, and F-statistic detectors to estimate the apparent horizontal phase velocity and incidence azimuth of arrivals in frequency–wavenumber space. Such methods generally proved effective only for infrasonic signals with signal-to-noise ratios (S/N) greater than ~ 4, and their performance varied with signal frequency. The Progressive Multi-Channel Correlation (PMCC) detector is a time–domain detector that uses the correlation between various groupings of three sensors to obtain an estimate of the consistency of specified closure relations. If the consistency is below a certain threshold, a detection is registered. This detector has performed well for S/N ratios that are close to unity and for all frequencies. Examples of infrasonic detections in Hawaii for small (Leonids 2001) and large bolides (23 April 2001), blasting activity, surf, thunder, severe storms, and low-frequency signals of unknown origin are presented.

3:45

4pPA9. Co-located atmospheric infrasonic and seismic recordings of rocket launches. Gerald L. D'Spain, Michael A. H. Hedlin, John A. Orcutt, William A. Kuperman, Catherine de Groot-Hedlin, Galina L. Rovner, and Lewis P. Berger (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 93940-0704)

Atmospheric infrasound data and co-located, three-component seismic data have been collected by the eight microbarometers of the International Monitoring System (IMS) station and the IRIS seismic station at Pinon Flat (PFO) plus five additional microbarometer/space filter systems at five Anza seismic stations located within 40-km range of PFO in Southern California. Characteristics of the infrasound and seismic recordings from this large-horizontal-aperture array of signals from 400-km-distant rocket launches at Vandenberg Air Force Base are analyzed using waveguide invariant theory. The Navy standard Gaussian Ray Bundle (GRAB) underwater acoustic propagation code (with slight modifications), along with launch trajectory information and atmospheric data collected at the time of the launches, is used to to examine the predictability of the signal arrival structure. The predictions take into account the signal-distorting effects caused by phase delays across the spatial aperture of the space filters, which cause each infrasound array element to be directional over the frequency band of interest. [Work supported by the Defense Threat Reduction Agency.]

4:00

4pPA10. Propagation variability and localization accuracy of infrasonic networks. David E. Norris and Robert G. Gibson (BBN Technologies, 1300 N. 17th St., #1200, Arlington, VA 22209)

Prediction of propagation variability induced by the environment is used to evaluate the localization performance of infrasonic networks. The dominant source of variability affecting infrasonic propagation is believed to result from gravity waves. A gravity wave spectral model based on scale-independent diffusive filtering is used to generate multiple wind perturbation realizations. A Monte Carlo simulation is executed where rays are traced through the perturbed environmental fields, and uncertainty in ray travel time and azimuthal deviation is calculated. The propagation uncertainties, along with uncertainty in infrasonic measurements, are then used to compute 90 percent confidence bounds of multi-station event localizations. Infrasonic data from the April 2001 Pacific bolide event are used to compute the performance of a five station network, and the network localization is compared to that found from satellites. [Work sponsored by Defense Threat Reduction Agency, Contract No. DTRA01-00-C-0063.]

4:15

4pPA11. Long-range infrasound propagation modeling using updated atmospheric characterizations. Robert G. Gibson and David E. Norris (BBN Technologies, 1300 N. 17th St., #1200, Arlington, VA 22209, rgibson@bbn.com)

Infrasonic waves can propagate thousands of kilometers in range and sample regions of the atmosphere from the ground up to and including the thermosphere. Conventional infrasound propagation modeling techniques rely on climatological models of mean temperatures and winds to characterize the environment. However, temperature and wind vary over temporal and spatial scales that are not captured by climatological models. Recent work addresses the integration of infrasound propagation models, such as three-dimensional ray tracing, with numerical weather prediction models, such as the Navy Operational Global Atmospheric Prediction System (NOGAPS). Propagation results are computed using both climatological and updated atmospheric characterizations, and comparisons are presented. Implications for global infrasound monitoring are discussed. [Work supported by the Defense Threat Reduction Agency.]

4:30

4pPA12. Determination of uncertainty in the estimation of velocity and direction-of-arrival for atmospheric infrasound signals. Curt A. L. Szuberla and John V. Olson (Geophysical Inst., Univ. of Alaska, 903 Koyukuk Dr., Fairbanks, AK 99775-7320)

Upon the detection of an atmospheric infrasound signal, the problem of precisely estimating the signal's velocity (v) and direction-of-arrival (θ) arises. Multiple sources, multipath, medium anisotropies, and other propagation effects can all degrade precision; however, uncertainty in the estimates of v and θ is fundamentally governed by array geometry and the estimation of time delays across the array. Typically, as in the Comprehensive Test Ban Treaty Organzation Provisional Technical Secretariat (CTBTO/PTS) specification for data from infrasound stations, the Cramér–Rao lower bound is invoked to ascertain the uncertainties associated with v and θ. As this theoretical lower limit is often overly conservative, a more general, and useful, approach to calculate these uncertainties is developed. Examples of this uncertainty determination are presented for typical impulsive and continuous atmospheric infrasound signals received at arrays in Windless Bight, Antarctica and Fairbanks, Alaska. Since the determination of v and θ serves as primary input to any propogation model, it is critical that uncertainties in these estimates be addressed. As an extension of this work, an interactive graphical tool is constructed to assist in the analysis of performance bounds for arbitrary array geometries and signal characteristics.

Session 4pSAa

Structural Acoustics and Vibration: Reconstruction of Acoustic Radiation from Vibrating Structures II

Earl G. Williams, Chair
Naval Research Laboratory, Code 7137, Washington, D.C. 20375

Chair's Introduction—1:00

Invited Papers

1:05

4pSAa1. A stopping rule for the conjugate gradient regularization method for ill-posed problems. Thomas DeLillo and Tomasz Hrycak (Dept. of Mathematics and Statistics, Wichita State Univ., 1845 N. Fairmount, Wichita, KS 67260, hrycak@math.twsu.edu)

We present a novel parameter choice strategy for the conjugate gradient regularization algorithm which does not assume *a priori* information about the magnitude of the measurement error. Our approach imitates the truncated singular value decomposition within the Krylov subspaces associated with the normal equations. Conjugate gradient is implemented using the Lanczos bidiagonalization process with reorthogonalization. We compare our method with one proposed by Hanke and Raus and illustrate its performance with numerical experiments, including an inverse problem of acoustic source detection.

1:35

4pSAa2. Multipole error estimation of an acoustical boundary-element solver combined with a dual-layer CHIEF approach. Martin Ochmann (TFH-Berlin, Univ. of Appl. Sci., Luxemburger Strasse 10, D-13353 Berlin, Germany, ochmann@tfh-berlin.de) and Alexander Osetrov (St. Petersburg State Electrotechnical Univ., 197376 St. Petersburg, Russia)

The sound radiation from vibrating structures is studied using an iterative boundary-element method. The starting point is a self-adjoint formulation of the Helmholtz integral equation. The generalized minimum residual method (GMRES) is used for solving the corresponding system of equations. Three aspects of the method are investigated. First, the accuracy of the iterative solver is checked by performing multipole error tests: The body is forced to vibrate with the corresponding surface normal velocity of a multipole. Hence, the radiated sound field is analytically known and can be used for determining the exact error caused by the boundary-element solver. Second, instabilities at irregular frequencies are suppressed by combining the iterative solver with a special variant of the CHIEF method, where the total boundary of the body is retracted to an auxiliary surface lying totally within the interior of the structure. CHIEF points can be placed on such an auxiliary surface. The effectiveness of this combined dual-layer GMRES–CHIEF approach and of the Burton–Miller method will be compared. Third, the radiation from a structure with mixed boundary conditions will be investigated. One part of the surface is coated with an absorbing impedance, another part is vibrating with a prescribed normal velocity.

2:05

4pSAa3. Iterative solution techniques for boundary element method (BEM) equations. Steffen Marburg, Stefan Schneider, and Has-Juergen Hardtke (Institut fuer Festkoerpermechanik, Technische Universitaet, D-01062 Dresden, Germany, marburg@ifkm.mw.tu-dresden.de)

Boundary-element discretization of the Kirchhoff-Helmholtz integral equation gives rise to a linear system of equations. This system may be solved directly or iteratively. Application of direct solvers is quite common but turns out to be inefficient for large scale problems with 10 000 unknowns and more. These systems can be solved on behalf of iterative methods. This paper is dedicated to testing performance of four iterative solvers being the Restarted Bi-Conjugate Gradient Stabilized algorithm (RBiCGStab), the Conjugate Gradient method applied to the normal equations (CGNR), the Generalized Minimal Residual (GMRes), and the Transpose Free Quasi Minimal Residual (TFQMR). For that, it is distinguished between internal and external problems. Performance of iterative solvers is investigated with respect to different parameters. It turns out that sophisticated preconditioning is required for the hypersingular external Burton and Miller formulation of nonsmooth meshes.

Contributed Paper

2:35

4pSAa4. A review optimization technique in structural acoustics. Steffen Marburg (Institut fuer Festkoerpermechanik, Technische Universitaet, D-01062 Dresden, Germany, marburg@ifkm.mw.tu-dresden.de)

In recent years, several activities on structural-acoustic optimization have been reported. The problem of structural-acoustic design optimization is characterized by nonlinearities. For that reason, it can only be dealt with numerically. Furthermore, the calculation of the objective function is usually very time-consuming since a multifield boundary value problem has to be solved. This paper summarizes results of a longer review paper. It will briefly discuss fast analysis techniques like FEM and BEM, objective functions, sensitivity analysis and optimization methods. Applications and parameter choices are categorized.

Session 4pSAb

Structural Acoustics and Vibration: Measurement and Analysis Techniques

Gerard P. Carroll, Chair
Naval Surface Warfare Center, Carderock Division, Code 7250, 9500 MacArthur Boulevard, Bethesda, Maryland 20817-5000

Contributed Papers

3:00

4pSAb1. Identification of very closely spaced modes using an iterative algorithm. Matt S. Allen, Jerry H. Ginsberg, and Aldo Ferri (G. W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

The algorithm of mode isolation is a frequency domain method for processing measured response data in order to identify the modal properties of a system. It has been shown [M. V. Drexel and J. H. Ginsberg, Proceedings of the 19th IMAC, Orlando, FL, 5–8 February 2001] to accurately evaluate a pair of damped modes whose natural frequencies differ by an amount that is commensurate with the bandwidth of either mode. Here, SIMO measurement of a two-degree-of-freedom system is simulated analytically by adding white noise to the computed response. This system is selected because the natural frequency difference can be made as small as desired by adjusting a parameter that has little effect on the modal damping ratios in the range of interest. Analytically, the two normal modes are unique and mutually orthogonal if the frequency difference is nonzero, while repeated frequencies lead to two arbitrary, and not necessarily mutually orthogonal, modes. The present work explores the degree to which AMI tracks the analytical behavior, specifically, its ability to detect and characterize the modes with decreasing frequency difference.

3:15

4pSAb2. The effect of high spatial resolution on techniques for experimental determination of loss factor. Gerard P. Carroll, Matthew A. Craun, and Daniel T. DiPerna (Naval Surface Warface Ctr. Carderock Div., 9500 MacArthur Blvd., Bethesda, MD 20817-5700, carrollgp@nswccd.navy.mil)

Using a common set of mobility functions, various methods for determining loss factor are compared. A scanning laser Doppler vibrometer was used to measure these mobility functions with high spatial resolution for a plate and beam sample. The advantages of high spatial resolution are explored by estimating the loss factor using the complete set of mobility functions, and then comparing the results to those obtained using smaller subsets of the measured data. The plate and beam specimens are treated with a free layer damping. The loss factor determination methods include: synthesized time decay, classical modal analysis methods, the power input method, and complex wave number analysis. The merits/limitation of the damping measurement methods and the possible advantages of high spatial resolution are compared, along with results for both test specimens.

3:30

4pSAb3. Autocorrelation of vibration data for flaw detection. Sally J. Pardue and Brahmaji Vasantharao (Dept. of Mech. Eng., Tennessee Technolog. Univ., Box 5014, Cookeville, TN 38505, spardue@tntech.edu)

The autocorrelation of a structural FRF (frequency response function) can be used to extract structural parameters, such as length from a transversely vibrating beam. Further work has shown that the technique can also detect flaws in the structure. The frequency spacing between resonant peaks is extracted from the autocorrelation process and relates to the length of the object. Indicators of flaws in the structure were observed in the autocorrelation values plotted against a length axis. The first test structure was a free–free nine feet long, one inch diameter aluminum rod. The flaws examined were machined slots of various widths cut into the rod. The concept was further explored numerically in ANSYS models of the same aluminum bar using much smaller flaw sizes. Now, the concept of autocorrelation of the FRF is being applied to composite beams, 8 ply, 10 in. long, 0.052 in. thick, 1 in. wide. Indicators in the vibration data are sought to relate to porosity levels of the manufactured composite beams.

3:45

4pSAb4. Phase determination of reverberant structural-acoustic systems using pole and zero distribution. Juan F. Betts (The Aerosp. Corp., P.O. Box 92957-M4/907, Los Angeles, CA 90009, Juan.F.Betts@aero.org) and Christopher Fuller (Virginia Tech, Blacksburg, VA 24061)

This study investigates the phase behavior of a simple coupled structural acoustic system with high modal overlap. It proposes that phase is simply accumulated across the components making up the structure. Classical dynamic solutions breakdown with high modal overlap and complex geometries, and standard Statistical Energy Analysis (SEA) solutions work well for high modal overlap but give no phase information. Many applications such as active control predictions and time-domain signal analysis require phase information. Results of the investigation show that the amount of damping in the system is very important in the estimation of reverberant phase. The amount of energy lost due to coupling with adjacent subsystems was also demonstrated to be important for both acoustic and structural subsystems. The reverberant phase of an acoustic subsystem was found to be affected more by the inclusion of coupling loss to an adjacent structure than the reverberant phase of a structure attached to an acoustic subsystem. The phase of the overall system is calculated by simply accumulating the phase of each of the subsystems.

4:00

4pSAb5. Communications/Navigation Outage Forecasting System (C/NOFS) statistical energy analysis. Juan F. Betts (The Aerosp. Corp., P.O. Box 92957-M4/907, Los Angeles, CA 90009, Juan.F.Betts@aero.org)

The objective of this study was the assessment of the dynamics response of the Communications/Navigation Outage Forecasting System (C/NOFS) spacecraft due to incident acoustic environments. C/NOFS is space vehicle (SV) launched using a Pegasus launch vehicle (LV). A statistical energy analysis (SEA) model of the SV was developed from a NASTRAN finite-element model (FEM). The study compared the dynamic response of the structural subsystems (panels and structural elements) due to incident acoustic Pegasuss protoflight environment [which is 3 dB above maximum expected flight environment (MEFL)] and compared it against the Mil-340 standard workmanship acoustic levels. A concern was raised as to whether the workmanship acoustics, required by Mil-340, imposed on C/NOFS may induce higher vibration responses than what C/NOFS was qualified for, and may result in unnecessary increases in component test levels. The two incident acoustic spectrums were introduced as reverberant loads on the model. The PSD results at the structural panels were compared to random vibration test levels performed on the components for C/NOFS.

4:15

4pSAb6. Effects of circumferential air gaps on the measurement of the absorption coefficient of poroelastic materials. Dominic Pilon and Raymond Panneton (GAUS, Dept. of Mech. Eng., Universite de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada, dominic.pilon@hermes.usherb.ca)

The influence of mounting conditions on the measurement of the absorption coefficient is investigated. More specifically, the effects of circumferential air gaps on poroelastic materials inside the standing wave tube are studied. The objective is to identify the materials, in terms of a ratio based on their physical properties, for which it is possible to measure the theoretical absorption coefficient using the tube. The difference, or error, between the measured and theoretical absorption coefficient is evaluated for a wide range of materials with various sample sizes. These errors are then sorted in terms of the chosen ratio. It is shown that for certain values of the ratio, the theoretical absorption coefficient can be efficiently measured using the standing wave tube. Through the use of this acousto-visco-inertial criteria, an experimenter can determine which absorption is going to be measured: the theoretical absorption coefficient or one that will be influenced by either the circumferential air gaps or the size of the sample.

4:30

4pSAb7. Experimental investigation of sound transmission through thin-walled structures due to nonuniform pressure–fluctuation fields. Part 2: Supersonic flow. Boris M Efimtsov, Alexey Yu Golubev, Alexander Ya Zverev (TsAGI, Acoust. Div., 17 Radio St., 107005 Moscow, Russia), Anders O. Andersson, and Sergey V. Kravchenko (Boeing, Seattle, WA 98124)

Theoretical models of the sound–energy transmission through thin-walled structures from nonuniform, convecting pressure–fluctuation fields were validated through experiments. The fields were created with separated flows at forward-facing and backward-facing steps and an airfoil as the shock generator. Thin plates and model fuselage panels were used and vibrations and acoustic radiation of the structures were investigated. The data permit a direct comparison with predicted results as well as an evaluation of the validity of the prediction-model assumptions. In particular, the elastic and inertial properties of the skin but not of the stiffeners need to be taken into account. The experiments support the validity of previously reported theoretical predictions.

4:45

4pSAb8. Transmission loss across a rectangular partition or aperture. Colin Fox and Hyuck Chung (Mathematics Dept., The Univ. of Auckland, PB 92019, Auckland, New Zealand)

We give closed-form solutions for sound propagation through a rectangular aperture, or plate, in a finite or infinite acoustically opaque barrier. This model is useful in descibing airborne excitation and reradiation through window openings or through floor/wall partitions that can be modeled in terms of bending stiffness and mass. While mode-matching solutions are routinely available for these problems, we are not aware of any previous analytic solutions. The value of the closed-form solution, other than the obvious ease of calculation, is in establishing scaling laws for the various regimes of structure-borne and air-borne sound. We derive the analytic solution using an extension of the Wiener–Hopf technique as applied to sound propagation in ducts with a partial rigid barrier.

THURSDAY AFTERNOON, 5 DECEMBER 2002 GRAND CORAL 3, 1:00 TO 5:00 P.M.

Session 4pSC

Speech Communication: Spanish and English in Contact and Other Cross-Language Studies (Lecture/Poster Session)

Winifred Strange, Cochair
PhD Program in Speech and Hearing, City University of New York Graduate School, 365 Fifth Avenue, New York, New York 10016

Ann R. Bradlow, Cochair
Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, Illinois 60208

Chair's Introduction—1:00

Invited Papers

1:05

4pSC1. Sensitivity to voiceless closure in the perception of Spanish and English stop consonants. Mary L. Zampini (Dept. of Spanish & Portuguese, Univ. of Arizona, P.O. Box 210067, Tucson, AZ 85721), Constance M. Clarke, and Linda W. Norrix (Univ. of Arizona, Tucson, AZ 85721)

The duration of voiceless closure that precedes the release of a stop consonant is a temporal cue that, like voice onset time (VOT), varies across languages. This talk will examine the interaction between VOT and voiceless closure in Spanish and English and will focus on monolingual and bilingual listeners sensitivity to changes in the duration of voiceless closure during perception. Experimental data will show that monolingual Spanish listeners mean VOT boundaries decrease as the duration of voiceless closure increases. This pattern is consistent with the finding that Spanish speakers produce voiceless stops with longer voiceless closure durations than voiced stops [K. P. Green, M. L. Zampini, and J. Magloire, J. Acoust. Soc. Am. **102**, 3136 (1997)]. It will also be shown that

monolingual Spanish listeners show greater sensitivity to voiceless closure than monolingual English listeners. Lastly, the impact that experience with both languages has on perception by late English–Spanish bilinguals will be discussed. It will be shown that the bilinguals under study are affected by their first language (English) while listening to tokens in isolation, but perform like monolingual Spanish listeners when listening to tokens in a Spanish sentence context.

1:35

4pSC2. Solving conflicts between vowel systems. Paola Escudero (Inst. of Linguist. UiL OTS, Utrecht Univ., 10 Trans, Utrecht 3512 JK, The Netherlands, Paola.Escudero@let.uu.nl)

Conflicts may arise when listeners encounter a language with a different number of vowels. For instance, Spanish has 5 vowels, English at least 9 and Spanish listeners have many difficulties perceiving the /i/–/I/ contrast in some (though not all) varieties of English. Spanish learners of English solve this few-to-many problem by incorporating a new auditory dimension (vowel duration) into their categorization system. Perhaps surprisingly, the reverse scenario, namely encountering a language with fewer vowels, is also problematic. For instance, Dutch listeners of Spanish (Dutch is comparable to English in its vowel inventory size) have a boundary mismatch when listening to the Spanish /a/–/o/ contrast, and they hear the extra category /I/ when listening to Spanish /i/–/e/. Dutch learners of Spanish shift their boundaries toward what would be appropriate for the target language and they gradually stop perceiving the extra category. It turns out that both in few-to-many and many-to-few situations learners are capable of changing their perception to become like or almost like the native speakers of the target language. This can be explained within a formal model of speech perception and its development.

2:05

4pSC3. Development of the Hearing In Noise Test (HINT) in Spanish. Sigfrid D. Soli, Andrew Vermiglio (House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057, ssoli@HEI.org), Karen Wen (House Ear Clinic, 2100 W. 3rd St., Los Angeles, CA 90057), and Carolina Abdala Filesari (Cochlear Corp. Latin America, Caracas, Venezuela)

Assessment of functional hearing ability is relevant to clinical outcome assessments, occupational health evaluations, and forensic applications. Communication with spoken language is a primary aspect of functional hearing ability. Comparable speech tests in multiple languages are required to make these assessments in a multilingual population. This presentation will report on an ongoing international research project to develop a Latin American Spanish version of the Hearing In Noise Test (HINT) for this purpose. The methods of selecting speech materials, recording the materials, synthesizing appropriate masking noise, equating the difficulty of the materials, forming sentence lists, norming the lists, and determining reliability and sensitivity will be discussed. Samples of Spanish, as well as other languages (English, Japanese, Mandarin, Cantonese, and Canadian French) will be used to demonstrate the procedures by which standard dialects of each language were selected. Cross-language comparisons of the spectral and temporal characteristics of the speech materials will be presented. Norms, reliability coefficients, and measurement errors for each language will also be reported. Methods for comparing and/or equating functional hearing ability across languages will be described. Finally, procedures for measuring functional hearing ability will be discussed.

2:35–2:40 Break

Contributed Papers

All posters will be on display from 2:40 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 2:40 p.m. to 3:40 p.m. and contributors of even-numbered papers will be at their posters from 3:40 p.m. to 4:40 p.m.

4pSC4. Within- and across-language acoustic variability of vowels spoken in different phonetic and prosodic contexts: American English, North German, and Parisian French. Winifred Strange (Speech and Hearing Sci., City Univ. of New York—Grad. Ctr., 365 Fifth Ave., New York, NY 10016-4309), Andrea Weber, Erika Levy, Valeriy Shafiro, and Kanae Nishi (City Univ. of New York—Grad. Ctr., New York, NY 10016-4309)

This project investigates the acoustic variability of vowels produced in multisyllabic nonsense words /cvC1VC2(v)/ in carrier sentences by speakers of American English (AE), Parisian French (PF), and North German (NG). Variables under examination are (1) immediate phonetic context (C1=b,d; C2=b,d,p,t), (2) sentence prominence (narrow focus versus postfocus), and speaking rate (normal vs rapid). Preliminary results show that for AE vowels, phonetic context produces large differences (2–3 bark) in midpoint $F2$ values for mid-to-high back vowels, reflecting allophonic fronting of these vowels in alveolar context. Much smaller increases in $F2$ ($\sim$1 bark) were found for both PF and NG mid/high back vowels in alveolar contexts. Most AE speakers showed little or no effect of sentence focus or speaking rate on the degree of fronting of back vowels in alveolar context. PF speakers produced more fronted mid/high back vowels in postfocused context than did NG and AE speakers. NG and PF speakers showed relatively greater F1 undershoot in low vowels in postfocus context than for AE speakers. These within- and across-language differences in the effect of prosodic and phonetic context on vowel acoustics reflect both language-universal constraints and language-specific rules for contextual warping of vowel spaces.

4pSC5. Perception of word-final /t/, /d/, /s/, and /z/ by beginning Spanish-speaking learners of English. Roberta Pikser (Ph.D. Prog. in Linguist., CUNY Grad. Ctr., 365 Fifth Ave., New York, NY 10016) and Winifred Strange (CUNY Grad. Ctr., 365 Fifth Ave., New York 10016)

Previous research has indicated that speakers of Spanish learning English have problems perceiving and producing the word-final /t/–/d/ contrast [Flege, 1992]. In this study we examined whether beginning learners

of English from Latin America could distinguish word-final voicing and manner contrasts among /d/, /t/, /s/, and /z/ when produced in three different vowel contexts by two speakers. As expected, subjects had difficulty correctly discriminating the voicing distinction in both stops and fricatives. On average, performance was poorer on /t/–/d/ than /s/–/z/. Subjects had no difficulty distinguishing voiced or voiceless consonants differing in manner. There was a clear vowel effect; performance was poorer on average for voicing contrast in the high front vowel context than for low central and high back contexts, although there were individual differences. There was also a strong speaker effect; subjects more easily identified the /d/–/t/ contrast of the male speaker, who had longer closure durations and stronger release bursts than those of the female speaker.

4pSC6. Spanish listeners' use of vowel spectral properties as cues to post-vocalic consonant voicing in English. Geoffrey Stewart Morrison (Dept. of Linguist., Univ. of Ottawa, 70 Laurier Ave. E., P.O. Box 450, Stn. A, Ottawa, ON K1N 6N5, Canada, geoff@japan.co.jp)

Mexican Spanish listeners who had just arrived in an Anglophone region of Canada were tested on an edited-natural-speech continuum for Canadian English /bit bɪt bid bɪd/ in which vowel duration and vowel spectral properties were varied, and in which consonant closures were silent. The Mexican Spanish listeners did not use vowel spectral properties to identify the vowel: Vowel identification was near chance level with a tendency for longer vowel stimuli to be identified as /i/. However, they did use vowel spectral properties to identify the consonant: Stimuli containing vowels with low $F1$ were identified as having voiceless consonants, and stimuli containing vowels with high $F1$ were identified as having voiced consonants. This presentation will consider possible explanations for this identification pattern for consonant voicing, and will also present the results of additional tests conducted with the same participants, namely: an identical perception test conducted 6 months after the participants' arrival in Canada, English production tests using the same words conducted at the same time as the two perception tests, and a Spanish production test using the words /bit bid bíti bídi bití bidí/ conducted at the same time as the first set of English tests.

4pSC7. Could lack of experience with a second language be modeled as a hearing loss? Monica Padilla (Dept. of Biomed. Eng., University Park Campus OHE 500, Univ. of Southern California, Los Angeles, CA 90089) and Robert V. Shannon (House Ear Inst., Los Angeles, CA 90057)

Previous studies have shown that listeners who learned a second language later in life have poorer speech recognition compared to native listeners, particularly under difficult listening conditions. We are interested in quantifying this demonstrated deficit experienced by non-native listeners as a function of the length of the exposure to the second language. In particular, we are interested in modeling it in terms of Plomp's D factor. In Plomp's model [J. Speech Hear. Res. **29**, 146–154 (1986)], D is defined as a hearing loss due to distortion. We tested listeners whose first language is Spanish with English phonemes, words and sentences in noise, and with reduced spectral information. Preliminary results showed that D increased with an increased loss of spectral resolution. D also increased with the age of immersion in the second language. These results suggest that a lack of experience with the second language may be modeled as a type of hearing loss. Other issues of interest to us in this study are: the effect of the conflicting vowel spaces of the two languages in vowel recognition and the use of context by non-native listeners compared to native English listeners. [Work supported by NIDCD.]

4pSC8. Nonsense syllable perception by monolingual and bilingual English speakers. Catherine L. Rogers (Dept. of Commun. Sci. & Disord., Univ. of South Florida, 4202 E. Fowler Ave., PCD1017, Tampa, FL 33620), Maria R. Brea-Spahn, Mei-Wa Tam, and Edna E. Nyang (Univ. of South Florida, Tampa, FL 33620)

Understanding speech in demanding environments is essential for daily communication. Previous research has shown that even highly proficient bilinguals may experience greater difficulty than monolinguals in understanding speech in noise. In the present study we further address this issue by examining the effects of varying task demand, fatigue, and practice on speech perception by bilinguals. One group of monolingual English listeners and three groups of Spanish–English bilinguals (bilingual since childhood, bilingual since teenager, and bilingual since adulthood) listened to nonsense syllables presented in noise and at increasing presentation rates. Listeners twice completed the two speech perception tasks on two days of testing. On one day the speech tasks were preceded by and on the other day followed by approximately 40 minutes of testing on non-speech auditory tasks. Monolingual and bilingual listeners' overall performance, their performance across the two days of testing (practice effect), and their performance across test orders (fatigue effect) will be compared. Preliminary results from two monolingual and two highly proficient bilinguals (bilingual since age 16 or earlier) show similar overall performance on both speech tasks and no effect of fatigue (test order), but suggest a greater practice effect for bilinguals' perception of syllables in noise.

4pSC9. Other-perception, self-perception and production of English phonemes by adult speakers of Spanish. Ellie Hanlon (Dept. of Educational Psych., CUNY, 365 Fifth Ave., New York, NY 10036, ehanlon1@gc.cuny.edu)

A causal model of second-language production was tested. It was hypothesized that correct production results from accurate perception of unfamiliar second language (L2) phonemes when spoken by others (other-perception) and the learner (self-perception). For the correct production to develop, the learner must form new phonetic categories for the novel L2 phonemes by listening to and observing native speakers. The categories are then used by the learner as an internal standard against which to judge production attempts. The participants' production was measured by recording a nonword syllable reading task. Self-perception was measured by asking participants to listen to their recordings and identify phonemes using a two-alternative forced choice format. Other-perception was measured the same way as self-perception except the stimuli were recorded by a native English speaker. The path analysis shows the squared multiple correlation coefficient for production when other-perception and self-perception are used as predictors is 0.83 (83% of the variance in production is accounted for by other- and self-perception). However, the relationship between other-perception and production is nonsignificant when controlling for the effects of self-perception. This suggests that learners are developing an internal standard through other-perception skill and using this standard to adjust their attempts at pronouncing novel L2 phonemes.

4pSC10. Effects of native language experience in Spanish–English bilinguals' ability to discriminate between two similar cross-language phonetic categories. Janielle Lugo-Marin and Cynthia M. Connine (Psych. Dept., SUNY-Binghamton, P.O. Box 6000, Binghamton, NY 13902-6000)

Two phoneme monitoring experiments examined bilingual listeners' (Spanish/English) ability to discriminate similar L2 (American English) phonemes (''sh'' and ''ch'') when L1 (Spanish) contains only one of the phonemes (''ch''). Experience with the ''ch'' phoneme in L1 was predicted to function as a perceptual magnet for the English phoneme ''sh,'' which is also a voiced palatal consonant. Both ''sh'' (Experiment 1) and ''ch'' (Experiment 2) targets were used. Target phonemes occurred in the initial or

final position of a carrier word. Response times for the ''sh'' targets in initial position were slower than for the ''ch'' targets. In word-final position, the ''ch'' targets were detected more slowly than ''sh'' targets. Accuracy rates were comparable for the ''sh'' and the ''ch'' targets in both positions. The results are discussed in terms of categorical generalization of L2 phonemes that are similar to L1 phonemes, knowledge of phonological rules about the likelihood of a voiced palatal consonant in word-final position in L1, and experiential factors with L2.

4pSC11. Does difficulty perceiving American English /r/ and /l/ affect Japanese listeners' lexical confusion of these phonemes? Ryo Komaki (ATR Human Information Sci. Labs., 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan, komaki@atr.co.jp)

Phonetic environment has proven to be one of the most important factors affecting non-native English speakers' perception of American English /r/ and /l/ [e.g., Lively *et al.*, J. Acoust. Soc. Am. **96**, 2076–2087 (1994)]. In the present study native Japanese speakers' lexical confusion of English /r/ and /l/ are examined through a word translation experiment which employed pairs of English words minimally contrasting /r/ and /l/ in various word positions. Results indicate that English words in which the /r/ or /l/ is located in a perceptually difficult to identify position (e.g., initial consonant cluster: ''branch'' versus ''blanch'') are more likely to be mistranslated than those in which the phoneme is located in an easy to identify position (e.g., final singleton: ''compare'' versus ''compel''). This finding suggests that errors in phoneme perception and lexical confusion might be related. Pedagogical implications for second language learning will be discussed with respect to results from both learning and retention experiments. [Work supported by TAO, Japan.]

4pSC12. Effect of age on perceptual learning of second-language phonetic contrasts. Reiko Akahane-Yamada, Tomoko Takada, and Rieko Kubo (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, yamada@atr.co.jp)

This paper is a mid-term report on an examination of the effect of age on learning new second language (L2) phonetic categories. Approximately 1000 Japanese speakers of various ages, from children (10 years old) to elderly (70 years old), were trained to identify English words produced by native speakers of American English. The words used minimally contrasted in /r/–/l/ for all age groups (10–70), and words contrasting in /b/–/v/ and /s/–/th/ were additionally used only for the younger age groups (10–22). A pretest–post-test design was administered in which the effect of the training was evaluated by the amount of improvement in accuracy from the pretest to the post-test. Results obtained at this point suggest that: (1) in general, the amount of improvement decreases with age; (2) during the early phases of training, the /r/–/l/ distinction is an extremely difficult one for Japanese children; and (3) elderly subjects, even in their 60s, can improve significantly. Implications for the theories of acquisition of speech perception will be discussed. [Work supported by TAO, Japan.]

4pSC13. Does semantic context help Japanese speakers identify English /r/ and /l/? Amanda Rothwell and Reiko Akahane-Yamada (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, rothwell@atr.co.jp)

Second language learners experience difficulty distinguishing phonemic contrasts that do not occur in their native language. In this study we investigated the effect of providing Japanese speakers with a semantic context when identifying /r/s and /l/s in English words. Questions addressed were: (1) whether semantic context improves performance; and (2) if context helps, whether the written or spoken context is most helpful. Japanese speakers participated in a two alternative forced-choice identification task in which participants identified words they heard by choosing one of two words contrasting in /r/ and /l/. Words were presented either: (a) in isolation; or (b) within contextual sentences in a variety of written and spoken forms. Preliminary results show that Japanese speakers' identification accuracy of the /r/–/l/ contrast was significantly better when words containing the contrast were presented in contextual sentences than when they were presented in isolation. The amount of improvement in performance was almost equal, whether the context was provided in written or spoken form, or a combination of thereof. These results suggest that a semantic context, both written and spoken, helps second language learners' identification of difficult phonemic contrasts. Implications of this work on second language learning will be discussed. [Work supported by TAO, Japan.]

4pSC14. Perceptual assimilation patterns and the perception of English /l/ and /ɹ/ by Japanese speakers. Katsura Aoyama (Div. of Speech & Hearing Sci., Univ. of Alabama at Birmingham, CH20, 1530 3rd Ave. S., Birmingham, AL 35294, aoyama@uab.edu) and James E. Flege (Univ. of Alabama at Birmingham, Birmingham, AL 35294)

This study examined the relationship between the perceptual assimilation patterns and the discrimination of English /l/ and /ɹ/ by Japanese speakers. The participants were 50 Japanese adults differing in their length of residence (LOR) in the U.S. In an identification task, they usually identified both English /l/ and /ɹ/ as Japanese /ɾ/. In a rating task, English /l/ was judged to be significantly more similar to Japanese /ɾ/ than English /ɹ/ was. The participants with a longer LOR gave lower similarity ratings for both English /l/ and /ɹ/ than those with a shorter LOR, suggesting that they became more aware of phonetic differences between Japanese /ɾ/ and English /l/ and /ɹ/. An exploratory analysis on similarity ratings was conducted for a subset of participants ($N=28$) based on their scores on /l/–/ɹ/ discrimination. Both relatively poor and relatively good discriminators judged /l/ to be more similar to Japanese /ɾ/ than /ɹ/, but the degree of difference between /l/ and /ɹ/ in terms of similarity ratings as Japanese /ɾ/ was smaller among poor discriminators than better discriminators. This suggests that the degree of ''goodness-of-fit'' as an L1 category may affect the ability to discriminate two L2 consonants. [Work supported by NIH.]

4pSC15. Neural activation during the discrimination of native and non-native speech contrasts by American and Japanese adults. Feng-Ming Tsao, Patricia K. Kuhl, and Huei-Mei Liu (CMBL 357988, Univ. of Washington, Seattle, WA 98195)

Language experience influences speech perception in adults. Adult speakers have been shown to experience difficulty when discriminating non-native consonants that are not phonemic in the native phonetic repertoire. The present study examined behavioral and neural differences in the processes underlying the perception of native and non-native speech sounds. American-English and Japanese native-speaking adults participated in a functional magnetic resonance imaging (fMRI) study. The study focused on three consonants: the American English /r–l/ consonants and Japanese /ɾ/. The investigation focused on neural correlates of both Kuhl's perceptual magnet effect (PME) and the boundary effects typical of categorical perception (CP). Subjects performed in a behavioral task in which they identified provided goodness ratings, and discriminated stimuli varying in formants 2 and 3. Results will be discussed in relation to cortical activation patterns related to the PME and CP boundary effect for native and foreign language contrasts, and the effects of language experience on brain organization. [This work was supported by NIH grant (HD37954) and Talaris Institute.]

4pSC16. Examining developmental patterns of neural activation during the discrimination of native and non-native speech contrasts by American–English adults and children. Jesica C. Pruitt, Tobey L. Nelson (Ctr. for Mind, Brain and Learning & Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98105), John S. Pruitt (Microsoft Corp., Redmond, WA), and Patricia K. Kuhl (Univ. of Washington, Seattle, WA)

To better understand the effects of language experience on speech perception, the present study examined developmental differences in the neural processes underlying the perception of native and non-native speech contrasts. Native American-English speaking adults and pre- and post-pubescent children participated in a functional magnetic resonance imaging (fMRI) study in which they discriminated non-native consonants from Hindi, Spanish, and Mandarin. The non-native contrasts differed in place of articulation, voicing, and manner. For comparison with the neural data, behavioral data were collected outside the scanner for the non-native and a native contrast (AE r/l) using a child-friendly computer-based testing method that presented the same contrasts to all subject groups. Results are discussed with regard to developmental differences in the location and extent of activation patterns for the native and non-native contrasts and the level of difficulty for the individual contrasts. [Work supported by NICHD grant HD37943 to Patricia K. Kuhl and Talaris Research Institute.]

4pSC17. Brain potentials in response to native and non-native speech contrasts in 11-month-old American infants. M. Rivera-Gaxiola, J. Silva-Pereyra, A. Garcia-Sierra, L. Klarman, and P. K. Kuhl (Speech and Hearing Sci. Dept. and Ctr. for Mind, Brain and Learning, M.S. 357920, Univ. of Washington, Seattle, WA 98195)

Behavioral studies show that 6- to 8-month-old infants discriminate both native and non-native phonetic differences, but that by 10–12 months of age, infants' perceptual sensitivities decline to many non-native contrasts. Kuhl *et al.* (1997) and Tsao *et al.* (2000) have also shown an enhancement in sensitivity to native contrasts, which was also shown by Cheour *et al.* (1998) using electrophysiological methods. Using both behavior and ERPs, Rivera-Gaxiola *et al.* (2000) reported MMNs in native English speaking adults for a non-native contrast that the same adults did not overtly discriminate. We here report auditory ERPs to native and non-native VOT contrasts in 11-month-old monolingual American infants ($N = 16$). Stimuli were presented as a double-oddball paradigm. Continuous EEG was recorded using electrocaps. ERPs were obtained off-line using standard methods. The results show that 11-month-olds display differential brain activity for native and non-native contrasts. A clear, widely distributed phonetic MMN can be observed for the native contrast, while the non-native acoustic differences pre-attentively detected are shown as differences in the first positive peak of the complex. A smaller centro-parietal negativity was also observed for the non-native contrast. [Work supported by NIH (HD 37954) and by the Talaris Research Institute.]

4pSC18. Native and non-native perception of phonemic length contrasts in Japanese: A categorization study. Hiroaki Kato and Keiichi Tajima (ATR Human Information Sci. Labs., Keihanna Science City, Kyoto 619-0288, Japan)

Segment length is distinctive in Japanese, for example, /kaze/ (''wind'') versus /kaze:/ (''taxation''). Such length contrasts are not necessarily categorical for non-native speakers. To study this property precisely, a series of perception experiments were conducted. Stimuli were based on a pair of nonsense Japanese words, /erete/ and /ere:te/. A continuum gradually varying between these words was synthesized using STRAIGHT, a high-fidelity speech analysis, synthesis, and manipulation system [Kawahara *et al.*, Speech Commun. **27**, 187–207 (1999)]. Three subject groups were tested: (1) English speakers with no Japanese experience; (2) English speakers who had spent 1–6 months in Japan; and (3) native Japanese speakers. All subjects participated in a single-stimulus two-alternative forced-choice identification task. Subjects in the second group were tested a second time after five days of perceptual training on phonemic length contrasts during which they were exposed to 1800 pairs of real Japanese words that minimally differed with respect to vowel length. Results suggest that identification boundaries of English speakers are less sharp than those of native speakers, and that identification boundaries sharpen with both exposure to Japanese and with perceptual training. Implications of these findings for second-language learning will be discussed. [Work supported by TAO, Japan.]

4pSC19. Native and non-native perception of phonemic length contrasts in Japanese: Effect of identification training and exposure. Keiichi Tajima, Amanda Rothwell (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, ktajima@atr.co.jp), and K. G. Munhall (Dept. of Psych., Dept. of Otolaryngol., Queen's Univ., Kingston, ON K7L 3N6, Canada)

Japanese distinguishes between words by the presence or absence of several types of mora phonemes, often realized as a contrast in segment duration, e.g., /hato/ (pigeon) versus /hat:o/ (hat). Several studies suggest that such contrasts are difficult for English listeners to perceive [e.g., R. Oguma, in *Japanese-Language Education Around the Globe*, 2000, Vol. 10, pp. 43–55]. In this study we investigated the effect on the perceptual ability of both Japanese exposure and perceptual identification training. Three groups of subjects were tested: (1) English speakers with no Japanese experience; (2) English speakers who had spent 1–6 months in Japan; and (3) native Japanese speakers. Subjects participated in a forced-choice identification task in which they heard words and nonwords produced by Japanese speakers and identified which word they heard by choosing among items that minimally differed with respect to these contrasts. Additionally, subjects in the second group underwent five days of perceptual training during which they received immediate feedback, repeating trials until they responded correctly. Results suggest that although the overall performance was relatively high, identification accuracy improved with exposure to Japanese and with perceptual identification training. Implications of this work on theories of second-language learning will be discussed. [Work supported by TAO, Japan.]

4pSC20. Learning stop-consonant voicing in English: The effect of perceived cross-language similarity. Wendy Baker (Dept. of Linguist., Brigham Young Univ., Provo, UT 84602) and Pavel Trofimovich (Dept. of Educational Psych., Univ. of Illinois at Urbana–Champaign, Urbana, IL 61801)

Previous research has demonstrated that perceived similarity between native- (L1) and second-language (L2) sounds overall influences how accurately L2 sounds are produced. No study, however, has examined whether certain phonetic cues to an accurate production of L2 sounds are susceptible to the influence of perceived similarity between L1 and L2 sounds more than others. In English, vowel duration is the primary cue to distinguishing word-final voiced and voiceless stops (e.g., bad versus bat). A secondary cue to stop-consonant voicing is the duration of the stop-gap closure of word-final voiced and voiceless stops. In this study we examined how perceived similarity between L1 and L2 sounds influences adult L2 learners' production of these two cues to stop-consonant voicing. Thus specifically examined were vowel and stop-gap closure duration in beginning, intermediate, and advanced native Korean adult learners' productions of six English vowels in 18 CVt and CVd words. The results of these acoustic analyses indicated that: (1) these two cues to stop-consonant voicing are differentially susceptible to the influence of perceived similarity between L1 and L2 sounds; and (2) perceived similarity between L1 and L2 sounds influences how accurately and quickly these cues are learned.

4pSC21. Short-term immersion can change the location of a phonetic category boundary. Anders Damgren-Hojen (Dept. of English, Univ. of Aarhus, Jens Chr. Skous Vej 7, DK-8000 Aarhus, Denmark)

In this study the effect of immersion in a second-language (L2) speaking environment on the perception and production of English sibilants by native (L1) speakers of Danish is examined. Three groups of subjects participated. Fourteen Danish au pairs or exchange students were tested before and after a 6–12 month immersion period. Eleven Danish control subjects were tested twice without immersion. Five native English baseline subjects were tested once. Danes tend to assimilate English /s/–/sh/ to two different Danish sound categories. However, the Danish /sh/ is palatalized and has more high frequency energy than English /sh/. If the immersed subjects changed their mental specifications for /sh/ after massive input of the more low frequency English /sh/ during immersion, one might expect a perceptual category boundary shift for /s/–sh/. This shift was indeed found, using an interpolated synthetic /s/–/sh/ continuum for the identification and discrimination in both an English and a Danish language set. Production was measured using spectral moments. However, there was no clear indication of a more authentic English pronunication of /sh/ after immersion than before. The results pertain to questions regarding the reorganization of phonetic categories after L2 experience, interactions between L1 and L2 phonetic systems, and the relation between perception and production.

4pSC22. Perception of American English glide consonants by Danish listeners. Catherine T. Best (Dept. of Psych., Wesleyan Univ., Middletown, CT 06459, cbest@wesleyan.edu) and Ocke-Schwen Bohn (Aarhus Univ., DK-8000 Aarhus C, Denmark)

Previous research has found both phonological and phonetic influences from the native language on perception of American English (AE) /r–l/, /w–r/, /w–j/ by Japanese and French speakers. This study examined native language effects on perception of the same stimuli by Danish listeners with minimal English experience. Danish contrasts /r–l/, but realizes them differently than AE, as pharyngeal versus nonvelarized lateral approximants. It has /j/ but lacks /w/; however, /v/ is realized as an unrounded labiodental approximant similar to /w/. Thus, on purely phonological grounds, Danes should have greater difficulty with AE /w–j/ and /w–r/ than /r–l/. On phonetic grounds, however, all three contrasts should be assimilated to Danish contrasts, and therefore categorized and discriminated quite well. Forced-choice identification was highly categorical for all three contrasts. However, A×B tests revealed a native-like discrimination peak only for /w–r/; /r–l/ showed a lower and less well-defined peak, whereas /w–j/ was discriminated at ceiling except for the /j/ end of the continuum. These results are consistent with previous evidence that both phonological and phonetic properties of the native language affect nonnative speech perception. The findings will be compared to prior cross-language research, and discussed in light of phonological differences among the listeners' languages.

4pSC23. Does one plus one equal three in early simultaneous bilingual speech perception? Megha Sundara (School of Commun. Sci. & Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, msunda@po-box.mcgill.ca)

Over half the world's population is bilingual, yet we know little about organization of phonetic abilities in early simultaneous bilinguals. In this study we tested bilingual French–English adults who had learned and used both languages simultaneously since birth. They were presented dental and alveolar stop consonants excised from real-word productions of French /t/ and English /d/, selected to insure that VOT values overlap. The alveolar-dental place distinction is not phonemic in either French or English. However, bilingual individuals are systematically exposed to this place distinction across their two native languages. Recent findings [Sundara and Polka, J. Acoust. Soc. Am. **110**, 2685 (2001)] indicate that simultaneous bilinguals clearly produce this place distinction. Perception was assessed using a categorical AXB task with tokens produced by multiple talkers in both French and English. Assimilation data were also obtained using a keyword identification task in both languages. Performance of bilinguals was compared and contrasted with those of respective monolingual groups and with native Malayalam listeners (dental-alveolar distinction is phonemic in Malayalam). The findings provide insights into perceptual organization in simultaneous bilingual adults by addressing whether an emergent contrast, not evident in either monolingual group is observed in both perception and production of bilinguals.

4pSC24. Age of immersion as a predictor of foreign accent. Miles Munro and Virginia Mann (3151 SSPA, Univ. of California, Irvine, CA 92697-5100, mmunro@uci.edu)

This study examined the relationship between Age of Immersion (AOI) and the Degree of Phonological Accent (DPA) that native-English-speaking judges perceived in the speech of Mandarin speakers who learned English as a second language. AOI and speech samples of variable length and linguistic context (single words, sentences, a short paragraph, and a self-generated picture narration) were collected from the target group (n=32, AOI=3–16) and from native speaker controls (n=4, AOI =0). These were judged by a native speaker panel (n=14) using a continuous scale, along with measurements of confidence. Judging was broken over 3 separate visits. A set of similar samples and speakers was repeated at the onset and termination of the main rating task of each visit to ensure reliability. No single AOI was found for a ''Critical Period'' (CP). Instead, DPA was found to increase in a highly linear manner with AOI. In general, ratings of DPA began to increase at about AOI=5, maxing out as early as AOI=7, although near-native ratings were given to some speakers whose AOI was greater than 5, especially in the case of female speakers. Thus DPA can be linearly predicted with AOI, though individual speakers are fairly variable after AOI=5.

4:40–5:10

Discussion
Winifred Strange, Discussion Leader

Session 4pSP

Signal Processing in Acoustics and Physical Acoustics: Seismic Signal Processing: Detection, Estimation and Inversion Methods

Alan W. Meyer, Cochair
Lawrence Livermore National Laboratory, L-154, 7000 East Avenue, Livermore, California 94550

Moyses Zindeluk, Cochair
Mechanical Engineering Department, COPPE/University of Rio de Janeiro, Caixa Postal 68 503, Rio de Janeiro 21945-000, Brazil

Invited Papers

1:00

4pSP1. Deconvolution and signal extraction in geophysics and acoustics. Leon H. Sibul, Michael J. Roan, and Josh Erling (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804-0030)

Deconvolution and signal extraction are fundamental signal processing techniques in geophysics and acoustics. An introductory overview of the standard second-order methods and minimum entropy deconvolution is presented. Limitations of the second-order methods are discussed and the need for more general methods is established. The minimum entropy deconvolution (MED), as proposed by Wiggins in 1977, is a technique for the deconvolution of seismic signals that overcomes limitations of the second-order method of deconvolution. The unifying conceptual framework MED, as presented in the Donoho's classical paper (1981) is discussed. The basic assumption of MED is that input signals to the forward filter are independent, identically distributed non-Gaussian random processes. A forward convolution filter ''makes'' the output of the forward filter more Gaussian which increases its entropy. The minimization of entropy restores the original non-Gaussian input. We also give an overview of recent developments in blind deconvolution (BDC), blind source separation (BSS), and blind signal extraction (BSE). The recent research in these areas uses information theoretic (IT) criteria (entropy, mutual information, K–L divergence, etc.) for optimization objective functions. Gradients of these objective functions are nonlinear functions, resulting in nonlinear algorithms. Some of the recursive algorithms for nonlinear optimization are reviewed.

1:30

4pSP2. Earth imaging using seismic migration. Michael Fehler, Lianjie Huang, Douglas Alde, and Hongchaun Sun (Los Alamos Natl. Lab., MS D443, Los Alamos, NM 87545)

Most subsurface imaging of the Earth is accomplished using seismic migration, which uses the entire wave forms recorded at the surface. The traditional method for performing migration was to use the Kirchhoff integral using a ray-based Green function to backproject the recorded data to positions of reflectors within the subsurface. Recent developments in Kirchhoff migration include the use of multivalued traveltime tables calculated with methods like wave-front construction ray tracing. Additional advances include the use of ray amplitudes calculated using dynamic ray tracing. Recently, more attention has been given to the development of wave-equation-based methods in which the recorded wave field is backprojected using some approximation to the wave equation. Wave-equation-based methods have grown in acceptance because better implementation of the backprojection operators have been developed and computing capability has allowed the generally slower methods to become more feasible.

1:50

4pSP3. 3-D tomographic imaging of the geologic structure in the Salton Sea Geothermal Field. Alan W. Meyer, Lawrence J. Hutchings, and Paul W. Kasameyer (Lawrence Livermore Natl. Lab., P.O. Box 808, Livermore, CA 94551-0808)

A three-dimensional tomographic reconstruction of the Salton Sea Geothermal Field is presented. This reconstruction is developed from data gathered in the course of one year between 15 September 1987 and 30 September 1988 using a microearthquake network. This geothermal field is important not only due to understanding potential energy sources but also because it is the result of a tectonic spreading zone bounded between two transverse fault systems: The San Andres system to the North and the Brawley fracture zone (BFZ) to the South. Here magma has penetrated into the crust to a depth of at least 8 km. This magmatic source is responsible for the microearthquakes generated along the BFZ as well as providing the thermal source for the geothermal activity. Using techniques for both blind source estimation as well as blind deconvolution, a travel time tomographic algorithm is applied to these data. The objective is to characterize the subterranean geological structure and estimate the fracturing that supports the geothermal field. These studies are a necessary foundation for future research into the energy capacity of this field. [Work performed under the auspices of the Department of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.]

2:10

4pSP4. Identification of a reversed-phase compressional in dipole borehole logging. Physical principle, theoretical modeling, and field examples. Lucio N. Tello (Computalog Research, Inc., 500 Winscott Rd., Fort Worth, TX 76126) and Marek Kozak (Magnetic Pulse, Inc., Fremont, CA 94538)

Dipole array sonic tools are particularly suited for estimating acoustic shear wave velocities in soft, unconsolidated, slow formations in boreholes. Dipole sources induce flexural waves, which are sustained in such formations. By contrast, monopole sources create pseudo-Rayleigh waves, related to shear wave velocity, which, in slow formations, refract away from the borehole and vanish before reaching the receivers. In addition, monopole sources create compressional and Stoneley waves. The typical order of these arrivals is compressional, shear (when present), and Stoneley. In an ideal situation dipole sources excite only borehole-flexural waves. The propagation velocity of the flexural wave is near that of the shear wave. Field experience backed by theoretical modeling indicates that in real situations dipole sources might also generate (reversed-phase) compressional and Stoneley waves. Theoretical simulations using well parameters taken from other logs render nearly identical wavetrains. Verification of the results with seismic measurements further validates the processing.

2:30

4pSP5. Transmission mode acoustic time-reversal imaging for nondestructive evaluation. Sean K. Lehman (L-154, Lawrence Livermore Natl. Lab., 7000 East Ave., Livermore, CA 94550) and Anthony J. Devaney (Northeastern Univ., Boston, MA 02115)

In previous ASA meetings and JASA papers, the extended and formalized theory of transmission mode time reversal in which the transceivers are noncoincident was presented. When combined with the subspace concepts of a generalized **MU**ltiple **SI**gnal **C**lassification (MUSIC) algorithm, this theory is used to form super-resolution images of scatterers buried in a medium. These techniques are now applied to ultrasonic nondestructive evaluation (NDE) of parts, and shallow subsurface seismic imaging. Results are presented of NDE experiments on metal and epoxy blocks using data collected from an adaptive ultrasonic array, that is, a ''time-reversal machine,'' at Lawrence Livermore National Laboratory. Also presented are the results of seismo-acoustic subsurface probing of buried hazardous waste pits at the Idaho National Engineering and Environmental Laboratory. [Work performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.] [Work supported in part by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the NSF (award number EEC-9986821) as well as from Air Force Contracts No. F41624-99-D6002 and No. F49620-99-C0013.]

2:50

4pSP6. Shallow subsurface applications of high-resolution seismic reflection. Don Steeples (Dept. of Geology, The Univ. of Kansas)

Shallow seismic reflection surveys have been applied to a wide variety of problems. For example, in many geologic settings, variations and discontinuities on the surface of bedrock can influence the transport and eventual fate of contaminants introduced at or near the ground surface. Using seismic methods to determine the nature and location of anomalous bedrock can be an essential component of hydrologic characterization. Shallow seismic surveys can also be used to detect earthquake faults and to image underground voids. During the early 1980s, the advent of digital engineering seismographs designed for shallow, high-resolution surveying spurred significant improvements in engineering and environmental reflection seismology. Commonly, shallow seismic reflection methods are used in conjunction with other geophysical and geological methods, supported by a well-planned drilling-verification effort. To the extent that seismic reflection, refraction, and surface-wave methods can constrain shallow stratigraphy, geologic structure, engineering properties, and relative permeability, these methods are useful in civil-engineering applications and in characterizing environmental sites. Case histories from Kansas, California, and Texas illustrate how seismic reflection can be used to map bedrock beneath alluvium at hazardous waste sites, detect abandoned coal mines, follow the top of the saturated zone during an alluvial aquifer pumping test, and map shallow faults that serve as contaminant flowpaths.

3:10–3:25 Break

3:25

4pSP7. Bayesian probability analysis for acoustic–seismic landmine detection. Ning Xiang (Natl. Ctr. for Phys. Acoust. and Dept. of Elec. Eng., Univ. of Mississippi, University, MS 38677, nxiang@olemiss.edu), James M. Sabatier, and Paul M. Goggans (Univ. of Mississippi)

Landmines buried in the subsurface induce distinct changes in the seismic vibration of the ground surface when an acoustic source insonifies the ground. A scanning laser Doppler vibrometer (SLDV) senses the acoustically-induced seismic vibration of the ground surface in a noncontact, remote manner. The SLDV-based acoustic-to-seismic coupling technology exhibits significant advantages over conventional sensors due to its capability for detecting both metal and nonmetal mines and its stand-off distance. The seismic vibration data scanned from the SLDV are preprocessed to form images. The detection of landmines relies primarily on an analysis

of the target amplitude, size, shape, and frequency range. A parametric model has been established [Xiang and Sabatier, J. Acoust. Soc. Am. **110**, 2740 (2001)] to describe the amplified surface vibration velocity induced by buried landmines within an appropriate frequency range. This model incorporates vibrational amplitude, size, position of landmines, and the background amplitude into a model-based analysis process in which Bayesian target detection and parameter estimation have been applied. Based on recent field measurement results, the landmine detection procedure within a Bayesian framework will be discussed. [Work supported by the United States Army Communications–Electronics Command, Night Vision and Electronic Sensors Directorate.]

3:45

4pSP8. Modeling heavy-tailed correlated clutter via wavelet packet bases with application to seismic event estimation. Paul J. Gendron (Naval Res. Lab., Acoust. Div., 4555 Overlook Ave. SW, Washington, DC 20375) and Balgobin Nandram (Worcester Polytechnic Inst., Worcester, MA 01609)

A model for correlated and heavy-tailed background noise in time series data is introduced. The approach is based on a multiresolution time adaptive wavelet packet expansion that assigns to each resolution and band a measure of spikiness associated with independent t distributed variates. The model is adaptive similarly in time, scale, and correlation structure within the dyadic structure of the wavelet packet framework. Parameter estimation is accomplished by recursions in time at each frequency band/scale. The model is applied to both ground motion data and shallow water ocean acoustic background noise. A comparison with the Gaussian model suggest its favorable advantage. Estimators for transient signals based on the new model are derived and show usefulness in extracting transients with unknown or weakly specified characteristics. The usefulness of the model and the estimator is demonstrated via a comparison with the fixed transform DWT on quarry blast seisms. The results show that the method outperforms DWT based estimation in MSE by a factor of 2 at a moderate SNR for these classes of transients. [Work supported by ONR.]

Contributed Papers

4:05

4pSP9. Use of the Bayesian Cramér–Rao bound in seismic exploration and research planning. Max Deffenbaugh (ExxonMobil Upstream Res. Co., P.O. Box 2189, Houston, TX 77252, max.deffenbaugh@exxonmobil.com)

Modern exploration for oil is heavily dependent on seismic data, which is used to detect and estimate properties of subsurface reservoirs. Seismic data is processed by applying a sequence of algorithms to filter, estimate velocity fields, beamform, and invert for reservoir properties. There are multiple algorithms to choose from at each processing step—ranging from simple and inexpensive to complex and costly. The Bayesian Cramér–Rao bound, combined with a statistical model of the signal and noise in seismic data, can be used to predict the information content of a specific data set at each stage in a proposed processing flow. This allows the exploration manager to identify the least-cost processing flow to achieve a certain information goal and to identify when the goal is unachievable at any cost. It can also guide research efforts to those processing steps where the greatest amount of information is being lost.

4:20

4pSP10. Correlation vibroacoustic imaging of subsurface objects. Alexander Ekimov, Nikolay Sedunov, Michael Tsionskiy, and Dimitri Donskoy (Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030)

Any subsurface object whose mechanical impedance is different from the surrounding media manifests itself on the surface by contributing to the surface mechanical impedance. Since most of the surfaces (except plates and membranes) can be characterized by the local impedance, such a manifestation can be interpreted as an object image projected on the surface. This image projection can be captured by measuring surfaces local impedance at a grid of points or pixels. One of the problems in implementing this imaging approach is that object impedance contribution can vary substantially depending on the frequency, depth, and medium properties. Therefore, multiple impedance images must be obtained at the different frequencies and, consequently, an image selection or more sophisticated image processing should be implemented at the expense of additional time and processing resources. The developed correlation technique utilizes broadband signals and performs convolution of the broadcasted signal with a measured vibration signal at each pixel. The resulted image has a high contrast and takes a fraction of time and processing power to create as compared with the image/frequency selection approach. Examples of practical implementation of the developed method are presented and discussed.

4:35

4pSP11. Field site evaluation for seismic mine detection. James S. Martin, Gregg D. Larson, Peter H. Rogers (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405), Waymond R. Scott, Jr. (School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332), and George S. McCallII (Georgia Inst. of Technol., Atlanta, GA 30332)

A system has been developed that uses audio-frequency surface seismic waves for the detection and imaging of buried landmines. The system is based on the measurement of seismic displacements immediately above buried mines using noncontacting vibrometers that interrogate the surface motion with either radar or ultrasonic signals. In laboratory tests and limited field tests the system has demonstrated the ability to detect a variety of inert antipersonnel and antitank mines with background contrast in excess of 20 dB. Current work on the system is focused on the transition from the laboratory into the field. To facilitate this, a series of experiments has been undertaken to measure the characteristics of several field test sites. The tradeoff between image contrast and scanning speed is of primary concern in evaluating the features of these sites. The field experiments have investigated the nature of ambient seismic noise, input impedance at the seismic source (a ground contacting shaker), modal content of the seismic interrogation signal, and the nature of the nonlinearities in the soil. Observed nonlinear phenomena have included harmonic generation, phase speed slowing, dispersion and spall. Although interesting, the differences between the field sites and the laboratory model do not appear to pose problems for seismic mine detection.

Session 4pUW

Underwater Acoustics and Acoustical Oceanography: David Weston Memorial Session II

Michael J. Buckingham, Cochair
Scripps Institution of Oceanography, La Jolla, California 92093-0213

Chester M. McKinney, Cochair
4305 Farhills Drive, Austin, Texas 78731

Contributed Papers

1:00

4pUW1. The feedback phenomenon applied to underwater acoustics. Philippe Roux, Jason E. Jordan, and W. A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238)

People are familiar with the feedback phenomenon that results in the loud sound heard when a musician plays an electric instrument directly into a speaker. Feedback occurs when a source and a receiver are connected both acoustically through the propagation medium and electrically through an amplifier in such way that the received signal is simultaneously and continuously added to the emitted signal. A resonance is then obtained when the emitter and the receiver are in phase. The resonant frequency appears to be highly sensitive to fluctuations of the propagation medium. The feedback phenomenon has been experimentally demonstrated as a means to monitor the temperature fluctuation of a shallow water environment [''Acoustic monitoring of the sea medium variability: experimental testing of new methods,'' by A. V. Furduev, Acoust. Phys. **47**, No. 3, 361–368 (2001)]. The goal of our work is to reproduce the feedback experiment using an alternative method that decomposes the feedback phenomenon into an iterative process. Successful reproduction of the feedback is accomplished using a step-by-step algorithm which details the evolution of the system from the initial signal to its steady-state form. These experimental and numerical results illustrate the potential of the feedback process for use in narrow-band acoustical tomography.

1:15

4pUW2. System-orthogonal functions for sound velocity profile perturbation. Wen Xu and Henrik Schmidt (MIT, 77 Massachusetts Ave., Rm. 5-204, Cambridge, MA 02139, wenxu@mit.edu.)

Empirical orthogonal functions (EOFs) are derived from direct measurements of the sound velocity profile (SVP) and they are orthogonal in regard to the statistics of the SVP uncertainty. Viewed from the sonar output end, however, the effect of an error in one EOF is usually coupled with the effect of the error in another due to the nonlinear physical constraints. Thus the traditional EOF is not an optimum basis set to characterize the SVP uncertainty in regard to sonar performance measure. To obtain a system-orthogonal SVP representation, a new set of orthonormal functions is derived for SVP perturbations in a water column. The performance measure used is the Cramer–Rao bound (CRB) for SVP expansion coefficients, and a full-field random Gaussian signal model is assumed. The derived functions make the CRB matrix diagonal, decoupling the errors in the estimation of the expansion coefficients. Compared to the traditional EOFs, the new set of SVP basis functions depends on both the statistics of the sound speed uncertainty and the sound waveguide propagation property. It also includes the noise effect in measurements. Using this system-orthogonal function set makes possible the investigation of the relative significance of the individual basis functions. [Work supported by ONR.]

1:30

4pUW3. The partial wave expansion for scattering in oceanic wave guides. Dalcio K. Dacol (Acoust. Div., Naval Res. Lab., Washington, DC 20375-5350)

The partial wave expansion for target scattering in oceanic wave guides is obtained via the pseudo-potential method and discussed in detail. The case of sound scattering by multiple objects in an oceanic waveguide will also be discussed. The Fermi pseudo-potential was introduced in quantum mechanics as a means of simplifying problems involving scattering by a multicentered potential including applications to many-body problems. As such, it has direct applications in acoustics. It can also be used to discuss scattering off targets in confining geometries such as the interior of cavities or in wave guides. [Work supported by ONR.]

1:45

4pUW4. A study of the connection between tidal velocities, soliton packets and acoustic signal losses. Stanley A. Chin-Bing, Alex C. Warn-Varnas, David B. King (Naval Res. Lab., Stennis Space Center, MS 39529-5004, chinbing@nrlssc.navy.mil), Kevin G. Lamb (Univ. of Waterloo, Waterloo, ON N2L 3G1, Canada), James A. Hawkins (Planning Systems, Inc., Slidell, LA 70458), and Marvi Teixeira (Polytechnic Univ. of Puerto Rico, Hato Rey, PR 00988)

Coupled ocean model and acoustic model simulations of soliton packets in the Yellow Sea have indicated that the environmental conditions necessary for anomalous signal losses can occur several times in a 24 h period. These conditions and the subsequent signal losses were observed in simulations made over an 80 h space–time evolution of soliton packets that were generated by a 0.7 m/s tidal velocity [Chin-Bing *et al.*, J. Acoust. Soc. Am. **111**, 2459 (2002)]. This particular tidal velocity was used to initiate the Lamb soliton model because the soliton packets that were generated compared favorably with SAR measurements of soliton packets in the Yellow Sea. The tidal velocities in this region can range from 0.3 m/s to 1.2 m/s. In this work we extend our simulations and analyses to include soliton packets generated by other tidal velocities in the 0.3–1.2 m/s band. Anomalous signal losses are again observed. Examples will be shown that illustrate the connections between the tidal velocities, the soliton packets that are generated by these tidal velocities, and the signal losses that can occur when acoustic signals are propagated through these soliton packets. [Work supported by ONR/NRL and by a High Performance Computing DoD grant.]

2:00

4pUW5. Seismo-acoustic propagation in environments that depend strongly on both range and depth. Donald A. Outing, William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180), LeRoy M. Dorman (Scripps Inst. of Oceanogr., La Jolla, CA 92093), and Michael D. Collins (Naval Res. Lab., Washington, DC 20375)

The parabolic equation method provides an excellent combination of accuracy and efficiency for range-dependent ocean acoustics and seismology problems. This approach is highly developed for problems in which the ocean bottom can be modeled as a fluid. For the elastic case, there

remain some accuracy limitations for problems involving sloping interfaces. Progress on this problem has been made by combining a new formulation of the elastic parabolic equation that handles layering more effectively [W. Jerzak, "Parabolic Equations for Layered Elastic Media," doctoral dissertation, Rensselaer Polytechnic Institute, Troy, NY (2001)] and a mapping approach that handles sloping interfaces accurately [J. Acoust. Soc. Am. **107**, 1937–1942 (2000)]. This approach makes it possible to handle problems involving complex layering and steep slopes, but the rate of change of the slope must be small. The method and its application to data will be described. Our immediate goal is to model propagation of seismic surface waves propagating across a transition between dry and marshy terrain. We have suitable data applicable to vehicle-tracking problems from Marine Corps Base Camp, Pendleton, CA. [Work supported by ONR.]

2:15

4pUW6. The split-step Padé–Fourier solution. Elizabeth Küsel, Flor A. Espinoza, William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180), and Michael D. Collins (Naval Res. Lab., Washington, DC 20375)

At lower frequencies, the split-step Padé solution is at least comparable in efficiency to the split-step Fourier solution [J. Acoust. Soc. Am. **100**, 178–182 (1996)]. For deep-water problems in the kilohertz regime, the split-step Fourier solution provides greater efficiency but is not as accurate as the split-step Padé solution. The split-step Padé–Fourier solution combines the attributes of the two split-step solutions to achieve greater efficiency for some problems without sacrificing accuracy. This approach is based on factoring the exact exponential operator by multiplying and dividing by the approximate exponential that is used in the split-step Fourier solution. The multiplicative factor is implemented using the classical algorithm based on the fast Fourier transform. The other two factors are handled together with a rational approximation. Since the oscillation rate of these two factors is much lower than for the exact exponential, an efficient solution can be obtained using a very large range step and a relatively small number of terms in the rational approximation. [Work supported by ONR.]

2:30

4pUW7. Internal solitary waves and coupled perturbed modes. Chris J. Higham and Chris T. Tindle (Phys. Dept., Univ. of Auckland, Private Bag 92019, Auckland, New Zealand)

Coupled perturbed mode theory combines conventional coupled modes and perturbation theory for fast range-dependent normal mode calculations. The method is applied to sound propagation through shallow water internal solitary waves, or solitons. The solitons considered are depressions of the thermocline, separating well-mixed upper and lower layers. Shallow water solitons lead to strong mode coupling. The theory gives physical insight into soliton coupling physics through explicit mode coupling formulas.

2:45

4pUW8. Coupled mode propagation in an elastic oceanic wedge. Ahmad T. Abawi (SPAWAR Systems Ctr., 53560 Hull St., San Diego, CA 92152-5001)

The elastic version of the one-way coupled mode propagation model [Abawi, J. Acoust. Soc. Am. **111**, 160–167 (2002)] is used to compute the propagation of waves in an ocean overlaying a shear-supporting wedge-shaped bottom. The range-dependent ocean is approximated by a set of stair-step elastic waveguides. The elastic modes are obtained from the solution of the equations of motion in each stair-step and the solution of the range-dependent problem is obtained by solving a set of coupled differential equations for the mode amplitudes as a function of range. Various field quantities such as the scalar and shear potentials, the compressional and shear pressures, and the displacements are computed and the results are compared with those obtained from the fast field propagation model, OASES.

3:00–3:15 Break

3:15

4pUW9. Influence of shallow water currents on the performance of a broadband time-reversing array. Karim G. Sabra and David R. Dowling (Dept. of Mech. Eng., Univ. of Michigan, 1231 Beal Ave., 2019 Lay Autolab., Ann Arbor, MI 48109-2121)

In recent shallow water experiments, the temporal and spatial focusing properties of time-reversing arrays (TRAs) were shown to be robust in a reciprocal medium and useful for underwater applications. The presence of oceanic currents in coastal environments leads to nonreciprocal acoustic propagation. In this case, time-reversal invariance is modified because the propagation speed inhomogeneity depends on the direction of acoustic propagation. Therefore, similarly to phase coherent reciprocal transmissions, a TRA will be influenced by the current-induced effects but not by the scalar contributions due to temperature or salinity. TRA performance, in the presence of steady currents, is investigated both theoretically using a simple first order normal mode formulation and numerically using a parabolic equation code for moving media, GCPEM [D. Mikhin, J. Acoust. Soc. Am. **105**, 1362 (1999)]. In a multipath shallow ocean environment, the retrofocus field is shifted relative to its location in a nonmoving medium. This shift depends on the current speed and the range-depth dependency of the ocean current profile because each acoustic mode is influenced differently. The possibility of using TRAs for monitoring coastal currents will be discussed.

3:30

4pUW10. A range-dependent propagation model based on a combination of ray theory and plane-wave reflection coefficients. Jens M. Hovem (Norwegian Univ. of Science and Technol., Trondheim, Norway) and D. P. Knobles (Univ. of Texas, Austin, TX 78713)

The paper describes a range-dependent propagation model based on a combination of range-dependent ray tracing and plane-wave bottom responses. The ray-tracing module of the model determines all the eigenrays between any source/receiver pairs and stores the ray histories. The received wave field is then synthesized by adding the contributions of all the eigenrays, taking into account the reflections from the bottom and the surface. The model can treat arbitrarily varying bottom topography and a layered elastic bottom as long as the layers are parallel. In the current version, the bottom is modeled with a sedimentary layer over an elastic half space, but more complicated structures are easily implemented. The new model has been tested against other models on several benchmark problems and also applied in the analysis and modeling of up-slope and down-slope propagation data recorded on a 52-element center-tapered array that was deployed at two locations about 70 miles east of Jacksonville, FL. The paper presents the results of these tests with an assessment of the potential use in connection with geo-acoustic inversion of range-dependent and elastic scenarios. [Work supported by Applied Research Laboratories, The University of Texas.]

3:45

4pUW11. Interference patterns in range dependent environments. Henry Cox and Kevin D. Heaney (ORINCON Industries, 4350 N. Fairfax Dr., Ste. 470, Arlington, VA 22203)

Interference patterns in space and frequency occur in ocean acoustics due to coherent multipath propagation. The wave guide invariant beta is commonly used to describe the behavior of these patterns [Brikhosvsky and Lysanov, *Fundamentals of Ocean Acoustics*, 2nd ed. (Springer-Verlag, New York, 1991)]. The standard derivation of this invariant uses a normal

mode expansion and is limited to range independent and weakly range dependent environments. A different approach is introduced that describes the interfering components in terms of their travel times. This leads to a very simple derivation of basic results that apply equally well to range dependent environments and to situations, such as Lloyd's Mirror, in which the normal mode representation is inconvenient. Beta can be expressed in terms of the local phase velocities and the difference in travel times for the interfering components. The travel time differences involve integration over the entire path trajectories from the source to the point where the interference pattern is observed. Numerical examples using ray theory will be presented.

4:00

4pUW12. Evidence of three-dimensional waveguide propagation in SWARM 95 data. Scott Frank, William L. Siegmann (Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180), and Mohsen Badiey (Univ. of Delaware, Newark, DE 19716)

During a period of the SWARM95 experiment, strong nonlinear internal waves passed across two tracks that had airgun pulses propagating along them. Environmental data for this period indicate that the angles between the tracks and the internal wave-fronts, which were roughly planar, were very different—one angle being close to zero, and the other approximately 42 deg. Two-dimensional PE simulations for these waveguides show dramatically different results for depth-averaged, pulse-integrated energy variations. Specifically, the observed levels can be reproduced for the waveguide with the large incidence angle, but not for the one with the small incidence angle. For the latter case, data show significant variations in pulse shapes and in the integrated energy ($\approx$5 dB), while simulations show very small changes in both of these characteristics. Results from several recent computational and theoretical studies suggest that the cause may be three-dimensional effects from horizontal refraction and modal interference due to the nonlinear internal waves. The adiabatic mode parabolic equation [Collins, J. Acoust. Soc. Am. **94** (1993)] is used to quantify the three-dimensional influence of the internal waves on the integrated energy variations. The results demonstrate experimental evidence of three-dimensional effects from strong nonlinear internal waves. [Work supported by ONR.]

4:15

4pUW13. Computing the two-point correlation function directly from the transport equation using split-step Padé solutions. Chen-Fen Huang, Philippe Roux, and W. A. Kuperman (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238, chenfen@mpl.ucsd.edu)

The two-point correlation function that describes the correlation between an acoustic field at two depths as a function of range obeys a transport equation. Without making a frequency or narrow angle approximation, this equation can be solved by parabolic equation methods. The solution algorithm is split into two depth-steps using the alternating direction implicit (ADI) method to advance one range-step. Then, at each depth substep, the split-step Padé solution [M. D. Collins, J. Acoust. Soc. Am. **96**, 382–385 (1994)] is used. Computations are confirmed against intensities computed from the indirect method that constructs the two-point function by an outer product of pressure vectors. Examples of the propagation of the vertical correlation for deep, shallow, and stochastic environments are presented.

4:30

4pUW14. A study on simplification of the Biot–Stoll model in water-saturated porous medium. Keunhwa Lee and Woojae Seong (Dept. of Ocean Eng., Seoul Natl. Univ., Seoul 151-742, Korea)

A practical approximate method based on the Biot–Stoll model to represent acoustic waves in the ocean in the presence of porous sediment is presented. The ocean sediment can be described individually by four models; visco-fluid, visco-elastic, equivalent fluid, and water-saturated porous medium. The first two models are approximate forms of the Biot–Stoll model. The equivalent fluid model of Zhang and Tindle [J. Acoust. Soc. Am. **98**, 3391–3396 (1995)] derived for the fluid/solid interface is extended to the fluid/porous-medium interface for analytical derivation of the reflection coefficient. Then, the reflection characteristics as functions of frame stiffness (Kb/Kr) and frequency region (f/fc) are discussed for all four models. In weak stiffness and high-frequency region, the equivalent fluid and visco-fluid bottom models describe ocean bottom better than the visco-elastic bottom model. Two-phase property is dominant in strong (high) stiffness or high-frequency region. This demonstrates that the visco-fluid and equivalent fluid models are useful approximate models in weak stiffness porous material such as the marine sediment.

4:45

4pUW15. Asymptotic methods of evaluation of acoustic fields generated by underwater moving sources. Vladimir S. Rabinovich (Instituto Politecnico Nacional de Mexico, ESIME-Zacatenco, edif. 1, Av. IPN, Mexico, D.F. 07738)

The problem of acoustic waves propagation from sources moving in stratified and almost-stratified waveguides simulating real oceanic waveguides is considered. It is assumed that acoustic sources move under water in the ocean with a subsonic variable velocity. Our approach for the evaluation of the acoustic field is based on the combination of two asymptotic methods: (1) The asymptotic analysis of the Green's function for almost-stratified waveguides by the operator ray method with respect to a small parameter which characterizes the variation of waveguides in the horizontal direction; (2) a subsequent asymptotic analysis of the integral representation of the field generated by a moving source by the two-dimensional stationary phase method with respect to a large parameter which characterizes the smallness of variations of the amplitude of source and the smallness of the vertical component of the source speed. Asymptotic formulas for the acoustic pressure, for the mode and time Doppler effects, are obtained. These formulas have a clear physical meaning and are convenient for the numerical simulation of the problem.

5:00

4pUW16. Analysis of very high frequency propagation in sediments: Experimental results and modeling. Richard Brothers, Sarah Page, Gary Heald (QinetiQ, Newton's Rd., Weymouth DT4 8UR, UK, rbrothers@qinetiq.com), Tim Leighton, Matt Simpson (ISVR, Southampton Univ., UK), and Justin Dix (SOC, Southampton Univ., UK)

A current QinetiQ study is investigating the propagation of sound waves into sediment at frequencies higher than 300 kHz. Previous work has found notable discrepancies between model predictions and experimental results and comparisons are inconsistent and unreliable. This new work package investigates the development of new scattering theories for frequencies ranging from 300 kHz to 1 MHz. In particular, the application of a pseudospectral time difference approached is analyzed. The model originally developed for lower frequency applications, is set up in various geometrical scenarios and for varying very high frequencies. Results show the received simulated pulses obtained for hydrophones placed within the sediment and source colocated. Furthermore, simulations are compared in tank experimental data. The controlled tank experiments were conducted by Southampton University and data are analyzed and discussed for various conditions and frequencies.

Meeting of Accredited Standards Committee (ASC) S3 Bioacoustics

to be held jointly with the

ANSI-Acredited U.S. Technical Advisory Group (TAG) Meeting for:
ISO/TC 108/SC 4 Human Exposure to Mechanical Vibration and Shock

R. F. Burkard, Chair S3
Hearing Research Laboratory, State University of New York at Buffalo, 215 Parker Hall, Buffalo, New York 14214

J. Franks, Vice Chair S3
Robert A. Taft Laboratories, 4676 Columbia Parkway, Mail Stop C27, Cincinnati, Ohio 45226

D. D. Reynolds, Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 4, Human Exposure to Mechanical Vibration and Shock
3939 Briar Crest Court, Las Vegas, Nevada 89120

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 4 Human Exposure to Mechanical Vibration and Shock
1325 Meadow Lane, Yellow Springs, Ohio 45387

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/TC108/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 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 Thursday, 5 December 2002.*

Scope of S3: Standards, specifications, methods of measurement and test, and terminology in the fields of psychological and physiological acoustics, including aspects of general acoustics, shock, and vibration which pertain to biological safety, tolerance, and comfort.

Meeting of Accredited Standards Committee (ASC) S1 Acoustics

G. S. K. Wong, Chair S1
Institute for National Measurement Standards, National Research Council, Montreal Road, Bldg. M36, Ottawa, Ontario K1A OR6, Canada

J. Seiler, Vice Chair S1
U.S. Department of Labor, Mine Safety and Health Admin., P.O. Box 18233, Bldg. 38, Cochrans Mill Road, Pittsburgh, Pennsylvania 15236

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 Thursday, 5 December 2002.*

Scope of S1: Standards, specifications, methods of measurement and test, and terminology in the field of physical acoustics, including architectural acoustics, electroacoustics, sonics and ultrasonics, and underwater sound, but excluding those aspects which pertain to biological safety, tolerance, and comfort.

Session 5aAA

Architectural Acoustics: Room Reverberation Modeling, Measurements and Materials

Angelo J. Campanella, Cochair
Campanella Associates, 3201 Ridgewood Drive, Columbus (Hilliard), Ohio 43026-2453

Fausto Rodriguez, Cochair
Condor 260, Casa 8, Col. Las Aguilas 01710, D.F. Mexico

Contributed Papers

8:30

5aAA1. Solving the diffusion equation with a finite element solver: Calculation of diffuse sound field in room acoustics. Vincent Valeau, Anas Sakout, Feng Li (LEPTAB, Univ. La Rochelle, Av. M. Crepeau, 17042 La Rochelle, Cedex 01, France), and Judicael Picaut (LCPC, 44341 Bouguenais Cedex, France)

Over the last years, some publications [e.g., Picaut, Simon, and Hardy, J. Acoust. Soc. Am. **106**, 2638–2645 (1999)] showed that the acoustic energy density in closed or semiclosed spaces is the solution of a diffusion equation. This approach allows the nonuniform repartition of energy, and is especially relevant in room acoustics for complex spaces or long rooms. In this work, the 3-D diffusion equation is solved directly by using a finite element solver, for a set of long rooms and absorbing rooms. The stationary equation is first solved. A constant-power acoustic source is modelized by setting appropriate boundary conditions. The time-dependent problem is also solved to simulate the sound decay, with an impulse source defined in a subregion with relevant initial conditions. Results concerning sound attenuation and reverberation times match satisfactorily with other theoretical and numerical models. An application is also given for two coupled rooms.

8:45

5aAA2. A multipole method for computation of acoustics of a convex polyhedral room. Nail A. Gumerov, Ramani Duraiswami, Dmitry N. Zotkin, and Larry S. Davis (Perceptual Interfaces and Reality Lab., UMIACS, Univ. of Maryland, 115 A.V. Williams Bldg., College Park, MD 20742)

The classical method for constructing room impulse responses (Allen and Berkeley, 1979) works for the case of rectangular rooms, and for general polyhedral rooms heuristic techniques are used. A new method for computation of the room impulse response of a general polyhedral room is presented. The method is based on the decomposition of the acoustic field into a free field and a sum of fields caused by interaction with the walls. These are modeled using a few multipole sources located outside the room. First a solution is developed for the case when the room is convex, and the walls are sound hard. Here the solution can be obtained using symmetry conditions for the multipoles re-expanded at proper locations. Using multipole translation theory, and truncation, a closed system of linear equations for the expansion coefficients is obtained, which can be solved to obtain the room impulse response. The method is then generalized to include wall impedances, by using linear differential transforms of multipole solutions. The method is validated by comparison with known solutions based on the image method for particular cases, and listening tests. [Work supported by NSF.]

9:00

5aAA3. New software for room acoustics simulation: A study of its performance and validation by an international comparison. Thiago S. Camilo, Ludimila O. Medrado, and Roberto A. Tenenbaum (Lab. de Acústica e Vibrações, Dept. de Eng. Mecânica, CT/Univ. Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil CEP 21945-970)

Nowadays, with the increasing capacity of personal computers, software in room acoustics simulation is being developed all over the world by strong architectural acoustic groups, as a practical and reliable way to evaluate the sound quality of a room. However, software must be exaustively tested in order to be recognized as a trustworthy tool for modern acoustical design. In this paper, a new software tool in room acoustics simulation, RAIOS, is presented. A hybrid technique, which uses the ray-tracing method to simulate the first part of the impulse response and an energy transition algorithm to compute its reverberant tail, for a room with arbitrary geometry is described. All the parameters for room acoustics quality, following ISO 3382 standard, are also computed. The effects of filtering, reflection, and absorption on the walls and in the air are considered in the model. The software has participated in Round Robin 3, promoted by PTB, Germany. The results of phase three of RR3 (the most complex one) are presented in this paper. Comparison with other software is given. If the experimental data of the PTB studio are available before the end of this year, then comparisons will be made.

9:15

5aAA4. Acoustic simulations with a ray-tracing program: Reporting an experience. Milton V. Granado, Jr. (School of Architecture and Urban Planning, Mackenzie Presbiterian Univ., Sao Paulo, SP, Brazil)

The possibility of exporting to ray-tracing type programs the geometry generated by Auto-Cad, is very appealing for the architect who wishes to simulate the acoustic behavior of a room. It was found that a good knowledge of room acoustic is necessary for the selection of the program parameters and the objective acoustical measures outputted by the program for the assessment of the required acoustical quality. There is scarce information on the absorption coefficients of lining materials, but it is on the choice of the diffusion coefficients where lie the greatest uncertainties. Because of the considerable impact that these data have on the computational results, the user may feel insecure about the results of his efforts. Therefore it is advisable that trained personnel participate in the efficient use of this computational tool. It was found however that one form of the output of the program, that in which the acoustical parameters are color mapped, have qualitatively corroborated speech intelligibility results objectively measured in the field. (To be presented in Portuguese.)

9:30

5aAA5. Aperture size, materiality of the secondary room, and listener location: Impact on the simulated impulse response of a coupled-volume concert hall. Michael Ermann, Marty E. Johnson (Virginia Tech, 201 Cowgill Hall 0205, Blacksburg, VA 24061, mermann@vt.edu), and Byron W. Harrison (The Talaske Group, Oak Park, IL 60301)

By adding a second room to a concert hall, and designing doors to control the sonic transparency between the two rooms, designers can create a new, coupled acoustic. Concert halls use coupling to achieve a variable, longer, and distinct reverberant quality for their musicians and listeners. For this study, a coupled-volume concert hall based on an existing performing arts center is conceived and computer modeled. It has a fixed geometric volume, form, and primary-room sound absorption. Ray-tracing software simulates impulse responses, varying both aperture size and secondary-room sound-absorption level, across a grid of receiver (listener) locations. The results are compared with statistical analysis that suggests a highly sensitive relationship between the double-sloped condition and the architecture of the space. This line of study aims to quantitatively and spatially correlate the double-sloped condition with (1) aperture size exposing the chamber, (2) sound absorptance in the coupled volume, and (3) listener location.

9:45

5aAA6. Reverberation time measurement using integrated impulse response and sweep sine excitation. Marco Nabuco and Paulo Brando (Acoust. Testing Lab., Inmetro, Av. N. Sra. Gracas, 50-Xerem-D, Caxias-25 250-020-RJ, Brazil)

As the capacity and speed of digital processing systems becomes much higher, the integrated impulsive response for reverberation time measurements by the indirect method also becomes more feasible and faster. The MLS technique to obtain the impulse response for LTI has been developed during the last several years and it is very well reported by the bibliography. Some frequency analyzers available in the market are capable to generate and process MLS to get the impulse responses very easily. Sometimes, when the room to be tested is very reverberant, sequences of higher order and a certain number of average are necessary to assure acceptable signal-to-noise ratio. The sweep sine technique or the deconvolution method to obtain impulsive responses presents many new advantages, most of them still reported in various technical documents. This paper presents the results of application of this technique to measure the reverberation time in two different reverberation rooms. Comparisons with MLS, ensemble, and reverberation time averages are presented. The sweep sine technique repeatability was verified in a reverberation chamber for a polyurethane foam sample and showed smaller standard deviations when compared with other techniques. (To be presented in Portuguese.)

10:00–10:15 Break

10:15

5aAA7. Reverberation time measurement of a structure or a room using modified wavelet transform. Sang-Kwon Lee (Dept. of Mech. Eng., Inha Univ., Inchon, Korea) and Jin-Ho Park (ARD Div., KAERI, Taejon, Korea)

In structure damping or acoustic decay measurement using the third-octave band pass filter, there is inevitable experiment error produced by ringing at the tail part of the impulse response of the filter. This ringing gives rise to the distortion of the decay curve. To reduce this error, it has been recommended that the product of the 3 dB bandwidth B of the filter and the reverberation time T of the system is at least 16. In case of a system having a short reverberation time the decay curve cannot be measured reliably using the filter. In this paper, the modified wavelet transform (MWT) has been proposed to determine accurately the decay curve with a low value of BT. The MWT decomposes the decay signal of a system into time–frequency domain using the third-octave band wavelet filter bank. When the MWT is applied to the measurement of the decay curve, it is found that the requirement $BT<16$ can be replaced by the replacement $BT<4$. It was applied to the reverberation time measurement of a small room and the identification of impact location occurred by a loose part in the nuclear power plant. [Work supported by Nuclear Research Project in Korea.]

10:30

5aAA8. Calculating room acoustic parameters from pseudo-impulsive acoustic sources. Maria L. San Martin, Antonio Vela, Ricardo San Martin, and Miguel A. Arana (Dpto. de Fisica, Universidad Publica de Navarra, 31006 Pamplona, Spain)

The impulse response function provides complete information to predict the acoustic response of a room to an acoustic input of arbitrary characteristics. At this job study, small explosions of firecrackers are proposed to be used as pseudo-impulsive acoustics sources to determine some acoustic parameters of a room such as reverberation time, definition, and clarity, comparing these results to those obtained with other techniques. A previous characterization of these sources allows us to state that they can be used for this purpose because they are, in practice, omnidirectional, their temporary pattern is highly repetitive and their spectral power is, as well, repetitive and with enough power in octave bands from 125 Hz to 8 kHz. If the linear time-invariant system impulse response $h(t)$ is known, output signal $s(t)$ regarding any arbitrary signal $s(t)$ can be obtained. For our pseudo-impulsive sources, the output signal $s(t)$ has been taken as impulse response $h(t)$. Using the integrated impulse response method suggested by Schroeder, it has been stated that both the mean values and standard deviations for some parameters are practically identical to results obtained with other usual techniques. (To be presented in Spanish.)

10:45

5aAA9. Sound absorption in a scale model reverberation chamber. Ranny Nascimento, Moyses Zindeluk, and Jose Flavio Feiteira (Acoust. and Vib. Lab., Mech. Eng. Dept., Federal Univ. of Rio de Janeiro, CP 68503, 21945-970 Rio de Janeiro, Brazil)

Scale models of rooms are widely used to investigate architectural acoustics. However, when building a room model, the choice of materials to simulate the surfaces of the original room should follow carefully the similarity conditions. In practice, it is very difficult to model the acoustic parameters of the materials used in room scale models. This work consists in the construction and validation of a small-scale reverberation chamber, according to the international standards, developed in order to improve similarity conditions in scale modeling. Such a chamber is relatively low cost and its configuration can be easily modified. The volume of the model reverberation chamber is 0.389 cubic meters, equivalent at 1:8 scale to a real volume of 200 cubic meters, a standard size for reverberation chambers. Made in wood with one transparent acrylic side, the one-eighth scale reverberation chamber can be used to test theories about room acoustics as well as to test absorbing materials. Experimental results for the chamber qualification are presented and the search for octave band similar materials is reported.

11:00

5aAA10. Normal incidence sound absorption measurement of individual patches in a reverberation room. Paulo Massarani and Swen Muller (Acoust. Testing Lab., Inmetro, Av. N. Sra. das Gracas, 50 Xerm, 25250-020 Duque de Caxias, RJ, Brazil)

This work presents results of sound absorption measurement of plain patches inside a reverberation room. An *in situ* method has been applied (ISO 13472-1). The main purpose is to support the installation of diffusers in a reverberation room for random incidence sound absorption measurements. The *in situ* method is based on impulse response evaluation, in which the incident and reflected front waves can be windowed to isolate them from the room reflections. Due to the small areas of the patches (around 1 square meter) and the high reverberation time, some special steps had to be performed. In order to establish an impulse response measurement with high SNR, a long sweep sine signal with low-frequency emphasis has been used as the excitation signal. Applying the subtraction

technique permitted locating the microphone very close to the sample. The sound absorption coefficient of each patch allowed selecting the most uniform set of material to compose the area required by the ISO 354. The patches not used were tested in an impedance tube, showing a good agreement between the results from the two methods. Thus, the *in situ* method allows quick measurement of plain samples laid on the room floor while preserving the material integrity.

11:15

5aAA11. An experimental study of the variation of absorption coefficients for fiberglass with altitude. Richard D. Godfrey (Owens Corning, Sci. and Technol., 2790 Granville Rd., Granville, OH 43023)

ASTM C 423 specifies requirements for air temperature and relative humidity during decay rate measurements, but does not address variations in air density even though this parameter decreases with increasing altitude. At constant temperature, air density decreases by approximately 20% from sea level to 5000 feet altitude, resulting in about 20% decrease in its acoustic impedance. In a previous paper, an analysis based on Mechels design charts (normal and diffuse) showed significant changes in predicted absorption coefficients with altitude. Due to the difficulties in depressurizing a reverberation chamber, a normal incidence study was conducted to experimentally investigate this effect by placing an impedance tube in a vacuum chamber. Fiberglass materials were tested at various amounts of vacuum. The measured results followed reasonably well those predicted by the analysis for normal incidence. Based on these results, comparison of absorption coefficients measured at substantial altitude with those measured at sea level is discouraged. It is suggested that notes be added to the ASTM standards cautioning users about these limitations or limiting their use to ambient pressures within a couple inches Hg of standard sea level conditions.

11:30

5aAA12. Comparison of the uncertainty budget between the one- and the two-microphone method for reflection and absorption coefficient. Adolfo Esquivel D. (División de Vibraciones y Acústica, Centro Nactional de Metrologia, km. 4.5 carr. a los Cués, El Marqués, Qro., México, aesquive@cenam.mx) and Lorenzo M. Muñiz

The impedance tube is mostly used to measure the reflection and absorption coefficient at normal incidence based on ISO 10534-2. It is possible to use the one-microphone or the two-microphone method, where either a pseudorandom or random signal is used to produce a standing wave pattern inside the impedance tube in order to measure the transfer function and obtain the reflection and absorption coefficient. The transfer function in both methods of measurement can be expressed in three different mathematical ways: (a) by the transfer function given by the signal analyzer itself, (b) by the auto-spectrum and the cross spectrum between the two-microphone positions, and (c) by the coherence between the two-microphone positions. The reflection and absorption coefficient as a function of those variables, and its measurement uncertainty is discussed considering those three cases of the transfer function for the two methods of measurement in the frequency range from 50 Hz to 1.6 kHz. Likewise the analysis of the uncertainties sources and its sensitivity coefficients are studied using a linear approach model for both techniques, and for each of the proposed cases. (To be presented in Spanish.)

FRIDAY MORNING, 6 DECEMBER 2002 CORAL KINGDOM 2 AND 3, 8:00 TO 11:25 A.M.

Session 5aAB

Animal Bioacoustics: Bioacoustics as a Conservation Tool I

Renata de Sousa Lima, Chair
Lab. de Mamiferos Aquaticos LMA, INPA, C P 478, Manaus AM CEP 69083 000, Brazil

Chair's Introduction—8:00

Contributed Papers

8:05

5aAB1. Vocalization of migrating humpback whales over 14 years as an indicator of stock size. Douglas H. Cato (Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont 2009, Australia), Robert Paterson, and Patricia Paterson (Indooroopilly, Qld 4068, Australia)

Acoustic monitoring of migrating humpback whales has been carried out since 1981 in conjunction with visual surveys from Point Lookout, Stradbroke Island, on the east coast of Australia. Recordings were made from a drifting boat a few kilometers seaward of the observation point, during the peak of at least one of the migrations each year. This paper presents an analysis of song vocalization in relation to the observed movements of humpback whales in order to determine the vocalization rates and their value as an index of relative abundance. The proportion of whales singing as they passed Point Lookout was ~5% during the northern migration (towards the breeding grounds) and ~13% during the southern migration, the difference being statistically significant. There was no significant trend in the proportion of singers as the stock size increased by more than a factor of 3. From 1982 to 1993, the number of singers passing per 10 h during the southern migration increased at a rate of 10.6% (95 confidence interval 3%–19%), consistent with the rate of increase of 11.7% obtained from the visual survey over a similar period.

8:20

5aAB2. Acoustic surveys of humpback whales: Calibration experiments off the east coast of Australia. Michael J. Noad (Ecology Ctr., School of Life Sci., Univ. of Queensland, Queensland 4072, Australia) and Douglas H. Cato (Defence Sci. and Technol. Organisation, Pyrmont NSW 2009, Australia)

Male humpback whales are frequently heard singing during migration and on their breeding grounds, often continuously for many hours, and the sounds are audible for tens of kilometers. Acoustic surveying may be useful in quantifying whale movements in oceanic waters beyond the range of land surveying, and it might also be a useful addition to visual monitoring in areas where this is possible. For acoustic surveys to be of use, the acoustic cues of the whales must be quantified and calibrated against the numbers of whales in an area. A combined visual and acoustic survey of whales migrating close to shore was conducted off the coast of southeast Queensland. Song activity, measured using two indices, correlated significantly with the number of passing whales determined visually. Linear regressions of daily measures of song activity with numbers of whales seen lead to estimates of ratios of singers with whales seen at 0.127 ± 0.027 (95% confidence interval) and singer hours with whales seen at 0.288 ± 0.065. The possible use of these indices for conducting stand-alone acoustic surveys will be discussed.

8:35

5aAB3. 1998–99 blue whale call abundance off the central California coast using a bottom-mounted hydrophone. Anurag Kumar (Moss Landing Marine Labs., Moss Landing, CA 95039), Christopher W. Miller, and Ching-Sang Chiu (Naval Postgrad. School, Monterey, CA 93943)

Blue whale vocalizations were monitored off the central California coast during 1998–99 using the Naval Postgraduate Schools (NPS) Ocean Acoustic Observatory (OAO), a former US SOSUS array. The density of B calls were estimated using a matched filter detector based on a deverberated blue whale B call recorded from a towed array [T. C. Moore, Naval Postgraduate School, thesis (1999)]. The detectors' performance characteristic was factored into the estimation to come up with the associated errors. Diurnal, seasonal, and inter-annular trends in call density were investigated. Also, the inter-call interval of the B call was measured. An unexpected large call density was observed during the first week of January in 1999. Historically, blue whales are sighted in Monterey Bay from approximately late July to October. Results of the blue whale call density estimation are discussed.

8:50

5aAB4. A comparison of visual and acoustic marine mammal monitoring methods. Adam Frankel (Marine Acoust., Inc., 901 N. Stuart St., #708, Arlington, VA 22203)

Concern about the potential effect of increased oceanic noise on marine mammals has led to the consideration of a variety of at-sea marine mammal monitoring methods. A recent marine mammal survey utilized both visual observation and passive acoustic monitoring. Visual observations were conducted using 7×50 binoculars during the day and generation III night vision devices (NVDs) at night. Acoustic data were collected with a towed hydrophone array and analyzed with a customized, PC-based acoustic workstation. Seventy-seven visual sightings were made. Forty sightings were made at night and 37 were made in daylight. The effective range of the NVDs was estimated through observations of a target at known distances; 50% detection rating was achieved at 130 meters. There were 98 acoustic detections. Of the 77 visual sightings, 42 were also detected acoustically. Six sightings were made without comparable acoustic detection, and 19 sightings were made when there was no acoustic monitoring. Conversely, 56 acoustic detections were made without visual sightings. These data suggest that, for species that commonly vocalize, a combined acoustic/visual survey will increase detection probabilities. Furthermore, night vision devices have the potential to be an effective observation tool for marine mammals.

9:05

5aAB5. Passive three-dimensional tracking of sperm whales using two towed arrays during the 2001 SWAMP cruise. Aaron Thode (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238), David K Mellinger (Oregon State Univ., Newport, OR 97365), and Anthony Martinez (Southeast Fisheries Sci. Ctr., Natl. Marine Fisheries Service, Miami, FL 33149)

The capability to passively track multiple sperm whales simultaneously in three dimensions can provide useful information about the animals' acoustic interactions, sound production mechanisms, and potential responses to anthropogenic noise. In 2000 this capability was demonstrated using a single towed array in the Gulf of Mexico, by exploiting the multipath information from surface and bottom returns. The relative paucity of bottom returns, and the difficulties involved in automatically detecting these returns, limited the utility of that method. In this presentation data collected during the 2001 SWAMP cruise in the Gulf of Mexico are used to explore how two arrays deployed serially behind a survey vessel can provide a means for automatically extracting all the relevent information required for 3-D tracking. Using the signal-processing program Ishmael, and segments of time where two whales are vocalizing simultaneously, it is shown that enough information can be extracted to track the animals. However, the problem complexity is increased by sound-speed refraction effects and the lack of information about the array depths. [Work supported by SE Fisheries, National Marine Fisheries Service, Minerals Management Service, and the ONR.]

9:20

5aAB6. Automatic detection of sperm whale sounds from the Gulf of Alaska. David K. Mellinger (Cooperative Inst. for Marine Resources Studies, Oregon State Univ., 2030 S. Marine Science Dr., Newport, OR 97365, mellinger@pmel.noaa.gov), Kate Stafford (National Marine Fisheries Service, Seattle, WA 98115), Christopher G. Fox (Pacific Marine Environ. Lab., Newport, OR 97365), and Sue E. Moore (National Marine Fisheries Service, Seattle, WA 98115)

Records from whaling ships in the Gulf of Alaska indicate that sperm whales were present in summer but were much less common in winter. These records could have been biased, perhaps strongly, by seasonal conditions: ships were less likely to be present in winter than summer, and sighting conditions may have been worse in winter. To obtain an independent and newer estimate of sperm whale seasonality, six autonomous hydrophone recorders were deployed for one year. After retrieval, recordings were analyzed automatically to detect sperm whale sounds. The detection technique comprised several successive stages: spectral transform, noise cancellation, band-limited energy summation, periodic autocorrelation, and thresholding. The autocorrelation step was highly effective for distinguishing between sounds of sperm whales and noise sources. The detection threshold was set so that, in test data, no noise sources were detected and approximately 2/3 of known sperm whale sounds were detected. Use of such a high threshold allows for an accurate index of sperm whale seasonal and geographic occurrence. Results indicate that sperm whales are indeed present in winter in the Gulf of Alaska. Use of automatic acoustic detection allows results to be replicated in the future.

9:35

5aAB7. Automated two-dimensional passive tracking of free-ranging dolphins using two towed arrays and frequency-domain beamforming. Aaron Thode (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr., La Jolla, CA 92093-0238, thode@mpl.ucsd.edu), Eric S. Howarth (San Diego State Univ., San Diego, CA 92122), Anthony Martinez, and Jack Stamates (Natl. Marine Fisheries Service, Miami, FL)

Dolphin whistles form time-frequency contours that are complex, overlapping, and unpredictable. These features make automatic detection and localization of whistles difficult and labor-intensive whenever standard cross-correlation methods are used. A technique is presented for automatically extracting ranges and bearings of multiple free-ranging dolphins by post-processing acoustic data from two towed arrays. Each array had two elements spaced 2 m apart, and the distance between towed arrays was 350 m. The data are processed in the frequency domain, a natural framework for automation that also permits individual bearings to be extracted from whistles that overlap in the time domain. The frequency contours of the whistles are also automatically obtained as a byproduct of the analysis. Data from a February 2002 survey cruise off the Atlantic seaboard are used to illustrate the process. [Work supported by SE Fisheries, National Marine Fisheries Service.]

9:50

5aAB8. Recording spider monkeys with a microphone array. M. Wahlberg (Ctr. for Sound Commun., Aarhus Univ., Zoophysiology, C. F. Fellers Alle Bldg. 131, 8000 Aarhus C, Denmark, magnus.wahlberg@biology.au.dk), G. Ramos-Fernandez (Pronatura Peninsula de Yucatan A.C., Merida Yucatan 97100, Mexico), and F. Ugarte (Whale Works, Denmark)

The potential of using a GPS-linked microphone array [J. Acoust. Soc. Am. **109**, 434–437 (2001)] to monitor spider monkeys (*Ateles geoffroyi*) In the rainforest of the Yucatan peninsula, Mexico, was assessed. This species is of concern for conservation because of their requirement of large areas of relatively undisturbed tropical forest. The array was composed of six independent nodes. Each node consisted of a DAT recorder, a microphone, a GPS receiver and a small computer. The GPS data, containing both timing and location information, was converted into an audio signal and recorded on the tape parallel to the microphone signal. In this manner the recordings were synchronized and the receiver locations logged. The recordings were used to locate vocalizing monkeys. We measured the apparent source level and directionality of monkey vocalizations, and the transmission loss through the rain forest canopy. The system has a potential for studying the behavior of spider monkeys, as well as other sound-producing animals. This could have useful applications for the monitoring of important species for conservation. [Work supported by the Danish National Research Council and Pronatura, Mexico.]

10:05–10:25 Break

10:25

5aAB9. St. Lawrence blue whale (*Balaenoptera musculus*) vocalizations. Catherine Berchok (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, cberchok@yahoo.com), David Bradley, Thomas Gabrielson (Penn State Univ., State College, PA 16804), and Richard Sears (MICS, Inc., QC J4P 1T3, Canada)

The Mingan Island Cetacean Study has been conducting biological field research on St. Lawrence blue whales since 1979. Currently, they have photo-identified 384 individuals and determined the gender of 37% through skin biopsy sampling. In a typical field season, approximately 85% of the 60–90 blue whales identified have been seen before, so a larger percentage (65%) have known gender. During four field seasons (1998–2001), 134 h of acoustic recordings were made in the presence of this well-studied population of blue whales. A calibrated omni-directional hydrophone, suspended from a surface isolation buoy, was deployed from an inflatable boat positioned close to the whales. Simultaneous visual observations were made during the recording sessions, with individual identities confirmed through photo-identification work between sessions. The primary field site for this project was the lower estuary region of the St. Lawrence River, with most recordings made between mid-August and late October. Both infrasonic tonal call sequences as well as aubible downsweep calls were recorded. Temporal and spatial distribution of the two call types will be presented, and possible correlations with behavior and/or social grouping will be discussed. [Work supported by ARL/PSU, NSF, and the American Museum of Natural History.]

10:40

5aAB10. Geographic variations in the whistles of spinner dolphins (*Stenella longirostris*). Carmen Bazua-Duran (CCADET, UNAM, Cd. Univ., 04510 Mexico, D.F., Mexico, bazua@servidor.unam.mx), Whitlow W. L. Au (Univ. of Hawai'i, Kailua, HI 96734), and Julie N. Oswald (Univ. of California, La Jolla, CA 92093-0208)

In order to better understand the reasons for geographic variations in the whistles of spinner dolphins (*Stenella longirostris*), recordings made in the Eastern Tropical Pacific will be compared to those recorded off Midway Atoll, Northwestern Hawai'ian Islands, the island of O'ahu, main Hawai'ian Islands, and Mo'orea, French Polynesia using discriminant function analysis. The values of 10 variables were extracted from the fundamental frequency (contour) of each whistle selected for analysis (begin, end, maximum, minimum and peak frequencies, center and peak times, duration, and number of turns and steps). Geographic variations in the whistles of spinner dolphins were found when comparing animals recorded in waters of Midway Atoll, off the main Hawai'ian Islands, and off Mo'orea. Greater differences were found between the Midway and Hawai'i and Mo'orea groups. The distinctiveness of Midway whistles could be due to geographic isolation and/or differences in population structure (habitat adaptation). The results indicate that geographic differences may not occur solely due to geographic isolation. Other factors, such as population structure and group size may be important when looking at differences in the whistles of spinner groups.

10:55

5aAB11. Vocal repertoire of the freshwater dolphins *Inia geoffrensis* and *Sotalia fluviatilis* in Colombia, South America. Maria Claudia Diazgranados and Fernando Trujillo (Omacha Foundation, Carrera 5 #26a-47 (506) Bogot, Colombia, omacha@cable.net.co)

The Omacha Foundation conducted a project to identify and compare vocal repertoires of the freshwater dolphins *Inia geoffrensis* and *Sotalia fluviatilis* in the Amazon and Orinoco Rivers in Colombia. Field recordings of 2457 min. were made using an omnidirectional hydrophone connected to a cassette recorder. Six different types of sounds were identified for *Inia geoffrensis humboldtiana* in the Orinoco River. Low-frequency whistles were recorded (fundamental signal around 3 kHz, maximum frequencies of 13 kHz), when foraging/socializing dolphin groups consisted of more than five animals. Screams, squeaks, pulsed sounds, low-frequency sounds, and echolocation clicks were identified as additional sound types for the species. In the Amazon region, where both species of river dolphins are found, recordings were only analyzed when the species occurred separately. Four different types of vocalizations including echolocation clicks, screams, squeaks, and pulsed sounds were identified for *Inia geoffrensis geoffrensis*. No clear whistles were heard for this species. *Sotalia fluviatilis* whistled constantly, even when only two dolphins were sighted. Whistles varied in frequency contour and duration. Generally high frequencies were observed with maximum values of 21.7 kHz and a fundamental signal of 15.8 kHz. Squeaking sounds, echolocation clicks, and burst pulses were also recorded. (To be presented in Spanish.)

11:10

5aAB12. Marine Mammal Resource. Nick Felgate (Defence Sci. and Technol. Lab., Farnborough, Hampshire GU14 OLX, UK)

In the sonar world there is still a heavy reliance on the human ear for detection. To detect threats, the sonar operator must be familiar, not only with their acoustic signatures but also with the many and varied noises emanating from other sources in the ocean. These include noises of biological origin, particularly marine mammals. The Marine Mammal Resource is a hypertext document aimed at providing a reference set of sounds for training purposes. It contains textual descriptions on 18 different aspects of ecology for each of the species, summarizing the information available from the open literature. Subjects include vocalizations, habitat, distribution, hearing, and physical characteristics. The resource also contains images, diagrams, and movies for identification purposes. Data has been collated for over 40 species primarily found in the Atlantic Ocean. Information has been collated through a number of research organizations in America and Europe. The resource has application to the reduction of cetacean by-catch in trials, both in providing information in the derivation of environmental impact assessments pretrial and in the acoustic monitoring of the area during the course of a trial.

Session 5aAO

Acoustical Oceanography: Acoustic Remote Sensing

Bruce D. Cornuelle, Chair
Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0225

Contributed Papers

8:00

5aAO1. Use of high frequency acoustics to study the spatial distributions of bubbles. Eric Terrill and Ken Melville (Marine Physical Lab., Scripps Inst. of Oceanogr., Mail Code 0213, La Jolla, CA 92093-0213)

It is well understood that the presence of bubbles in the ocean can dramatically change the sound speed, attenuation, and scattering of underwater sound over a range of frequencies. Over the last few decades, rational theories have been developed and tested that describe the complex dispersion of sound through spatially homogeneous bubbly mixtures as a function of the sizes and densities of bubbles. However, it is clear that in the ocean, the size distribution of the bubbles will evolve with a number of different temporal and spatial scales as a result of both the physics which govern their formation (breaking wind waves, breaking waves in the surf, rain generated bubbles, or ship wakes) and the physical processes which control their lifetimes: turbulent mixing, bubble rise speed, and gas dissolution. One approach to measuring their distribution in space is the application of O(1) MHz range-gated sonars which can resolve scales of O(1) cm. These high frequencies are also advantageous to work with since the scattering cross section of the bubbles is approximately proportional to the second moment of the size distribution. Results of field measurement efforts to characterize the spatial scales of bubbles from O(1) m to O(1) cm will be presented and discussed. [Work supported by the ONR Underwater Acoustics and Marine Optics programs.]

8:15

5aAO2. Measurements of ambient bubble populations with a multibeam sonar. Thomas C. Weber, David Bradley, R. Lee Culver, Anthony P. Lyons, and Steven G. Adelman (Appl. Res. Lab. and the Grad. Prog. in Acoust., P.O. Box 30, State College, PA 16804, tcw141@psu.edu)

A 240 kHz multibeam sonar was deployed on the research platform FLIP off the coast of San Diego for two weeks in July, 2002. The sonar was oriented to measure backscattering from the ambient bubble field near the ocean surface. A bubble population density spectrum shape described by Novarini *et al.* [IEEE J. Ocean. Eng. **23**, 423–438 (1998)] is assumed, and the spectrum level is shifted up or down to fit the backscatter measurements. This allows the bubble population to be inferred from the single frequency measurement. The wide field of view for the sonar (1.5° beams distributed over a 150° swath) allowed the spatial structure of the bubble field to be observed over the time period, and compared with environmental measurements including wind/wave conditions and the spatial variation of temperature and salinity. [Work supported by ONR Code 321US.]

8:30

5aAO3. On the multifrequency method of the range dependent transversal current monitoring. Mikhail Charnotskii, Goseph Fuks, Konstantin Naugolnykh (NOAA, ETL/ZELTECH, 325 Broadway, Boulder, CO 80305), and Daniela Di Iorio (Univ. of Georgia, Athens, GA 30602)

Transverse flow of inhomogeneous current produces fluctuations of the acoustic signal passing through it. These fluctuations vary with the signal frequency due to the variation of the Fresnel zone size. Respectively, fluctuations of signals at two different frequencies are coherent in a low-frequency range of the spectrum and noncoherent in the high-frequency band when ocean inhomogeneities are smaller than the transverse size of the Fresnel zone. The frequency cutoff of the coherence function of signals at different frequencies depends on current flow characteristics. Di Iorio and Farmer [J. Acoust. Soc. Am. **100**, 814–834 (1996)] describe high-frequency acoustic scintillation measurements in a tidally forced flow through a Cordova Channel specifically designed to characterize the two-dimensional angle of arrival fluctuations in terms of the mean and turbulent properties of sound speed and current. We present new results from this data set which was collected in an alternating dual frequency mode. The multifrequency technique gives results which are consistent with previous results, but eventually will allow the measurement of current profiles across the channel.

8:45

5aAO4. Managing uncertainty in ocean-acoustic tomography. Jason Rudzinsky, Subramaniam Rajan, James Lewis, and Peter Stein (Sci. Solutions, Inc., 99 Perimeter Rd., Nashua, NH 03063, jrudzinsky@scisol.com)

The Pacific Missile Range Facility, located off the west coast of Kauai, offers a unique opportunity to perform perpetual tomographic imaging of the sound-speed structure in the region. The range covers over 1100 square miles of ocean and features 172 bottom-mounted hydrophones and 15 bottom-mounted projectors. Data collected during an August, 2001 feasibility experiment suggest that a small subset of the total number of acoustic propagation rays connecting sources and receivers can be used to image prominent oceanographic features such as internal tides. However, acoustic travel time fluctuations occurring over durations of only a few seconds are a potentially significant source of error. Uncertainty in bathymetric relief also prohibits effective utilization of ray paths with multiple surface interactions in the inversion. Results of research intended to better manage these and other uncertainties are presented. The basis of this research is the use of repeated measurements, made every few seconds over many hours, along with a sophisticated ocean circulation model of the region and error decomposition techniques to isolate errors in assumed bathymetry and travel time noise. Inversion improvements as well as approaches to assimilating tomographic measurements into the ocean model are also discussed. [Funding provided by CEROS and ONR.]

9:00

5aAO5. Adjoint methods for two-dimensional inversion and error analysis. Paul Hursky, Michael B. Porter (SAIC, 1299 Prospect St., Ste. 305, La Jolla, CA 92037), Bruce D. Cornuelle, William S. Hodgkiss, and William A. Kuperman (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093)

Adjoint methods provide a rapid technique for mapping observation errors back to the environmental errors that caused them. For instance, the acoustic pressure field can be measured in a complicated, unknown two-dimensional (2D) environment, and an acoustic model can be run for a simpler, presumed environment. The discrepancy can then be back-propagated to provide a corrected ocean structure. The adjoint model is particularly attractive when the observation space is small (few hydrophones) and the environmental space is large (thousands of points characterizing the range-depth sound speed field). In earlier work, we demonstrated this process for a range-independent scenario involving tracking of internal tides. Here we extend the technique to full 2D inversion. We also illustrate how the adjoint model can be used to construct a so-called representer for the range-depth sound speed function. The representers can be used to quantify the relationship between errors in the observations and errors in the resulting inversion. They can also be used to optimize the placement of assets such as XBT's, identifying locations where the sound speed has the most influence on the current observations. We will discuss these approaches and demonstrate them using a 2D scenario in which we image a solibore.

9:15

5aAO6. Remote sensing of sandy sediment ripples using high-frequency backscatter. Kevin Williams and Dajun Tang (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA)

The presence of bottom ripples changes the acoustic interaction with sediments. Two important ripple parameters are the height and spatial wavelength. In previous work, *in situ* measurements of ripple were made using stereophotography and electrical conductivity. Although these methods are accurate and dependable, they can only be carried out over limited areas (the order of 15 meters on the side). Using acoustic backscattering to remotely measure the ripple field, on the other hand, is fast and can cover much larger areas (the order of 50 m by 50 m). Data from a 300 kHz system with two closely spaced receivers was used to estimate ripple wavelength and height. Data analysis and numerical simulation show that, as long as the wavelength of the sound wave is much shorter than the ripple wavelength, it is straightforward to resolve the wavelength of the ripple field. To measure the ripple height, however, is more difficult. The best results were obtained when the horizontal projection of the sound beam was neither parallel nor perpendicular to the ripples. In simulations a composite model was used, where the large scale ripples and small scale roughness are treated as Kirchhoff and Lambertian scatterers, respectively. [Work supported by ONR-USA.]

9:30

5aAO7. Shear wave velocity measurements in high porosity marine sediment models using radiation impedance method. Masao Kimura (Dept. of Geo-Environ. Technol., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka 424-8610, Japan, mkimura@scc.u-tokai.ac.jp)

Acoustic measurements of surficial high porosity marine sediment are important to investigate the coastal environments. To do this, the measurement of shear wave velocity rather than longitudinal wave velocity is useful for determining the physical properties of surficial high porosity marine sediment. However, the direct measurement of shear wave velocity in high porosity marine sediment is very difficult, because the attenuation of the shear wave in high porosity marine sediment is quite large. In this study, shear wave velocities in four kinds of high porosity marine sediment models are measured using the radiation impedance method which has been proposed. From the measured results, it is seen that the shear wave velocity of below 1 m/s can be measured, and the shear wave velocity increases, as the porosity decreases. The relationship between the shear wave velocity and the porosity are different depending on the kind of marine sediment model.

9:45

5aAO8. Geoacoustic inversion of tow-ship noise via near-field matched-field processing. David J. Battle, Peter Gerstoft, William A. Kuperman, William S. Hodgkiss (The Marine Physical Lab., Univ. of California, San Diego, La Jolla, CA 92093-0238, davidb@mpl.ucsd.edu), and Martin Siderius (SACLANT Undersea Res. Ctr., La Spezia, Italy)

The estimation of geoacoustic parameters of marine sediments and sub-sediment layers involves the measurement of acoustic fields in the ocean waveguide for comparison with accurate forward propagation models. This paper discusses self-noise geoacoustic inversion of tow-ship noise data acquired via a horizontal towed array from the standpoint of near-field matched-field processing (MFP). In particular, the sensitivity of the reconstructed source power to perturbations in the parameters of a short-range propagation model are quantified. Through simulations and experimental results, it is shown that even very quiet ships radiate sufficient noise power to enable self-noise inversion of basic geoacoustic parameters such as effective bottom velocity. The experimental results presented are particularly encouraging in view of the high level of interference shown to be tolerated from nearby shipping.

10:00–10:15 Break

10:15

5aAO9. Acoustic remote sensing from aboard the Royal Caribbean Cruise Lines passenger ship *Explorer of the Seas*. David R. Palmer, Shailer R. Cummings, Jr., S. Jack Stamates, Peter B. Ortner, and Douglas Wilson (NOAA/Atlantic Oceanogr. and Meteorological Lab., 4301 Rickenbacker Cswy., Miami, FL 33149, David.R.Palmer@noaa.gov)

The NOAA/Atlantic Oceanographic and Meteorological Laboratory, Royal Caribbean International, and the University of Miami, with support from the National Science Foundation and the Office of Naval Research, have established modern, well-equipped oceanic and atmospheric research laboratories on the cruise ship ***Explorer of the Seas***. This novel relationship between commercial, academic, and federal maritime communities is enabling scientists to collect a number of unique data sets. The instrumentation in the oceanic laboratory includes two RD Instruments Inc. Ocean Surveyor Acoustic Doppler Current Profiler's (ADCP's) for monitoring ocean currents and biomass, and seawater sensors for measuring temperature, salinity, chlorophyll-a, dissolved organic material, and dissolved oxygen. Data are collected weekly along the ship's two repetitive cruise tracks through the Bahamas and the Caribbean. We summarize the acoustic observations made with the ADCP's during the first year of operation of the laboratories. These observations are applicable to ongoing and planned regional research projects including those focused on ocean current systems and frontal boundary dynamics.

10:30

5aAO10. A self-navigating, unmanned sailboat for studies in acoustical oceanography. David R. Palmer (NOAA/Atlantic Oceanogr. and Meteorological Lab., 4301 Rickenbacker Cswy., Miami, FL 33149, David.R.Palmer@noaa.gov), Ernest W. Schlieben (Von Research Engineers and Consultants, Ewing, NJ 08628), and James C. Hendee (NOAA/AOML, Miami, FL 33149)

Based on a patent by one of us (E.W.S., Patent No. 3, 556, 035) RCA developed during the 1970's the concept of a self-navigating, unmanned, sailboat. It was designed to be deployed from a ship or from shore, to sail for thousands of miles on a predetermined course, to collect upper ocean and meteorological data while station keeping, to telemeter those data back to shore by satellite, and then return to the point of departure, or some other point, up to a year later. It could carry any equipment ordinarily placed on buoys and could function in hostile weather and climatic conditions. This vessel seems ideally suited for the measurement of a number of ocean parameters using acoustical methods. We discuss the characteristics of the prototype that was built and tested with U.S. Navy support and compare its properties with Unmanned Air Vehicles (UAVs) and Autonomous Underwater Vehicles (UAVs) that are presently being used or are under development. We describe our plans for resurrecting the RCA effort and, in particular, a sea trial designed to test the automatic enroute sailing and station-keeping of the vessel in an ocean environment using GPS and modern data transfer capabilities.

10:45

5aAO11. Propagation through a stratified ocean wave guide with random volume and surface inhomogeneities, Part I. Theory: Attenuation, dispersion, and acoustic mirages. Purnima Ratilal and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, makris@mit.edu)

Analytic expressions for the mean field propagated through a stratified ocean with random volume or sufrace inhomogeneities of arbitrary size compared to the wavelength are derived from a wave guide scattering model stemming from Green's theorem. It is found that multiple scattering through inhomogeneities in the forward direction can be succinctly expressed in terms of modal attenuation and dispersion coefficients under widely satisfied conditions. The inhomogeneities can have an arbitrary distribution in depth so that the model can realistically apply to scattering from internal waves, bubbles, fish, seafloor and seasurface roughness as well as sub-bottom anomalies. An understanding of the coherence of the forward scattered field can be gained by analogy with the formation of optical mirages in low-grazing angle forward scatter from random surfaces.

11:00

5aAO12. Propagation through a stratified ocean waveguide with random volume and surface inhomogeneities, Part II. Application: Internal waves, bubbles, sub-bottom and seafloor anomalies. Tianrun Chen, Purnima Ratilal, Michele Zanolin, and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, makris@MIT.EDU)

Analytic expressions for the mean acoustic field propagated through a stratified ocean with a distribution of random volume and surface heterogeneities that may be large compared to the acoustic wavelength, presented in this session by Ratilal and Makris, are used to study the effect of internal waves, bubbles, and sub-bottom anomalies as well as seafloor and seasurface roughness on acoustic propagation in both shallow and deep water. Oceanographic and geologic data are used to probabilistically characterize the acoustic properties of the anomalies in terms of their size, shape, sound speed and densities. The expected scattering function for the anomalies are then computed as a function of depth. These are used to determine attenuation and dispersion coefficients as well as the forward propagated field as a function of range and depth.

11:15

5aAO13. Wave number inversion in a noisy shallow-water waveguide. Didier Cassereau (Laboratoire Ondes et Acoustique, 10 rue Vauquelin, 75005 Paris, France, didier.cassereau@espci.fr) and Philippe Roux (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238)

The calculation of wave numbers resulting from the propagation of an acoustic wave in a noise shallow-water waveguide like the ocean is addressed. There are no particular assumptions regarding the acoustic properties of the waveguide; we just start from an experimental measurement of the acoustic field created by a source and measured on different points of a receiving transducer array. In 1999, a high-resolution wave number inversion algorithm was proposed and compared to the Cramer–Rao lower bounds. Both methods have been implemented and compared using the same set of simulated data. We propose a complete overview of the main characteristics of these two algorithms, in terms of the accuracy, computation time, dependence on noise, etc. The same comparisons will be presented with simulated and experimental data.

11:30

5aAO14. Nonlinear magnetoacoustic wave propagation with chemical reactions. Timothy Scott Margulies (61 Broad Reach, North Weymouth, MA 02191)

The magnetoacoustic problem with an application to sound wave propagation through electrically conducting fluids such as the ocean in the Earth's magnetic field, liquid metals, or plasmas has been addressed taking into account several simultaneous chemical reactions. Using continuum balance equations for the total mass, linear momentum, energy; as well as Maxwell's electrodynamic equations, a nonlinear beam equation has been developed to generalize the Khokhlov–Zabolotskaya–Kuznetsov (KZK) equation for a fluid with linear viscosity but nonlinear and diffraction effects. Thermodynamic parameters are used and not tailored to only an adiabatic fluid case. The chemical kinetic equations build on a relaxing media approach presented, for example, by K. Naugolnukh and L. Ostrovsky [*Nonlinear Wave Processes in Acoustics* (Cambridge Univ. Press, Cambridge, 1998)] for a linearized single reaction and thermodynamic pressure equation of state. Approximations for large and small relaxation times and for magnetohydrodynamic parameters [Korsunskii, Sov. Phys. Acoust. **36** (1990)] are examined. Additionally, Cattaneo's equation for heat conduction and its generalization for a memory process rather than a Fourier's law are taken into account. It was introduced for the heat flux depends on the temperature gradient at an earlier time to generate heat pulses of finite speed.

Session 5aBB

Biomedical Ultrasound/Bioresponse to Vibration: Acoustic Radiation Force, Transient Elastography and Optoacoustic Imaging

Ibrahim M. Hallaj, Cochair
Department of Radiology/MRI, Brigham and Womens Hospital, 221 Longwood Avenue, Boston, Massachusetts 02115

Jorge Becerra, Cochair
Av. Juan de Dios Batiz s/n, Col. Lindavista 07070, D.F. Mexico

Contributed Papers

8:30

5aBB1. Local tissue displacements induced by a focused ultrasound beam. Cheri Deng (Riverside Res. Inst., 156 William St., New York, NY 10038)

The acoustic radiation force generated by an ultrasound beam in a medium where energy density gradient exists as the result of absorption or reflection can induce local tissue movement. Through energy absorption, an ultrasound beam can also cause local temperature rise in tissue, which can induce thermoelastic displacement because of thermal expansion. The radiation-force deformation is related to tissue elasticity and has been investigated for tumor detection in ultrasonic imaging and for monitoring therapeutic lesion size and location in high intensity focused ultrasound (HIFU) applications. In our study, theoretical and numerical models are developed to simulate the displacements induced by a focused ultrasound beam such as a HIFU beam and an imaging pulse, including both the radiation force deformation and thermoelastic displacement. Temperature rise resulted from the beam is computed using a 3D finite-difference algorithm that evaluates the bioheat equation. Thermoelastic displacements are obtained from the temperature rise. Radiation-force displacements are computed using finite-difference algorithm that solves the equations of motion subjected to such internal body force. The mechanical and thermal effects are compared and studied to elucidate how tissue displacements are related to factors including exposure characteristics and tissue properties such as absorption coefficients and Young's modulus.

8:45

5aBB2. Temperature rise and safety considerations for radiation force ultrasound imaging. Bruce A. Herman and Gerald R. Harris (FDA/CDRH, 9200 Corporate Blvd. HFZ-132, Rockville, MD 20850, bah@cdrh.fda.gov)

Current models for estimating temperature increase during ultrasound exposure calculate the steady-state rise, using time-averaged acoustic output, as the worst case for safety consideration. While valid for the typically very short (microsecond) pulses used by conventional diagnostic techniques, this analysis does not necessarily correspond to a worst case scenario for the longer pulses or pulse bursts used by a new method, radiation force imaging. Radiation force imaging, employing ultrasound pulse durations up to hundreds of milliseconds, produces and detects motion in solid tissue or acoustic streaming in fluids via a high intensity beam. Models that calculate the transient temperature rise from these pulses are developed for both the bone at focus and soft tissue cases. Based on accepted time–temperature dose criteria, it is shown that for pulse lengths and intensities utilized by this technique, temperature may increase to levels that raise safety concerns for bone at the focus of the ultrasound beam. Also, the impact on this modality of the current U.S. Food and Drug Administration output limits for diagnostic ultrasound devices is discussed.

9:00

5aBB3. Ultrafast imaging of shear waves induced by the acoustic radiation force for elasticity imaging in soft tissues. Jeremy Bercoff, Mickael Tanter, Sana Chaffai, and Mathias Fink (Laboratoire Ondes et Acoustique, E.S.P.C.I., C.N.R.S. UMR 7587, Universit Paris VII, 75005 Paris, France)

Shear wave propagation induced by the acoustic radiation force is currently under investigation. In our experiments, an ultrafast scanner is used both to generate and detect shear waves propagation. The scanner is first used to focus, at a chosen location, a high intensity ultrasonic beam. The resulting acoustic radiation force induces in the focal zone a shear source. The system allows us then to apply very high ultrasound frame rates (up to 5000 images/s) in order to image the shear wave propagation, which velocity is linked to local elasticity properties of the medium. The elasticity map is recovered by solving the inverse problem. Actually, the generation of shear displacement requires high level insonifications at the threshold of FDA norms. Thus, the great advantage of such an ultrafast scanner is to recover the complete propagation by generating a unique palpation inside the medium. Measurements have been performed in several soft tissue mimicking phantoms with different elastic properties and in beef muscles. The 3-D finite differences simulations have been used to support our experimental results: the delivered acoustic power and the resulting heating induced by the acoustic focused beam can be estimated and compared with the FDA requirements.

9:15

5aBB4. Study of elastic nonlinearity of soft solids with transient elastography: Landau coefficients and finite amplitude shear wave. Stefan Catheline, Jean-Luc Gennisson, Sana Chaffa, and Mathias Fink (Laboratoire Ondes et Acoustique, E.S.P.C.I., 10 rue Vauquelin, 75231 Paris Cedex 05, France, stefan.catheline@espci.fr)

Transient elastography allows one to accurately achieve the Youngs modulus of soft tissues from the measurement of the shear-wave velocity. The Youngs modulus is a second-order elastic moduli and describes the linear behavior of a medium. In order to achieve the third-order elastic moduli in soft tissues (which characterize the nonlinearity), we have used transient elastography in an acoustoelasticity-like experiment. It consists

in measuring shear-wave velocities in a uniaxial pre-stressed medium. This technique uses an ultrafast scanner and allows one to measure thenonlinear third-order elastic Landau coefficients A, B, C in an Agar-gelatin based phantom. The several orders of magnitude found between these three constants appear to be a characteristic of soft tissues. In a second part, we present the first experimental displacements of a finite amplitude transverse wave. The results are compared to a simulation based on the modified Burgers equation. In both approaches, it appears that contrary to longitudinal wave for which even harmonics are created during the propagation, only odd harmonics are present on the displacement spectra of a finite amplitude transverse wave.

9:30

5aBB5. Ultrafast imaging of transient shear waves for breast elastography: First clinical results. Jeremy Bercoff, Stefan Catheline, Sana Chaffai, Mickael Tanter, Laurent Sandrin, Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, CNRS UMR 7587, Universit Paris VII, Paris, France), and Martine Meunier (Institut Curie, Paris, France)

Transient elastography of soft human tissues has been developed in our lab for several years. This technique consists in following the shear wave propagation induced by a low frequency pulsed vibrator located at the surface of the body with an ultrasonic imaging system. As the shear wave speed can reach up to 10 m/s, an ultrafast imaging system made of 128 fully programmable and independent channels has been developed able to produce 5000 images/s. Such frame rates are attainable by limiting the emission to a plane wave illumination and achieving the beamforming only in the receive mode. The comparison of successive images by classical cross-correlation techniques allows us to produce a 2-D movie of the displacements induced by the shear wave deep inside the body. An inversion algorithm is then used to recover the shear modulus map from the spatiotemporal data. We present here the first clinical results obtained on 20 female patients. These preliminary results of quantitative imaging of breast elasticity are very promising. The limits of the system due to shear wave diffraction are discussed and the improvements planned for the next clinical experiments are discussed.

9:45–10:00 Break

10:00

5aBB6. Study of anisotropy by transient elastography. Jean-Luc Gennisson, Stefan Catheline, Sana Chaffa, and Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI 10 rue Vauquelin, 75005 Paris, France)

From the low-frequency (150 Hz) shear-waves' velocity, transient elastography can precisely measure the Youngs modulus in isotropic soft tissues. Using polarized shear waves, the character of an anisotropic media like fiber muscle is clearly visible. The polarization of elastic shear waves is obtained by the use of a rod (3×80 mm) as a low-frequency source. Indeed, the displacement field induced within a medium when an acoustic pulse (150 Hz) is applied at the surface by a rod is no more axisymmetric: the component of the displacement parallel to the rod is smaller than the component perpendicular to the rod. These polarized shear waves are well predicted by a numerical simulation of theoretical Greens functions. It is shown that the shear-wave velocities in a beef muscle (biceps femoris) strongly depend on the position of the rod as regard to the fibers. These results show that the anisotropic model for this muscle is the hexagonal system. Results *in vivo* on man biceps also show the existence of a low and a fast shear wave. We have started on this subject a collaboration with the myology institute at the hospital La Piti Salptrire in Paris.

10:15

5aBB7. Coherent nonlinear scattering mechanisms for ultrasound-stimulated vibro-acoustic spectrography. Nicholas C. Makris, Purnima Ratilal, and Jonathan Thierman (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Physical mechanisms by which responses at the difference frequency can arise from targets insonified by sources radiating at two distinct primary frequencies, as found in ultrasound-stimulated vibro-acoustic spectrography, Fatemi and Greenleaf [Science **280** (1998)], are described from first principles using coherent field theory. Three specific mechanisms are investigated. These are (1) nonlinear interactions at the difference frequency in the primary fields co-linearly scattered from the object, (2) linear scattering by the object of nonlinear evanescent waves incident on the object at the difference frequency due to interference of noncolinear primary incident waves, (3) the linear scattering of propagating nonlinear waves incident on the object from co-linear primary waves. The relative amplitude of waves from these various mechanisms will be investigated and shown to vary with the target's size, shape, acoustic properties as well as the measurement geometry.

10:30

5aBB8. Diffraction tomography applied to simulated ultrasound through breast tissue. David H. Chambers (Dept. of Elec. and Computer Eng., Univ. of California, Santa Barbara, Santa Barbara, CA 93106-9560)

Diffraction tomography is used to obtain images of sound speed and attenuation of a slice of breast tissue obtained from the Visible Woman data set. Simulated ultrasound data was generated using an acoustic propagation code run on the ASCI Blue Pacific computer at Lawrence Livermore National Laboratory. Data was generated for a slice of healthy tissue, and a slice with simulated lesions to determine the ability of the imaging method to detect various abnormalities in the breast. In addition, the time reversal operator for the slice was constructed from the data and the eigenfunctions backpropagated into the slice as first suggested by Mast [Mast, Nachman, and Waag, J. Acoust. Soc. Am. **102**(2)] to identify structures associated with each time reversal mode for both the healthy tissue and tissue with lesions.

10:45

5aBB9. Optoacoustic probe for prostate cancer diagnostics. Valeriy G. Andreev, Alexander A. Karabutov (Acoust. Dept., Phys. Faculty, Moscow State Univ., 119992 Moscow, Russia, andreev@acs366.phys.msu.su), and Alexander Oraevsky (Univ. of Texas Med. Branch, Galveston, TX 77555)

The optoacoustic probe for prostate cancer detection was developed and tested. The 10-ns pulses of the YAG:Nd laser were delivered by an optical fiber with a turning mirror at its tip. A fiber tip was placed above an ultrasonic array which was employed for the detection of acoustic transients excited inside prostate tissue. The increased infrared light absorption inside prostate tumors resulted in acoustic pulses with enhanced peak pressure providing 200%–300% optoacoustic contrast. The transducer array and the optical fiber were wrapped inside a 20-mm diameter thin cylindrical shell filled with ultrasonic gel transparent for infrared radiation. Each acoustic transducer was made of 0.05-mm thick PVDF film with

dimensions of 1 mm×12 mm. The frequency bandwidth of transducer array provided 0.3-mm axial in-depth resolution. The lateral resolution is defined by the array length and was estimated as 0.8-mm for 32-element array with 1-mm gap between transducers. Transducer sensitivity of 0.05 mV/Pa allowed the detection of 2-mm tumor located at 50 mm depth. The optoacoustic probe performance was evaluated via the acquisition of two-dimensional optoacoustic images of small absorbing spheres in prostate-tissue phantoms. [Work supported by NIH and FIRCA grants.]

FRIDAY MORNING, 6 DECEMBER 2002 CORAL GALLERY FOYER, 8:10 TO 11:30 A.M.

Session 5aEA

Engineering Acoustics: Underwater Acoustic Devices, Techniques and Measurements

Kim C. Benjamin, Cochair
NAVSEA Newport, Code 216, B-1171, 1176 Howell Street, Newport, Rhode Island 02841

Thomas R. Howarth, Cochair
NAVSEA Newport, Bldg. 1171, Code 20, 1176 Howell Street, Newport, Rhode Island 02841

Chair's Introduction—8:10

Contributed Papers

8:15

5aEA1. Force distortion in a coil/magnet vibration exciter for shallow structures subjected to large displacements. Olivier Thomas and Antoine Chaigne (UER de Mécanique–ENSTA, Chemin de la Hunière, 91761 Palaiseau Cedex, France)

An electromagnetic noncontact vibration exciter has been designed for investigating nonlinear vibrations of shallow structures such as gongs, cymbals, and spherical shells. It is composed of two devices: a coil driven by a current and a small magnet glued on the structure. One interest of this system is that the mass of the magnet is small compared to the mass of the structure, so that its dynamic properties are only slightly modified. The magnitude of the force applied to the structure via the magnet depends on the relative position of magnet and coil. As a consequence, an harmonic distortion appears in the force wave form as the magnitude of the magnet displacement increases. A theoretical study is first carried out in order to find an optimal geometry for the coil that minimizes the force distortion. In a second step, a simple degree-of-freedom model is used to evaluate the distortion rate. It is found that this rate not only depends on the exciter properties, but also on the dynamics of the system under testing. A multiple-scale technique is used for the resolution of the nonlinear model that governs the distortion. Experimental results will be presented which confirm the theoretical predictions.

8:30

5aEA2. Design of a coreless electromagnetic projector for an A-size sonobuoy. Joseph A. Clark (CDNSWC, Code 734, Bethesda, MD 20084, clarkja@nswccd.navy.mil) and Jane A. Young (CDNSWC, Bethesda, MD 20084)

An electromagnetic projector has been designed that uses a pair of closely spaced pancake coils to produce low-frequency, high-power, steady-state and transient underwater pressure waves. The design requires no magnetic core materials with their associated weight and saturation limits. The projector is inexpensive to build, robust in construction, and small enough to fit with a power source in an A-size sonobuoy. An electromechanical circuit analysis of the design shows that high efficiencies can be achieved at low frequencies because the large surface motions, required to generate useful acoustic flows, do not cause significant stressing of the projector. [Research supported by ONR.]

8:45

5aEA3. Testing of a coreless electromagnetic projector for an A-size sonobuoy. Jane A. Young (CDNSWC, Code 725, Bethesda, MD 20084, youngja@nswccd.navy.mil) and Joseph A. Clark (CDNSWC, Bethseda, MD 20084)

The feasibility of using coreless electromagnetic coil projectors as underwater sources of coherent sound was explored by building and testing a projector in an acoustic test facility. Electrical measurements of the projector determined the resistance and inductance of the unit as a function of frequency and also the change in inductance as a function of spacing between two coils. Voltage and current measurements of the projector while driven by a power amplifier determined electrical power input, while acoustic pressure measurements near the face of the projector and at a range of 1 yard characterized the acoustic power output. Both long and short duration tonebursts were used to study the operation of the projector and to obtain data for comparison with a theoretical analysis of the projector design. [Research supported by ONR.]

9:00

5aEA4. The design, fabrication, and acoustic calibration of a 1-3 piezoelectric polymer composite-based two-dimensional array for UUV applications. Kim C. Benjamin and Jeffrey A. Szelag (Naval Sea Systems Command Div. Newport, 1176 Howell St., Newport, RI 02841)

A two-dimensional planar transducer array containing 52 1-3 piezocomposite elements was acoustically calibrated in the frequency range 10–100 kHz. The active substrate thickness was 25.3 mm giving the elements a single low-Q half-wave resonance frequency around 60 kHz. This was the only in-band resonance for this array. The construction approach described is scaleable with respect to the size and number of elements and could accommodate singly and doubly curved apertures. The injection-molded ceramic preform used in this study would allow one to realize an element center-to-center spacing of approximately 9.5 mm. Also, this construction approach would enable element apertures of arbitrary geometry (i.e., nonrectangular). This paper will present the fabrication approach, associated tooling, and measured acoustic results. It will also discuss the use of Navy Type VI piezoceramic material under high drive and high duty cycle conditions. [Work sponsored by the U.S. Navy.]

9:15

5aEA5. A plano-concave, 1–3 piezocomposite transducer for swept-frequency calibration of ultrasound hydrophones below 2 MHz. Gerald R. Harris (Food and Drug Administration, 9200 Corporate Blvd., HFZ-132, Rockville, MD 20850) and Paul M. Gammell (Gammell Appl. Technologies, Exmore, VA 23350)

Rapid calibration of hydrophones used in biomedical ultrasound is possible with swept-frequency techniques such as time-delay spectrometry. However, the range below 2 MHz has largely been neglected because of insufficient source transducer bandwidth [P. M. Gammell and G. R. Harris, J. Acoust. Soc. Am. **106**, L41–L46 (1999)]. Calibration in this range is important when recording pressure waveforms distorted by nonlinear propagation, especially for accurate measurement of the peak rarefactional pressure. Therefore, a 1–3 piezocomposite transducer designed for this range has been developed and tested (poled piezocomposite element, Smart Material Corp., Sarasota, FL; transducer assembly, UTX, Inc., Holmes, NY). The element has plane front and spherically concave back faces. The nonuniform thickness disperses the thickness resonance, and the composite structure suppresses radial-mode resonances. Also, the composite's lower acoustic impedance provides a more efficient match to water. A 4-cm-diameter transducer having maximum and minimum thicknesses of 1.5 cm and 1 mm produced a smooth pressure spectrum at an axial distance of 35 cm that varied <15 dB from 0.1–2 MHz. Compared to plano-concave, single-element designs having similar dimensions, the new transducer had equivalent transmit sensitivity, and its response was significantly more uniform, possibly due in part to decreased contributions from radial modes.

9:30

5aEA6. Piezocomposites 1–3 with two periodicities to avoid the lateral modes. Nicolas Perez and Carlos A. Negreira (Laboratorio de Acustica, Ultrasonora–Instituto de Fisica Facultad de Ciencias-Montevideo 11400, Uruguay, can@fisica.edu.uy)

In the frequency range 1–5 MHz the lateral modes of the periodic piezocomposites 1–3 disturbs the optimal performance of the transducer in the thickness mode and produces a defocalization of the ultrasonic field. In previous works nonperiodic piezocomposites 1–3 were built to suppress the Lamb lateral modes. In this paper an analysis is proposed for the vibration modes in piezocomposites 1–3–1 samples three-phase PZT–polymer-holes constituting a network with two periodicities. This kind of configuration has a system of double stop-band and greatly improves the electromechanical coupling factor. This fact reduces drastically the lateral Lamb modes and broadens the bandwidth of the principal mode. This was obtained in a different way to the usual one, that is increasing the backing, which leads to a reduction of the sensitivity. To avoid the distortion in the vibration produced by the holes, the samples were covered with a flexible thin film electrode. In order to obtain the vibration modes, the electrical admittance was measured and ultrasonic spectroscopy was done to the samples surface. These results were compared with the numerical simulations performed by finite elements. The diffracted field by this kind of structure was also experimentally analyzed.

9:45

5aEA7. A frequency-domain technique for noninvasive, ultrasonic fluid detection in submerged sealed canisters. Scott MacIntosh (Dept. of Phys., Portland State Univ., P.O. Box 751, Portland, OR 97207-0751, scott.macintosh@mail.com), Dipen N. Sinha, Gregory Kaduchak, and Christopher S. Kwiatkowski (Los Alamos Natl. Lab., Electron. and Electrochemical Mater. and Devices, Los Alamos, NM 87545)

There are situations where canisters containing specialized parts and/or materials are submerged underwater for storage or safety purposes. In some cases, leaks in the canisters can be detrimental and must be detected as early as possible. The present research describes an ultrasonic method for determining the presence of fluid in sealed, submerged canisters. In this research, the canisters are metallic and are filled with a random array of loose objects. The technique is based on two phenomena: the radiation damping effect of the internal fluid on the container wall and spectral signatures that result from internal reflections within the canister. The experiment utilizes a pair of piezoelectric transducers with a lateral separation of 27 mm. The transducers are placed at a standoff distance of 5 mm from the container wall and traverse the vertical dimension of the canister on a stepper-slide assembly. The transmitter is swept between 0.8–4.0 MHz and a frequency spectrum is collected. By analyzing the contributions to the spectrum from internal radiation damping and internal reflections separately, it can be determined whether the container wall is fluid or air backed. Current results demonstrate that fluid levels as low as 1 cm can be detected.

10:00

5aEA8. Use of complex source points to simplify Gaussian beam synthesis. Stephen Forsythe (Naval Undersea Warfare Ctr., 1176 Howell Ave., Newport, RI 02841)

It is often desirable to generate the acoustic field due to a so-called Gaussian beam. One way to do this is to use the free-space Greens function for the acoustic field and to sum small area sources over a circular plate with the appropriate shading for the desired Gaussian beam. For very high frequencies and narrow beams, the computation time to give an accurate sum can be large when calculating the sum for many points in the acoustic field. An alternate approach comes from the use of a single point source with complex coordinates $R=[Xr+iXi, Yr+iYi, Zr+iZi]$. When this complex source point is used in the free-space Greens function, the formal expressions for pressure and particle velocity can be used if careful attention is paid to the interpretation of the complex distance, r, that arises in the $\exp(ikr)/r$. The singularity is no longer a single point in the case of a complex source, but a circular disk. The far field of a complex source point is a good approximation to a Gaussian beam. Several computational uses of the technique will be demonstrated. [Work supported by ONR.]

10:15

5aEA9. Methods of experimental investigation of acoustic interaction between electroacoustical transducers in array. Boris Aronov, Tetsuro Oishi, and David A. Brown (Acoust. Res. Lab., Univ. of Massachusetts, Dartmouth, North Dartmouth, MA 02747)

Determination of acoustic loading conditions on projectors in underwater arrays is needed to predict the performance of the array. Nonuniform distribution of the acoustic loading over the radiating surface of the array is generated due to acoustic interactions between the projectors, and often causes performance degradation of the directivity patterns and the transmit frequency response. In the electromechanical equivalent circuit model of such an array, the influence of the neighboring transducers can be considered as the mutual radiation impedance that is inserted in series with the self-radiation impedance. The mutual radiation impedance depends on the configuration of the transducer, the mode of vibration, velocity distribution, and the geometry of the array. We present experimental techniques to determine the mutual radiation impedance as an intrinsic property of a given transducer and array configuration. Two experimental methods are presented: the Z method and the V method. The methods are based on the measurements of the input impedances and on the output voltages of the interacting transducers. For the two methods, analytical investigations based on the electromechanical equivalent circuit analysis will be presented along with supporting experimental results on cylindrical transducers. [Work supported in part by ONR 321SS Lindberg and BTECH Acoustics.]

10:30

5aEA10. High-frequency laser sonar system. Benjamin A. Cray, Ashwin Sarma, and Ivars P. Kirsteins (Naval Undersea Warfare Ctr., Div. 1176, Howell St., Newport, RI 02841)

A set of measurements recently completed at the Naval Undersea Warfare Center (NUWC) demonstrated that a laser-based sonar system can be used to detect acoustic particle velocity on the surface of a thin acoustically-compliant plate embedded beneath a standard acoustic window. The theoretical acoustic and measured surface particle velocity varied by less than 1 dB (reference m/s) over a wide frequency band (10 kHz to 100 kHz). However, the Polytec Model PSV-100 Scanning Laser Vibrometer System (SLVS) used in the experiments had relatively poor acoustic sensitivity, presumably due to high electronic noise, speckle noise, stand-off distance, and drifting laser focus. The laser's acoustic sensitivity appears to be inversely proportional to the backscatter signal level. The existing SLVS can sample a grid of 512 by 512 points, with each grid point having a spot size of approximately 10 mm (0.0004 in.). Such fine sampling may be used to create essentially a continuous aperture, eliminating acoustic grating lobes at all frequencies of practical sonar interest.

10:45

5aEA11. Development of a dynamic Young's modulus measurement system for compliant polymers. Francois M. Guillot and David H. Trivett (George W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, francois.guillot@me.gatech.edu)

A new system to experimentally determine the complex Young's modulus of elastomers at elevated hydrostatic pressures and as a function of temperature is presented. A sample cut in the shape of a rod is glued to a piezoelectric ceramic shaker and mounted vertically inside a pressure vessel equipped with glass windows. Two independent measurement methods are then used: a resonant technique, to obtain low frequency data, and a wave propagation technique, to obtain higher frequency data. Both techniques are implemented utilizing laser Doppler vibrometers; one vibrometer detects sample resonances through a window located at the bottom of the pressure vessel, and a set of two separate vibrometers monitors the speed of longitudinal waves propagating in the sample (by measuring symmetric lateral displacements), through windows located on the sides of the vessel. The apparatus is contained inside an environmental chamber for temperature control. Using this approach, Young's modulus data can be obtained at frequencies ranging typically from 100 Hz to 5 kHz, under hydrostatic pressures ranging from 0 to 2.07 MPa (300 psi), and at temperatures comprised between -2 °C and 50 °C. Experimental results obtained on commercial elastomers will be presented. [Work supported by ONR.]

11:00

5aEA12. Nondestructive evaluation of conformal sonar hull array manufacture and installation. John W. Doane, David H. Trivett, Jacek Jarzynski (Dept. of Mech. Eng., Georgia Inst. of Technol., 771 Ferst Dr., Atlanta, GA 30332-0405, john.doane@gtri.gatech.edu), and Mehdi Batel (Bruel and Kjaer, Norcross, GA 30071-1588)

The Conformal Array VElocity Sonar (CAVES) consists of an array of neutrally buoyant accelerometers mounted on a compliant layer that is adhered to the hull with an approximately one inch thick layer of Versathane. The performance of this sonar is dependent upon having uniform material properties and on having a bond to the hull free of voids and delaminations. Due to the high propagation loss in the compliant layer, conventional ultrasonic techniques cannot be utilized. There is also the need to directly measure the physical properties of the compliant layer since it has been demonstrated that QA samples do not have the same properties. Two non-destructive evaluation techniques are presented. In the first, a point drive is used to excite elastic waves in the compliant layer. The resulting motions are measured with a Laser-Doppler Vibrometer and an optimization technique is used to determine the complex (real and imaginary components) material properties. This method also lends itself to finding macrovoids within the compliant layer. In the second technique the vibration source excites the bond layer from a free edge. The array sensors are then used to determine the presence of bond failures. Both theory and experimental results are presented.

11:15

5aEA13. Matching of ultrasonic transducer: Interest of the glue layer. Dorothee Callens, Christian Bruneel, and Jamal Assaad (IEMN DOAE, Univ. of Valenciennes, Le Mont HOUY, 59313, Valenciennes, Cedex 9, France)

Our objective in this study is to propose a new solution to the acoustic impedance mismatch between a PZT transducer and a water load. Broad bandwidth matching to water is often perforrmed theoretically by using a quarter wave layer with a characteristic acoustic impedance of 7 MRayl. Practically, it is difficult to find material with this particular impedance. Moreover, performances of this matching can be reduced by the thickness of the glue coating between the quarter wave layer and the transducer. The approach we proposed consists in using the glue coating as a controlled matching layer. Once the glue is taken as the first matching layer, the range of available material for the second layer is very large. In our configuration, the front face matching is achieved using glass. The thickness of each matching layer is then calculated as a function of the impedance of the constitutive layer of the multilayer PZT/Layer 1: Glue/Layer 2: Glass/Water. The new two-layer matching configuration is compared to the ideal quarter wave layer matching. It will be shown that this new configuration provides performances as good as those obtained with conventional quarter wavelength matching.

Session 5aMU

Musical Acoustics: Guitar Acoustics

Thomas D. Rossing, Cochair
Physics Department, Northern Illinois University, De Kalb, Illinois 60115

Ricardo R. Boullosa, Cochair
CCADET, UNAM, Circuito Exterior Cd. Universitaria, Apdo. Postal 70-186, C. P. 04510, Mexico, D. F. Mexico

Invited Papers

9:00

5aMU1. Modal analysis of guitars. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115, rossing@physics.niu.edu)

The complex vibrations of a guitar body are conveniently described in terms of normal modes of vibration, which involve the coupled motion of the component parts, mainly the top plate, back plate, ribs, and air cavity. These normal modes are generally determined from observing operating deflection shapes at frequencies near the principal resonances. We compare the normal modes, generally obtained by holographic interferometry, of several different types of guitars by various makers. The effects of bracing pattern and other design features on acoustical behavior are discussed.

9:30

5aMU2. Acoustics of the guitar at the National Autonomous University of Mexico. Ricardo R. Boullosa (CCADET, UNAM, Circuito Exterior Cd. Universitaria, Apdo. Postal 70-186, C. P. 04510, Mexico D. F., Mexico, rrb@aleph.cinstrum.unam.mx)

Ongoing work in the field of Musical Acoustics (among other areas of interest) has been carried out at the Acoustics Laboratory at the Centro de Investigaciones Aplicadas y Desarrollo Tecnológico in the Universidad Nacional Autónoma de México (UNAM). This research began almost 15 years ago and although there have been times when no work was made, due to the effort involved in the characterization of the chambers at this laboratory and the development of other areas, some results are worth mentioning. In this paper the main work done in the laboratory in the field of musical acoustics, in particular, that related to the guitar is summarized. The topics studied so far deal with the objective and subjective evaluation of the sound of the guitar, the role of high frequency radiation, the string–body interaction and tuning problems, and the fretboard's influence on the sound radiation, among other things. Most of this work has been published elsewhere, but the aim in this presentation is to give a general view of the work done by the Acoustics Laboratory, to the author's acknowledge, the only group in México working in this area.

10:00

5aMU3. Time–domain simulation of a guitar. Gregoire Derveaux (UMA, 32 Bd Victor 75739 Paris Cedex 15, France, derveaux@ensta.fr), Antoine Chaigne (ENSTA-UME, 91761 Palaiseau cedex, France), Patrick Joly, and Eliane Becache (INRIA-Projet ONDES, 78153 Le Chesnay Cedex, France)

Our purpose in this study is to present a time–domain numerical modeling of the guitar. The method consists in the modeling of the vibratory and acoustical phenomena from the initial pluck to the 3-D radiation. An idealized plucking force is acting on a 1-D damped string model. The string is coupled to the soundboard via the bridge. The soundboard is modeled as an orthotropic heterogeneous damped Kirchhoff–Love plate, with a soundhole, clamped at its boundaries. The other parts of the body (back, neck, sides, . . .) are assumed to be perfectly rigid. The plate radiates both inside the cavity and in the external free field. We obtain a system of partial differential equations which are solved numerically in the time–domain. The fluid–structure interaction, in particular, is solved by using a fictitious domain method, which preserves the efficiency of the finite difference method for the 3-D radiation and approximates precisely the shape of the instrument. This technique allows us to observe the evolution of the acoustic field with time inside and outside the cavity. From a musical viewpoint, the computation of important quantities such as sound intensity, directivity, sound power, or energy transfer between the cavity and free field are made possible.

10:30

5aMU4. A guitar in vacuum: How does it sound? Bertrand David (Ecole Nationale Supérieure des Télécommunications, 46 rue Barrault, 75634 Paris Cedex 13, France) and Xavier Boutillon (Laboratoire d'Acoustique Musicale, Paris, France)

The loudness of the radiated sound is one prominent quality feature of guitars. However, few studies have been devoted to the ratio between the radiated energy and that which is lost in the wooden structure. In order to evaluate this ratio experimentally, yielding the acoustical efficiency of the instrument, we have compared the vibrations of a guitar at different ambient pressure levels. The guitar is installed in a large tank where the open bass string is automatically plucked. The pressure is varied between atmospheric level and vacuum, which implies a proportionally vanishing radiated power. The decay time of the most energetic component of each partial is estimated. It turns out that decay times are sometimes shorter in vacuum. This can be explained through a model of the coupling

between the string and the bridge, the important parameter being the ratio of the real part of the mechanical impedance of the bridge to its imaginary part. The efficiency is measured every 82 Hz between 80 and 2500 Hz. Values are widely scattered between 15% and 60%. A comparison will be presented with acoustical efficiencies of other structures (plates, piano soundboard) and with values obtained on guitars by other methods and authors.

Contributed Papers

11:00

5aMU5. Modal analysis of an archtop guitar. Cassandra Wiese and Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115, rossing@physics.niu.edu)

Although archtop guitars are very popular with jazz musicians, their acoustical properties have not been studied nearly as much as their flattop cousins. Their plates are carved, like those of violins or cellos, so internal bracing is minimal. Bridges are usually not fixed to the top plate. We describe the modes of vibration of an archtop guitar, determined with the help of holographic interferometry, with those of a flattop guitar. Some modes show a close relationship, while others are quite different.

11:15

5aMU6. The acoustic effects of guitar components. Ra Inta (Univ. of New South Wales, Sydney, Australia), Gerard Gilet (Gilet Guitars, Sydney, Australia), John Smith, and Joe Wolfe (Univ. of New South Wales, Sydney, Australia)

The guitar is a complex oscillatory system made up of many vibrating components. Because of the variable mechanical properties of wood, it is not easy for makers to reproduce good instruments. Reproducibility can be improved if we know how the mechanical properties of the components interact to produce the sound of the completed instrument. Three steel-string acoustic guitars were constructed, in parallel and as similarly as possible, the only design difference being the timber used for the top-plates. Prior to construction, the Young's moduli, densities, and moisture contents of a selection of top-plate brace, neck, and bridge materials were measured and the most similar were retained for creating the three instruments. Transfer functions and Chladni modes of the top-plates were measured at seven stages of construction, and the radiation patterns and acoustic efficiencies of the finished instruments measured. The effects of brace scalloping and neck attachment systems are reported. These results, and the behavior of some simple systems, are compared with finite element simulations that include scalloped bracing and glue bonding. [Work supported by the Australian Research Council and Gilet Guitars, Australia.]

11:30

5aMU7. Modal analysis of a classical guitar. David Cohen (J. Sargeant Reynolds Community College, Richmond, VA 23261, dcohen@jsr.vccs.edu) and Thomas D. Rossing (Northern Illinois Univ., DeKalb, IL 60115)

Using holographic interferometry, we have determined the modes of vibration of a classical guitar (by the first author) having an asymmetrically-braced top plate and a crossed braced back of unique design. The vibrational modes and acoustical properties are compared with other classical guitars.

FRIDAY MORNING, 6 DECEMBER 2002 CORAL KINGDOM 1, 8:00 TO 9:40 A.M.

Session 5aNSa

Noise: Pan-American Noise Standards

Samir N. Y. Gerges, Cochair
Mechanical Engineering, Federal University of Santa Catarina, Campa Universitario Trindade, Florianopolis SC 88040 900, Brazil

Robert D. Hellweg, Jr., Cochair
Compaq Computer Corporation, MR01-3/D3, 200 Forest Street, Marlborough, Massachusetts 01752

Chair's Introduction—8:00

Invited Papers

8:05

5aNSa1. Standards 101; the ASA standards program. Paul D. Schomer (Standards Director, Acoustical Society of America, 2117 Robert Dr., Champaign, IL 61821)

ASA supports the development of standards by serving as the secretariat for standards committees of the American National Standards Institute (ANSI). The program is organized through four ANSI 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) the development of National Standards (ANSI Standards), (2) the national adoption of an international standard (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, Psychological and Physiological Acous-

tics, etc. Overall, the Standards Program involves more ASA members than does any other single function of the Society except meetings and it is the biggest outreach function of ASA affecting the health, welfare, and economic well-being of large segments of the population, the business and industrial community, and government at all levels.

8:30

5aNSa2. Noise labeling in Brazil. Marco A. N. de Araujo, Paulo M. Massarani, Jose A. J. de Azevedo (INMETRO, Div. de Acoust., Av. N. Sra. das Gracas, 50-Xerem, Doque de Caixias, RJ, Brazil, CEP:25250-020), and Samir N. Y. Gerges (Federal Univ. of Santa Catarina, Florianopolis, SC, Brazil)

The Brazilian Silence Program, created in 1990 by the Brazilian Ministry of Environment, advocates the production and use of equipment with lower noise level. The subcommittee of Noise Labeling of the Brazilian Committee of Certification is composed of INMETRO acoustic specialists to organize and implement the Brazilian Labeling Program. This subcommittee elaborated the label form and test procedure. The noise-labeling program will first concentrate on the following household devices, both manufactured in Brazil or imported from abroad; mixers, blenders, hairdryers, refrigerators, and vacuum cleaners. The label should contain the sound-power level in dBA. INMETRO or other credited laboratories are responsible for the measurements. The ISO 4871, 3740 (1 to 5), ISO 8960, and IEC 704 (1 to 4) and also the equivalent Brazilian standards are used for the measurements, such as ABNT NBR 13910-1. The main objective of the label is to inform the consumer about the emitted noise level. The label offers the noise parameter to be used by the consumer when comparing devices, considering price, performance, and now also noise. No restriction for noise level was established.

Contributed Papers

8:55

5aNSa3. The Brazilian National Standards in Acoustics. Walter E. Hoffmann, Paulo M. Massarani (INMETRO, Electroacoustics Lab., RJ, Brazil), and Samir N. Y. Gerges (Federal Univ. of Santa Catarina, Cx.P. 476, Florianopolis, SC, Brazil)

The elaboration of technical standards is a difficult and expensive activity, not always necessary, since there are international standards that can be used. Therefore, the elaboration of technical standards in Brazil is, in general, linked to several programs and regulations developed in this country. Among these are the Brazilian labeling program, which intends to label domestic appliances; the approval of new car models and the annual control of the existent fleet; occupational noise control programs and the growing requirement for the accreditation of calibration and testing laboratories. In particular, in the electroacoustics area, in 1998 a Technical Committee formed of 3 Working Groups was created, which developed 11 technical standard projects relating to specific programs in the process or in the elaboration phase such as the introduction of the conformity tests for equipment used for acoustic measurements, the creation of a conformity label for audio equipment, and an intercomparision program at the national level for the calibration of audiometers. This contribution aims to show a general vision of the state of the art of Brazilian standardization as well as the several programs for which these standards will serve as support.

9:10

5aNSa4. Prevention, monitoring, and correction law for urban acoustical contamination in the province of Buenos Aires. Senator Plana, Ana Mara Rizzo La Malfa, Luis Forte (Honorable Camara de Senadores de la Provincia de Buenos Aires, Argentina), Ariel Velis, Horacio Bontti, Antonio Mendez (Laboratorio de Acustica y Luminotecnia de la Comisin de Investigaciones Cientficas de la Provincia de Buenos Aires, Argentina), Francisco Ruffa, Pablo Kantor (Facultad de Ingenieria de la Universidad de Buenos Aires, Argentina), and Gustavo Basso (Universidad Nacional de La Plata, Argentina)

In the last decades, many cities of the Buenos Aires Province, in Argentina, have grown significantly. This expansion was a disorderly process, and produced important damage to environmental quality, expressed among other things, as high levels of acoustical contamination. Not many measures have been taken in order to obtain an effective solution to this problem. The cities' governments could only take limited action using several regulations which only partially answer the people's complaints. Regarding this situation it becomes necessary to create legal regulations for the jurisdictions mentioned before, in order to include the whole matter from a holistic and systematic point of view. The purpose of this law is to create an easy tool for the cities' governments to use, thus allowing them to deal with the acoustic contamination problem in a global manner, and not in a fragmented way as it is now. The proposed law is oriented to the prevention, monitoring, and correction of the acoustic contamination for the whole provincial jurisdiction, and brings together as an innovation the fields of environment, planning, and noise. (To be presented in Spanish.)

9:25

5aNSa5. Elements for a conformity assessment system in acoustics, vibrations and ultrasound in Mexico. Salvador Echeverria-Villagomez and Alfredo Elias-Juarez (Centro Nacional de Metrologa, Mexico)

Conformity assessment in acoustics, vibrations, and ultrasound have great relevance for human health, safety, and environmental protection. Due to this fact, it is usually the government and public agencies that promote, together with the National Standards Institute (NSI) and representatives from the whole society, the development and continuous updating of standards and regulations. Besides appropriate regulations, conformity evaluation requires the existence and fitness of, at least, two other elements: adequate measuring capabilities and a proper system of consequences. The measuring capabilities are embodied in the infrastructure that goes from the National Metrology Institute (NMI) to the calibration and testing laboratories and verification units. The system of consequences, the means by which compliance with regulations can be verified and required, can be established by the same government and public agencies (GPA) of the field of work in which the regulation has been developed. In Mexico, a conformity assessment system of this kind has been evolving rapidly during the past 10 years, since the establishment of CENAM. The paper will present a proposed conformity assessment system, arising from a comparison of the Mexican system with those of other countries.

Session 5aNSb

Noise and Architectural Acoustics: Soundscapes and Community Noise I

Paul D. Schomer, Cochair
Schomer and Associates, Incorporated, 2117 Robert Drive, Champaign, Illinois 61821

Brigitte Schulte-Fortkamp, Cochair
Institute of Technical Acoustics, Technical University Berlin, Secr TA 7, Einsteinufer 25, 10587 Berlin, Germany

Chair's Introduction—10:00

Invited Papers

10:05

5aNSb1. For purposes of environmental noise assessment, A-weighting needs to be retired. Paul D. Schomer (Schomer and Assoc., Inc., 2117 Robert Dr., Champaign, IL 61821)

Currently, most countries use some form of A-weighted equivalent level (ALEQ) to assess the annoyance of most noises. But A-weighting fail to properly assess multiple noise sources such as planes, trains, and automobiles relative to one another, and it fails to properly assess sound with strong low-frequency content. It performs better outdoors than indoors even though the receivers are indoors. It certainly cannot be used for room noise criteria. A-weighted LEQ cannot assess the audibility of sound, and, in fact, LEQ in fractional octave bands cannot be used to assess the audibility of sounds at low frequencies. A-weighting needs to be retired. There are better measures for all of these functions such as loudness-level weighting using ISO 226. At low frequencies, data show some people (about one-third) are ''C-weighted'' listeners. For all noise, it may be that one model just does not fit all. Experiments show that a majority of listeners make categorical judgments and merely count events based on level with the minority of subjects fitting three other models. There are many ways to clearly move forward but we must give up our A-weighting, it has now reached old age.

10:25

5aNSb2. Not another airport noise metric. Nancy Timmerman (Nancy S. Timmerman, P.E., 25 Upton St., Boston, MA 02118-1609, nstpe@hotmail.com)

The assessment of airport noise in communities has been fraught with confusion by the proliferation of metrics. This paper will discuss the origins and uses of unusual acoustic metrics such as Time Above, Community Noise Index (CNI), Noise per Seat Index (NPSI), D-weighting, and exposure (EXP); and compare them with more ordinary ones such as Effective Perceived Noise Level (EPNL), Day–Night Equivalent Sound Exposure Level (DNL), A-weighting, maximum level (Lmax), and signal-to-noise ratio (S/N). The metrics will then be scrutinized from the technical, airport, and community viewpoints, and the strengths and weaknesses of each will be discussed. Nonacoustic factors, such as listener expectations, will also be discussed. It will be seen why so many different types of decibel have led to confusion.

10:45

5aNSb3. Physical and perceptual assessment of glass plate boundary conditions. Cathy Marquis-Favre and Julien Faure (LASH/DGCB, URA, CNRS 1652, ENTPE, rue M. Audin, 69518 Vaulx-en-Velin, France, cathy.marquis@entpe.fr)

Building structures act as input filters to outer noise pollution. Windows constitute one of the weak points of the building envelope despite their increasing acoustic performances. The latter depend on the structural parameters and have been achieved thanks to the understanding of physical phenomena. Our concern is to improve acoustic comfort by integrating sound perception and people's expectation. Actually, this work aims at studying the influence of the structural parameters on sound perception. The approach consists in investigating relationships between the vibroacoustic behavior of the structure and the perceptual attributes of sounds it transmits. The understanding of these relations will lead to improved sound quality of the structure by optimizing its structural characteristics. This communication will focus on the influence of various boundary conditions of a glass plate on sound perception. Broad-band sounds are synthesized from physical model and submitted by pairs to a panel of subjects for dissimilarity and preference judgments. Coupling of both vibroacoustic analysis and perception tests results will be presented. [Work supported by ADEME, the French Agency for Environment and Energy Management.]

11:05

5aNSb4. The role of soundscape in children's learning. David Lubman (David Lubman & Assoc., 14301 Middletown Ln., Westminster, CA 92683-4514) and Louis C. Sutherland (Consultant in Acoust., Rancho Palos Verdes, CA 90275-3908)

The soundscapes of classroom and play areas influence children's behavior and learning. For example, the sustained high noise levels of shouting children frequently reported in day care centers are averse to behavior, learning, and teaching. To some extent this Lombard effect, or reflex to the reverberant buildup of room noise, can be tamed by generous application of sound absorption treatment to room surfaces. Low ambient noise allows children to perceive and enjoy quiet sounds, such as the subtle sounds of pets

housed in classrooms. The benefits of improved interior soundscape may justify reducing reverberation time below the maximum value of 0.6 s permitted in the new classroom acoustics standard ANSI S12.60-2002. Exterior play areas are also vital to children's socialization, informal learning, and environmental awareness. Reducing noise in outdoor play areas will promote the development of soundscapes favorable to learning. This paper suggests uses of natural outdoor soundscapes to promote learning. Also shown are uses of acoustical technology for sound playgrounds, such as those pioneered by Bill and Mary Bucken. Using such devices as sound focusing structures and sound transission channels built into playground equipment, sonic playgrounds stimulate learning, increase sonic awareness, and provide experiential enjoyment of the soundscape.

11:25

5aNSb5. The uncertainty of engineering grade community noise surveys and predictions relative to two particular variables. Robert A. Putnam (Siemens Westinghouse Power Corp., MCQ2-388, 4400 Alafaya Trail, Orlando, FL 32826)

The uncertainty inherent in the estimation of the sound power level of any noise source, coupled with the variability of assumptions regarding the modeling of typical noise sources combine to yield an A-weighted sound level predicted uncertainty on the order of 3 dB(A). This will be related to the expected uncertainty of an Engineering Grade sound survey suggested in international standards, such as ISO 6190. However, since ISO 6190 uncertainty is intended to be the measurement uncertainty, the question arises as to how we are to evaluate such combinations of measurement and predictive uncertainties. The author will discuss and defend the view that such uncertainties are additive and will propose the method of combining these effects. The modeling variables to be considered will be confined to the nature of the distributions of the source sound power level over the surface of the actual sound source, and the effects such assumptions have when considered in conjunction with barriers.

11:45

5aNSb6. Acoustical sensory profiles: The bridge between measurements and preference. Richard H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138)

Our judgments about all kinds of sounds depend on the sound itself and also on our experience, situational context, and expectation. The response is multimodal in that vibration, colors, temperature, and other environmental factors have an influence on our reactions. In the limited area of product sounds there has been an effort to relate consumer judgments to measurements through the use of descriptive words to create an ''Acoustical Sensory Profile'' or ASP, particularly for products that have sound as a primary feature. Examples include musical instruments, audio products, and concert halls. More recently, other products for which sound is an ancillary feature have been evaluated using ASPs. This presentation discusses this background and projects the use of ASPs, and their relationship to physical measurements and consumer judgments in particular.

FRIDAY MORNING, 6 DECEMBER 2002 CORAL ISLAND 1 AND 2, 8:00 TO 11:30 A.M.

Session 5aPA

Physical Acoustics: General Linear Acoustics

Preston S. Wilson, Cochair
Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Elvira Viverios, Cochair
Department Arquititura (ARG)-CTC, Federal University of Santa Catarina, Campa Universitario Trindade, Florianopolis SC 88040 900, Brazil

Contributed Papers

8:00

5aPA1. Refractive acoustic devices for airborne sound. J. Sánchez-Dehesa, D. Caballero (Autonomous Univ. of Madrid, Cantoblanco, 28049 Madrid, Spain), F. Cervera, L. Sanchis, J. V. Sánchez-Perez, R. Martinez-Sala, C. C. Rubio, F. Meseguer (Unidad Asociada CSIC–UPV, Polytech. Univ., 46022 Valencia, Spain), and C. Lopez (Inst. of Mater. Sci. ICMM–CSIC, Cantoblanco, 28049 Madrid, Spain)

It is shown that a system made of periodic distributions of rigid cylinders in air acts as a new material which allows the construction of refractive acoustic devices for airborne sound. This kind of crystal, named sonic crystal, has low impedance and transmits the sound at subsonic velocities. Here, the fabrication and characterization of a convergent lens is presented. Also, the acoustic analog of the Fabry–Perot interferometer based on this crystal is analyzed. It has been concluded that refractive devices based on sonic crystals behave in a manner similar to that of optical systems. [Work supported by CICyT of Spain.]

8:15

5aPA2. Influence of a focusing lens on the electroacoustic response of ultrasonic transducers. Marechal Pierre, Levassort Franck, Tran-Huu-Hue Louis-Pascal, and Lethiecq Marc (EIVL LUSSI, F. Rabelais Univ., Rue de la Chocolaterie, BP 3410, 41034 Blois Cedex, France)

Focusing lenses are currently used in most ultrasonic transducers for medical and immersion NDT applications. In particular, high frequency medical single element transducers for applications such as skin and eye imaging require focusing lenses to obtain high resolutions. The effect of a lens on the radiation pattern has been widely studied through impulse diffraction theory, but its influence on the electrical impedance and electroacoustical response is not possible using classical models which are based on single-mode and plane wave hypothesis. Here, the effect of a lens is investigated as a function of its radius of curvature, minimum thickness, wave velocity, and acoustic impedance. The method combines a finite element code (ATILA software), which allows us to calculate the

displacements near the lens surface, with a radially symmetrical impulse diffraction calculation based on the Hankel tranform to obtain the transducer response at all locations in the propagation media. This allows to observe simultaneously electrical, electroacoustical and radiation properties of the device. Finally, a multidimensional unconstrained nonlinear minimization method is used to identify the geometrical and acoustical properties of the elements which will provide the highest overall transducer performance. [Work supported by the EU PIRAMID project.]

8:30

5aPA3. Real time focusing using an ultrasonic "one channel time-reversal mirror" coupled to a solid structure. Nicolas Quieffin, Ros Kiri Ing, Stephan Catheline, and Mathias Fink (10 rue Vauquelin, 75231 Paris Cedex 05, France, nicolas.quieffin@loa.espci.fr)

Focusing and beam steering is achieved using a time-reversal process and a unique transducer coupled to a solid structure. This low cost technique allows one to focus acoustic energy anywhere on a 3-D domain with a spatiotemporal resolution comparable to multiple transducer's array. We first record the signal emitted by the transducer and detected by a hydrophone needle at a reference point. The signal received is then time-reversed and remitted using the same transducer. At the reference point one can observes a spatiotemporal recompression. Moreover, it is shown how the experimental Green's functions at the surface of the cavity can be used to control the emitting ultrasonic field. A careful study of this phenomena leads us to better understand the resolution and the signal to noise ratio of the focusing system. Indeed the side-lobe level as well as the focal width are no more dependent on the transducer aperture but on the dimensions of the solid structure and the multiple paths covered by the acoustic waves in the solid. Those results are supported both by experiments and numerical modeling.

8:45

5aPA4. Harmonic generation of Lamb waves. Ros Kiri Ing (Laboratoire Ondes et Acoustique, Univ. Denis Diderot / CNRS UMR7587 -10, rue Vauquelin, 75231 Paris Cedex 05, France)

Lamb waves are dispersive waves that propagate following a number of distinct modes that depend on the values of the central frequency and frequency band. According to such properties and using the time-reversal process, it is shown that the hyperfocusing effect may be experienced [R. K. Ing and M. Fink, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **45**, 1032–1043 (1998)]. Such a focusing effect both relates the time recompression of the dispersive Lamb waves and constructive interference on the focus point of the modes involved. The hyperfocusing effect is interesting because it allows the amplitude of the Lamb waves to reach huge values on the focus point. In our experiments, Lamb waves with normal amplitudes of micrometer values have been achieved on the free surface of a Duralumin plate of 3 mm thickness. By analyzing the Lamb waves in the neighborhood of the focus point using the 2-D Fourier transform technique, a nonlinear process of harmonic generation is then observed—the fundamental frequency component is centered at 1.5 MHz. This nonlinear process is under study and quantified.

9:00

5aPA5. An investigation on source distribution characteristics obtained with the equivalent sources method. Yves Gounot, Ricardo E. Musafir, and Jules G. Slama (Dept. of Mech. Eng./COPPE Universidade Federal do Rio de Janeiro C.P. 68503, 21945-970 Rio de Janeiro, Brazil)

Although the equivalent sources method is frequently used to solve acoustics radiation and scattering problems, very few investigations focus on the behavior of source distributions obtained as solutions, as well as on the general properties and restrictions these solutions must fulfill. In spite of the fact that the equivalent sources are highly dependent on the direction of the incident wave, for the rigid body scattering problem, the combined use of three different source distributions, satisfying the relevant energy requirements, may provide a satisfactory source system. The advantages of using such a system are investigated using three-dimensional numerical experiments, both for monopole configurations and for an expansion in spherical harmonics. Procedures for choosing between these two types of equivalent sources for a general problem are discussed. Characteristics of the source distributions found, using two particular variants of the equivalent sources method—one using a ''least square minimization'' technique and the other being Ochmann's ''full-field equations'' method—are compared and discussed.

9:15

5aPA6. A theoretical elucidation for the experimentally observed backward displacement of waves reflected from an interface having superimposed periodicity. Nico F. Declercq (Soete Lab., Dept. of Mech. Constr. and Prod., Ghent Univ., Sint Pietersnieuwstraat 41, 9000 Gent, Belgium, NicoF.Declercq@rug.ac.be), Joris Degrieck (Ghent Univ., 9000 Gent, Belgium), Rudy Briers (KATHO, 8820 Torhout, Belgium), and Oswald Leroy (IRC, 8500 Kortrijk, Belgium)

It has been shown experimentally by Breazeale and Torbet [M. A. Breazeale and M. A. Torbett, Appl. Phys. Lett. **29**(8), 456–458 (1976)] that a backward beam displacement can occur when a beam is reflected from an interface having superimposed periodicity. The angle at which the phenomenon occurs was predicted by Bertoni and Tamir [H. L. Bertoni and T. Tamir, Appl. Phys. **2**, 157–172 (1973)] with a classical approach dated before the inhomogeneous waves era. However, during that era it has been shown by many scientists that the ultimate theory to describe critical phenomena, such as the generation of surface waves or the existence of a beam displacement, is the inhomogeneous wave theory. In the present research, one applies the inhomogeneous wave theory and shows that the theory predicts extremely well the phenomenon of backward displacement when a bounded beam reflects from a periodically rough surface. The agreement reaches far beyond a prediction of the angle of occurrence. One is able to simulate the entire physical phenomenon as it happens in the experiments of Breazeale and Torbet. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

9:30

5aPA7. Phase matching methods and their use to approximate resonances from various elastic targets. M. F. Werby (Phys. Dept., Catholic Univ. of America, Washington, DC 20064 and NRL, Stennis Space Center, MS 39529) and H. Überall (Catholic Univ. of America, Washington, DC 20064)

This paper considers acoustic scattering from submersed, elongated elastic objects, i.e., solid spheroids and finite-length cylinders. Surface waves are generated on the objects by incident acoustic waves; they can cause resonances in the objects' vibrations which become evident in the scattered acoustic amplitudes. As it has been shown previously, the resonance frequencies can be obtained by a phase matching condition for multiply circumnavigating surface waves as they propagate on the object over closed paths. Comparing with calculated resonances in the scattering amplitude, it is found that the phase matching condition furnishes correct values for the calculated resonance frequencies for the case of Rayleigh-type surface waves. For spheroids and cylinders of larger aspect ratio, it is shown that bar-wave induced eigenfrequencies may also be used to approximate the calculated resonance frequencies.

9:45–10:00 Break

10:00

5aPA8. Acoustic wave spectroscopy inside a reverberant media with moving scatterers. Julien de Rosny, Philippe Roux, and Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, Universite Paris 7, CNRS UMR 7587, 10 rue Vauquelin, 75251 Paris, France)

In this work, we present an original method in order to measure static and dynamic parameters of moving scatterers confined inside a strong reverberant cavity. In such media, wave propagation is very complex due

to multiple scattering and multiple reflections. In a previous article (de Rosny *et al.*, J. Acoust. Soc. Am. (2001)], we have shown that the number or the elastic cross section of moving scatterers, like a school of fish in an aquarium, are deduced from the pressure field averaged over different and independent positions of the scatterers cloud. In this talk, we present an extension of this work which also takes into account the scatterer's dynamics, e.g., mean speed. Basically, the idea is to relate the correlation of the pressure field to the scatterer's motion. A DAWS-like (Diffusive Acoustic Wave Spectroscopy) theory based on a statistical approach of this system is presented. The theory is compared with experimental data obtained with scatterers whose motion and elastic cross section are known. Finally this method is applied in order to survey the dynamic behavior of a school of fish inside an aquarium under some environmental modifications as luminosity or temperature.

10:15

5aPA9. Bistatic scattering from the edge of a wedge with an impedance coating. R. Hughes, J. Niemiec, Y. J. Stoyanov, and Herbert Überall (Naval Surface Warfare Ctr., Carderock Div., Bethesda, MD 20817-5700)

We calculate the amplitude of acoustic monostatic and bistatic scattering from a solid wedge, thereby extending some previous work of ours that assumes a line source parallel to the edge of the wedge. The surface of the wedge is treated as an impedance surface, and the expression for the impedance corresponding to an absorbing, multilayered coating which had also been obtained by us previously is utilized here for calculating the scattering amplitude from the wedge coated with such an impedance material. Results are obtained as a function of frequency for various incident and bistatic angles. Calculations show in general a highly detailed interference structure except for the case of monostatic backscattering from an impenetrable wedge.

10:30

5aPA10. Diffraction by polyhedral angle: Field in vicinity of singular ray. Vladimir A. Borovikov (Benemerita Universidad Autonoma de Puebla, F.C.F.M., Ciudad Universitario de la BUAP, Colonia San Manuel, Apt. Postal 1152 Puebla, Pue 72570, Mexico)

If the plane acoustic wave hits a polyhedral angle, the scattered field consists of an incident wave, of its reflections from the angle's facets, of primary edge waves excited at angle's edges, and of a spherical wave excited at the angle's vertex. It is supposed for simplicity that the geometry of the problem is such that there is no secondary and subsequent edge waves. There are two types of light–shadow boundaries: first for incident and reflected waves and second for edge waves. As is well known, the field in penumbras is expressed via a Fresnel integral. The incident wave's ray passing through the vertex and its reflections from facets are *singular rays* belonging to the light–shadow boundaries of both types. It will be shown that the field in the vicinity of a singular ray is expressed via a Fresnel intergal and via a generalized Fresnel integral $G(p,q)$ which was introduced in [P. S. Clemmov and T. B. A. Senior, Proc. Cambridge Philos. Soc. **9**, 570–572 (1953); see also V. A. Borovikov, *Uniform Stationary Phase Method*, IEE Electromagnetic Series 40, London, 1994]. This result is valid for an elastic incident wave as well.

10:45

5aPA11. Sound field representation by virtual potential near apex and its application to boundary value problem. Mitsuhiro Ueda, Ryohei Funatsu, and Lawu Tjundewo (Dept. of IDE, Tokyo Inst. of Technol., O-okayama, Meguro-ku, Tokyo 152-8552, Japan, ueda@ide.titech.ac.jp)

A new scheme for representing 2D sound field diffracted by a rigid polygon is proposed where the sound field is described by a sum of two potentials, that is, one that depends on the geometrical arrangement of a point source, an observation point and the polygon and the other that depends on the contribution from the potential at the surface of the scatterer. The latter potential is calculated by assuming virtual potentials in the vicinity of every apex of the polygon and the sound field is calculated using the virtual potential even in the vicinity of the apex using virtual discontinuity principle of diffraction that has been proposed by us before. The virtual potential is fixed so that the calculated potential is continuous across the apex. Consequently two sampling points are necessary in the vicinity of every apex to assure the continuity of the potential across the apex. Thus, the total number of sampling points can be reduced remarkably in this new method. The new method is applied to a semi-infinite plane and a strip and the estimated boundary values are compared with rigorous solutions. Remarkably good agreements are obtained using few sampling points.

11:00

5aPA12. Scatter of ultrasonic waves by rough crack edge. Vladimir Borovikov (Benemerita Universidad Autonoma de Puebla, Facultad de Ciencias Fisiko-Matematicas, Ciudad Universitario, Col. San Manuel, Puebla, PUE, Mexico) and Larissa Ju Fradkin (South Bank Univ., London SE1 0AA, UK)

Our aim in this talk is to describe statistical properties of an ultrasonic field scattered from a rough crack edge assuming that the statistics of the rough crack edge are known. The crack is embedded into a homogeneous solid. In order to develop the necessary algorithms we assume that the crack surface is planar and the crack edge is a polygonal line. The problem is approached in three stages: First, we derive approximate expression for the scattered field in the form of an integral over the crack edge. To this end, we use a version of the Geometrical Theory of Diffraction (GTD) which is called the Integral GTD (IGDT). Second, we obtain approximate analytical expressions for the mean and variance of the diffracted field in terms of statistics of the corner points. Finally, we use the Monte Carlo method for simulating different realizations of the crack edge using a prescribed statistical probability distribution and study statistics of the resulting diffracted field. The model is rather unrealistic but it gives us a rule of thumb for assessing the influence of edge roughness on the pulse echo which we hope to extend later to cases of practical interest.

11:15

5aPA13. Diffraction by systems of finite strips. Yuri Karlovich (Facultad de Ciencias, Universidad Autonoma del Estado de Morelos, Av. Universidad 1001, 62210 Cuernavaca, Morelos, Mexico, karlovich@buzon.uaem.mx)

The talk is devoted to wave diffraction problems for the two-dimensional Helmholtz equation. Boundary-transmission problems for a union of intervals (i.e., strips of a negligible thickness) which lie on a system of one or several parallel lines are considered in a Sobolev space setting. These problems include the mixed Dirichlet/Neumann and impedance boundary and transmission problems. Such problems for a union of intervals are reduced to equivalent systems of Wiener–Hopf equations on Lebesgue spaces over the positive real semiaxis. The matrix symbols of corresponding Wiener–Hopf operators are semialmost periodic, that is, they are continuous on the real line and have, in general, different almost periodic behavior at semineighborhoods of infinity. The solvability of obtained systems of Wiener–Hopf equations is studied on the basis of factorization techniques for oscillating and sectorial matrix functions and a corona problem approach developed in M. A. Bastos, Yu. I. Karlovich, and A. F. dos Santos [Integr. Eq. Oper. Th. **42**, 22–56 (2002)], A. Bottcher, Yu. I. Karlovich, and I. M. Spitkovsky [*Convolution Operators and Factorization of Almost Periodic Matrix Functions* (Birkhauser-Verlag, Basel, 2002)], and L. P. Castro and F.-O. Speck [Integr. Equat. Oper. Th. **37**, 169–207 (2000)].

Session 5aSC

Speech Communication: Speech Production (Poster Session)

Amalia Arvaniti, Cochair
Linguistics Department, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0108

Fred Cummins, Cochair
Department of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland

Contributed Papers

All posters will be on display from 8:30 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:30 a.m.

5aSC1. Exercises in "absurdity:" Interactions between flapping and tongue position in American English with implications for speech motor planning. Bryan Gick (Dept. of Ling., Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada and Haskins Labs., New Haven, CT)

Articulatory conflicts can prevent the achievement of production targets, forcing decisions regarding the resolution of these conflicts. Observing which strategies are chosen can help us to understand relative priorities in speech planning and phonology [Wood, J. Phonetics (1996)]. One particularly felicitous case of such a conflict is the interaction between tongue tip position for English /r/ and vowels and the flapping of /t/ and /d/. This study presents data from ultrasound imaging of tongue kinematics and acoustic output during productions of words containing sequences of vowels, flaps and /r/'s. Results show that for at least some speakers of North American English, a flap may be produced either up (from tip-down to retroflex, as in "Jupiter") or down (from retroflex to tip-down, as in "puberty"). For the one speaker analyzed to date, r-flap-r ("frankfurter") and V-flap-V ("taffeta") begin in the correct position for the leftmost segment in each sequence (retroflex and tip-down, respectively). However, "double flaps" start in tongue down position ("Saturday," "editor"), even when it forces compromise of multiple vowel/liquid tongue positions ("absurdity"). Flap kinematics thus take precedence over tongue position (and therefore acoustic realization) for vowels/liquids. Additional results will be presented. [Research supported by NSERC and NIH.]

5aSC2. Measures of variability in child and adult speech: Test–retest results. Laura L. Koenig (Haskins Labs., 270 Crown St., New Haven, CT 06511 and Long Island Univ.) and Jorge C. Lucero Univ. Brasilia, Brasilia, Brazil)

Elevated measures of speech production variability are commonly observed in children and in clinical populations. Methods of assessing variability thus have utility in understanding processes of speech development and breakdown. Valid interpretation of variability measures for a given speaker requires understanding the stability of such measures over time. Here, we present test–retest data on aerodynamic measures of consonants in two 5-year-old children and two women, each recorded two times approximately 3.5 months apart. Two methods of assessing variability, the spatio-temporal index (STI) and Functional Data Analysis (FDA), were applied to the data. STI yields a single value over repeated productions, incorporating both spatial and temporal aspects of the data. FDA yields separate measures of warping (phasing) and amplitude. Both FDA and STI analyses yielded higher values in children than adults at both times, as expected. The test–retest analysis indicated that both STI and FDA values were significantly correlated over time, suggesting that either method provides a reliable measure of speech production variability. The r values for STI were higher than for FDA, whereas FDA values showed more differentiation across consonants. Our discussion focuses on the relative merits of these two approaches for assessing speech variability.

5aSC3. Intraoral pressure trajectories during voiced and voiceless stops in women and children. Gabrielle Rothman (Haskins Labs., 270 Crown St., New Haven, CT 06511 and Long Island Jewish Medical Ctr.), Laura L. Koenig (Haskins Labs., New Haven, CT 06511 and Jorge C. Lucero (Univ. Brasilia, Brasilia, Brazil)

Vocal-fold vibration requires that subglottal pressure (Psub) exceed intraoral pressure (Pio) values by a threshold amount. The increasing Pio that accompanies stop consonants can thus halt voicing. Muller and Brown, studying Pio trajectories in men, found that most voiced stops were produced with slowly rising (concave) contours, whereas voiceless stops had fast-rising (convex) contours. Modeling work indicated that concave trajectories resulted only when supraglottal volume was increased, and the authors argued that their subjects must have used active cavity expansion maneuvers to slow Pio increase during voiced stops. Here, we extend the measurement techniques of Muller and Brown to Pio data from women and children (5 and 10 years). Although Psub values are comparable between men and women, supraglottal volumes are systematically smaller in women. Children have both higher Psubs and smaller supraglottal cavity sizes than adults. Fully voiced stops are rare in young children, perhaps because children have limited abilities to manipulate oral volumes. Our discussion focuses specifically on the extent to which Muller and Brown's original results are valid for women and child populations, and more generally on how consonantal voicing may be affected by anatomical and physiological differences between speakers as functions of age and sex.

5aSC4. Investigation of labio–lingual coordination in the production of formant frequencies for /u/: Normal speakers and speakers with dysarthria. Gary Weismer, Yana Yunusova, and Jeff Berry (Waisman Ctr. and Dept. of Communicative Disord., Univ. of Wisconsin—Madison, Madison, WI 53705-2280)

Articulatory behavior for the vowel /u/ is interesting because motions of the lips and tongue dorsum can both produce the effect of a relatively low second formant frequency ($F2$). Previous studies have suggested that speakers may trade off labial and lingual contributions to this low $F2$, using less labial contribution when the tongue is making a strong contribution, and *vice versa*. Moreover, if the lips are experimentally con-

strained, the tongue will make up the difference and produce an acoustic output relatively close to the desired value. What is not clear is how the motions of the lips and tongue are coordinated across the vocalic nucleus to produce the *F*-pattern typically associated with /u/. Are the two gestures synchronized? Is there a predictable articulatory organization, one that is fairly consistent across a large group of speakers? We present x-ray microbeam and acoustic data for four groups of speakers (two groups of neurologically normal speakers, and two groups with dysarthria associated with Parkinson's disease and amyotrophic lateral sclerosis) suggesting that the answer to both questions is a qualified yes. Somewhat surprisingly, articulatory coordination in dysarthric speakers does not appear to be qualitatively different from that in normal speakers. [Work supported by NIH.]

5aSC5. Studying articulatory variability using functional data analysis. Jorge C. Lucero (Dept. of Mathematics, Univ. of Brasilia, Brasilia, DF 70910-900, Brazil, lucero@mat.unb.br) and Anders Lofqvist (Haskins Labs., New Haven, CT)

This paper examines articulatory variability across repetitions using functional data analysis. In this analysis, signals are time-normalized using a smooth function of time. The magnitude of the time-warping function provides an index of phasing (temporal) variability; a separate index of amplitude variability is calculated from the time-normalized signal. Here, lip, jaw, and tongue movements in vowel–consonant–vowel sequences recorded using a magnetometer were used for the analysis. Preliminary results suggest articulator specific patterns of variability. Thus vertical jaw movements showed small amplitude variability for alveolar consonants, like /s, t/. As expected, the amplitude variability of the tongue tip was small during alveolar consonants, where the tip is in contact with the alveolar ridge. However, the amplitude of the tongue tip movement varied considerably at the release of alveolar stops. The horizontal and vertical movements of the same articulator often showed a different amount of amplitude variability. For example, during the movement from the first to the second vowel in /api/, the tongue body showed small vertical and large horizontal variability. The results will be discussed in terms of articulatory, acoustic, and perceptual constraints on speech production. [Work supported by CNPq, Brazil, and NIH.]

5aSC6. Effects of upright and supine orientation on tongue position during silence. Maureen Stone, Martin Sutton (Dept. of OCBS and Orthodontics, Univ. of Maryland Dental School, 666 W. Baltimore St., Rm. 5A-12, Baltimore, MD 21201), Vijay Parthasarathy, Jerry Prince (Johns Hopkins Univ., Baltimore, MD 21201), Min Li, and Chandra Kambhamettu (Univ. of Delaware, Newark, DE)

Previous work has examined the effects of supine orientation on tongue position during speech using ultrasound image sequences [Stone *et al.*, J. Acoust. Soc. Am. **111**, 2476(A) (2002)]. That work compares upright and supine ultrasound image sequences during repetition of words. Results indicate backward tongue rotation during supine position, plus several phonemic effects. The present project extends that work by examining silences between the words. The work is of interest because MRI data are collected in supine position. Use of that data to develop models and support theories of speech production calls for an understanding of the differences between supine and upright tongue orientation. Cine-MRI is collected by repeating a short word at timed intervals, such as one second. The word lasts for about half the second followed by silence. During the speech, the effects of gravity on tongue behavior are coupled with the speech demands. During the silence, there are fewer constraints on tongue behavior. The tongue is anterior and reflects neighboring phonemes and small periodic inhalations. Thus, the tongue may resist the gravitational pull to keep the airway patent and facilitate inhalation. This study will compare multiple subjects and words regarding tongue location, mobility, and variability in upright and supine orientation.

5aSC7. Control of closure duration in labial stop consonants. Anders Lofqvist (Haskins Labs., 270 Crown St., New Haven, CT 06511, lofquist@haskins.yale.edu)

This paper examines lip and jaw kinematics in the production of labial stop consonants where the duration of the oral closure is varied for linguistic purposes. The subjects were speakers of Japanese and Swedish, two languages that have a contrast between short and long consonants. Lip and jaw movements were recorded using a magnetometer system. Based on earlier work showing that the lips are moving at a high velocity at the oral closure, it was hypothesized that speakers could control closure duration by varying the position of a virtual target for the lower lip. According to this hypothesis, the peak vertical position of the lower lip during the oral closure should be higher for the long than for the short stops. This would result in the lips staying in contact for a longer period. Results from five subjects suggest that this is the case for those who produce nonoverlapping distributions of closure duration for the two categories. However, the peak velocity of the lower lip raising movement did not always differ between the two categories, suggesting that differences in muscle activation duration rather than virtual target position is another potential control variable. [Work supported by the NIH.]

5aSC8. Phonemic errors are not always categorical: Articulatory evidence. Marianne Pouplier and Louis Goldstein (Haskins Labs., 270 Crown St., New Haven, CT 06511 and Yale Univ., marianne.pouplier@yale.edu)

Recent speech error research has challenged the traditional view that the temporal mislocation of a segment leaves its structural phonological integrity unaffected. Results of several studies suggest the existence of subsegmental, phonologically ill-formed errors (e.g., Mowrey and MacKay, J. Acoust. Soc. Am. **88**, 1299–1312 (1990)]. While Pouplier *et al.* [J. Acoust. Soc. Am. **106**, 2242 (1999)] could demonstrate on the basis of articulatory data that errors involve systematic movements of individual speech gestures, their data are limited to tongue-twister-like stimuli. This leaves open the possibility that the errors they observe might arise at a low execution level given the repetitive nature of the task. The present study employs the SLIP elicitation technique [Motley and Baars, Commun. Quart. **24**, 28–24 (1976)] which is traditionally hypothesized to reflect errors at an abstract phonemic planning level (subjects produce target word pairs in isolation). Results show that also in presumably abstract ''planning errors,'' often two gestures are coproduced (one erroneous, one appropriate), resulting in an illicit phonological structure. These findings call into question the assumption that segments are primary units of speech production. [Work supported by NIH.]

5aSC9. The vocal-tract "potential function:" Many-to-one mappings and compensatory articulations in speech production. Barbara J. Forbes (Dept. of Environment and Mech. Eng., The Open Univ., Walton Hall, Milton Keynes MK7 6AA, UK)

A fully quantitative, mathematical framework for the analysis of compensatory articulations is described. It is shown that the physics of speech production is best described by a spatial ''potential function,'' which, unlike cylindrical *n*-tube approximations, accounts for dispersive wave phenomena in regions of rapidly varying cross section. It is demonstrated that many possible area functions, and hence articulatory configurations, correspond to a single potential function, and so it is claimed that the potential-function descriptor is a both more accurate and more compact basis for the scientific investigation of phonetic systems. This work demonstrates that a 27-vowel system may be simulated from just 8 binary, potential-function parameters. The acoustic correlates of the individual

''bits'' are discussed, and it is found that they characterize the main vowel classes in a transparent way. Transformations from high to mid, and from mid to low, vowels are obtained, along with those involving the tense–lax and round–unround dimensions. It is shown that vocal-tract area functions familiar from the literature can be deterministically recovered from the 8-bit strings, although many others are also possible. It is claimed that the potential-function vocal-tract model has uniquely abstract and compact properties, and is suitable for development as a full phonological representation.

5aSC10. A parametric area function model of three female vocal tracts based on orthogonal modes. Brad Story (Dept. of Speech and Hearing Sci., Univ. of Arizona, P.O. Box 210071, Tucson, AZ 85721-0071)

Image sets corresponding to 11 vowels were obtained from each of 3 female speakers using Magnetic Resonance Imaging. From these images, vocal tract area functions for each vowel were determined. Each speaker's 11 area functions were then decomposed into a set of orthogonal basis functions or ''modes'' and mean vowel area functions. Similar to previous work with male vocal tract data [Story and Titze, J. Phon. (1998)], two orthogonal modes accounted for nearly 90% of the variance within each speaker's vowel set, while the mean area functions serve as neutral, but unique, tract shapes around which the vowels are formed. For each speaker, the two most significant modes and the mean area function were used to develop a parametric model that allowed for calculation of an arbitrary area function based on coefficients of the modes. With these models a hypothetical vowel space was generated in terms of $F1$, $F2$ pairs (first and second formant frequencies). The vowel spaces for each speaker were then used to map time-varying formant frequencies extracted from natural speech to the modal coefficients of their respective parametric models. This effectively produces a time-varying area function suitable for synthesis of speech. [Work supported by NIH R01-DC04789.]

5aSC11. Microanalysis of perturbation-induced temporal changes in repetitive and discrete utterances. Elliot Saltzman (Boston Univ., 635 Commonwealth Ave., Boston, MA 02215 and Haskins Labs., New Haven, CT, esaltz@bu.edu) and Anders Lofqvist (Haskins Labs., New Haven, CT)

In earlier work, we used phase-resetting techniques in which mechanical perturbations are delivered to the articulators during repetitive speech utterances (/. . . pepepe . . . /), and analyzed the resultant temporal changes in order to assess the manner in which the peripheral articulators are coupled to ''clocklike'' processes in the central nervous system. We found that the perturbations induced systematic steady-state shifts in the timing between successive bilabial opening gestures for /p/, demonstrating a resetting of a central clock for these utterances. These results were limited to analyzing durational changes in the opening–closing bilabial ''cycle'' periods (defined between the onsets of successive bilabial opening gestures). However, in work on perturbed discrete utterances, we also found that most of the perturbation-induced durational changes (shortening) occurred during the lip opening phase of the opening–closing bilabial period, while the closing phase was relatively resistant to temporal distortion. We have therefore conducted a more fine-grained set of analyses on the repetitive utterances of our phase-resetting data to localize the parts of the opening–closing gestural cycle in which shortening occurs, and to compare these detailed repetitive kinematics with those obtained in the perturbed discrete utterances. Preliminary results indicate that the repetitive and discrete utterances behave similarly. [Work supported by NIH.]

5aSC12. Efficacy of proprioception in achieving vowel targets when acoustic feedback is blocked: An EMMA study of production under contrasting noise conditions. Mark K. Tiede (Haskins Labs, 270 Crown St., New Haven, CT 06511 and MIT–RLE, Cambridge, MA, tiede@haskins.yale.edu), Melanie Matthies (Hear. Res. Ctr., Boston Univ., Cambridge, MA), Joseph Perkell, Majid Zandipour, and Ellen Stockmann (MIT–RLE, Cambridge, MA)

In a magnetometer study of six normal hearing native English speakers producing vowels in an /hVd/ context, masking noise was used to interfere with acoustic self-monitoring of speech production. Data were obtained under two conditions. In the first, subjects received unmodified audio through earphones. In the second condition audio feedback was replaced by white noise at 95 dB SPL. Under both conditions subjects were instructed to use a real time SPL display for regulating loudness to avoid Lombard effects. Midsagittal transducer data, corrected for head movement, were obtained for the lower incisors, lips, and three points on the tongue. A consistent measurement point was obtained by first delimiting each vowel using RMS thresholding of the acoustics, then selecting the 1/3 offset between those limits. Results showed, contrary to expectation, that neither formants nor the measured articulator locations under the noise condition were more dispersed than when acoustic feedback was available; for some subjects they were in fact less variable. The observed tighter grouping under the noise condition will be discussed in the context of increased stiffness of the production system in response to loss of normal acoustic monitoring. [Work supported by NIH.]

5aSC13. Geometrical analysis of the tongue muscles based on MRI and functional modeling of the tongue. Sayoko Takano (JST-CREST/ATR, Japan), Kiyoshi Honda (ATR, Kyoto, Japan), and Dang Jianwu (JAIST/ATR)

The geometry of the tongue muscles during production of Japanese five vowels was analyzed using MRI toward a realistic design of our physiological articulatory model [Dang *et al.*, Acoust. Sci. Technol. **22**, 415–425 (2001)]. Articulatory changes in shape of the muscle bundles and of the lingual arteries indicated two findings. (1) The styloglossus (SG) was previously thought to cause high-back tongue position by pulling its insertion point in the tongue. However, the MRI-based geometry of the extra-lingual part of SG was found to be almost constant across vowels. This implies that SG may contribute to tongue deformation mainly by the contraction of its intra-lingual part. (2) The right and left arterial landmarks within the tongue tissue indicated tissue deformation across vowels in the lateral direction. The distance between the landmarks was larger in front vowels, and smaller in back vowels. That is, front vowels are associated with lateral expansion, and back vowels are accompanied by medial compression. With the above results, the physiological model of the tongue was revised to have more layers so that the natural geometry of the muscles could be implemented.

5aSC14. Improvements of a physiological articulatory model in construction and control strategy. Jianwu Dang (School of Information Sci., JAIST, Ishikawa, Japan and 923-1292 /ATR, Japan) and Kiyoshi Honda (ATR, Kyoto 619-0288, Japan)

A physiological articulatory model that represents midsagittal regions of the speech organs has been used to simulate articulatory movements during speech. A fast simulation method was realized by modeling both soft tissues and rigid organs by mass-points and springs. It is, however, difficult to exactly represent characteristics of a continuous body using the network consisting of mass-points and volume-less springs. To model soft tissues accurately, the springs are replaced by viscoelastic pillars with both length and thickness, and the masses are distributed over the pillars. Each eight-vertex mesh of the soft tissue is filled with 28 pillars that connect the vertices one to another on the edges and surfaces, and inside diagonally. The initial thickness of the pillars for each mesh is calculated by making the total volume of the pillars equal to the mesh volume. During tongue deformation, pillars change in the length and thickness according to the

physical laws in continuous bodies. The articulator model is driven by muscle contraction parameters. Muscle activation signals are generated based on a target-reaching task, and fed to drive the model to approach the targets. This study further accounts for the agonist–antagonist mechanism of the muscles in the control strategy according to EMG analysis. [Work supported in part by TAO, Japan.]

5aSC15. An articulatory and acoustic study of Pima coronals. Heriberto Avelino and Sahyang Kim (Dept. of Linguist., UCLA, 3125 Campbell Hall, Box 951543, Los Angeles, CA 90095-1543)

The obstruents of Pima, a Uto-Aztecan language spoken in Arizona, have been impressionistically described as involving a contrast between dental and retroflex sounds (Saxton, Saxton, and Enos, 1983; Saxton, 1963). The current study aimed to refine the description of these sounds by investigating palatographic and linguographic data with a concurrent acoustic analysis of recordings from 6 Pima speakers. The results from articulatory data demonstrated that the traditionally so-called retroflex sounds are produced with the apex of the tongue in the alveolar region while the dentals are produced with the blade of the tongue. This evidence suggests that Pima retroflex sounds are not similar to the corresponding sounds in languages such as Hindi or Toda, which are articulated as apical post-alveolars or sub-apical palatal, respectively (Ladefoged and Bhaskararao, 1983; Shalev, Ladefoged, and Bhaskararao, 1994; Ladefoged and Maddieson, 1996). Rather, the Pima retroflexes are more similar to the corresponding sounds of Papago (Dart, 1993). The acoustic analysis showed that the apico-alveolar sounds in contrast to laminal-dentals present a lowering of higher formants (F3, F4), an effect produced by a larger sublingual cavity. The results indicate that in Pima, retroflex sounds are better characterized as apico-alveolar, and the dental as laminal.

5aSC16. Neutralization of word-final voicing in Russian. Miriam Shrager (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405)

This paper reexamines the well-established view in phonology that in Russian there is a voicing neutralization of obstruents in word final position. An experiment was conducted with Russian data that included coronal stops in final position. For each target word four measurements were taken: (1) the duration of the vowel that precedes the target obstruent, (2) the duration of the final obstruent, (3) duration of voicing into the consonant closure, (4) the burst duration of the final stop. The results of Anova statistics test demonstrated that the only variable that had a significant difference between the two groups was the burst duration. However discriminant analysis showed that about 74% of tokens were correctly classified. These results mean that an incomplete neutralization effect exists in Russian in word final position.

5aSC17. Temporal differences in Serbian palatal affricates. Amanda Miller-Ockhuizen and Draga Zec (Dept. of Linguist., Cornell Univ., 203 Morrill Hall, Ithaca, NY 14853)

We report on temporal cues associated with contrastive palatal affricates in Serbian in minimal pairs such as veche evening (apical) and vec'e bigger (laminal). These sounds have been difficult to distinguish using articulatory features. We investigated risetime, rate of rise (ROR), closure duration, release duration and segmental duration in four vowel contexts and two word positions. The apical affricate displays shorter risetime, ROR and closure duration, with longer release duration, while the laminal affricate displays longer risetime, ROR and closure duration, with shorter release duration. While risetime differentiates fricatives from affricates (Howell and Rosen, 1983), and ROR differentiates stops from fricatives (Weigelt *et al.*, 1990), this is the first study to demonstrate risetime and ROR distinguishing two affricates. Rate of rise distinguishes these two affricates from the palatal fricative and alveolar stop allowing for the maintenance of contrast over the three-way distinction. Risetime correlates with closure proportional providing evidence for a universal physical relationship between closure duration and risetime. Differences in closure duration are compensated for by differences in release duration, allowing the overall segmental duration of the two classes of affricates to be similar.

5aSC18. Frequency and sentence position in the production of word-final /d/ in North-Central Peninsular Spanish. Carolina Gonzalez (Dept. of Linguist., USC, GFS 301, 3601 Watt Way, Los Angeles, CA 90089-1693)

In Spanish, /b, d, g/ are usually pronounced as voiced approximants. In North-Central Peninsular Spanish, these speech sounds are frequently fricated and devoiced. Previous research on this dialect using nonce words showed that word-final /b, d, g/ are fricated more often in stressed syllables than in unstressed syllables [Gonzalez (2001, 2002)]. This study investigates whether other factors are responsible for frication and/or voicing of these sounds. Specifically, the role of frequency and sentence position is examined in the pronunciation of word-final /d/ in real words. Eight speakers of this dialect read a word list with 15 common words like verdad (truth) and 15 rare words like efod (ephod), both in sentence medial and sentence final position. The word list included eight tokens for each word and position, and it was randomized for each speaker and repetition. The effects of frequency and word position will be discussed, and the results will be contrasted with previous research. [Work supported by Del Amo and the Basque Government.] (To be presented in Spanish.)

5aSC19. Coarticulatory fundamental frequency effects associated with Ju/hoansi guttural consonants. Amanda Miller-Ockhuizen (Dept. of Linguist., Cornell Univ., 203 Morrill Hall, Ithaca, NY 14853)

This study presents the results of an investigation of the fundamental frequency of vowels following guttural clicks in Ju/hoansi. Ju/hoansi contains an extremely rich set of consonantal phonation types that behave as gutturals, having contrastive aspirated, uvularized, epiglottalized, and glottalized clicks. Additionally, closure voicing is contrastive for oral aspirated, nasal aspirated, uvularized, and epiglottalized consonants, making the language a good test case for the effect of voicing on fundamental frequency of the following vowel. Fundamental frequency traces were created from the inverse of glottal period durations, which were calculated between adjacent Epoch marks (point of maximum glottal closure) determined by the ESPS Epochs utility associated with XWAVES. This methodology avoided pitch halving and doubling that is often found following nonmodal consonant phonation types. Previous studies have shown lower $F0$ following voiced unaspirated and voiced aspirated consonants, which have led to the expectation that voicing is associated with lowering of fundamental frequency. The results here show that voicing only lowers fundamental frequency when the release phase of a consonant is voiced as is the case for the voiced aspirated clicks, and not when there is only closure voicing, as is found in the case of the voiced epiglottalized clicks.

5aSC20. Assessing the time course of learning new acoustic–motor relationships using a frequency-altered feedback paradigm. Jeffery A. Jones (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, jones@atr.co.jp)

Speakers rely on a learned mapping between their speech articulations and the resulting auditory feedback. In a previous study, subjects exposed to frequency-altered feedback adapted to the novel articulatory–acoustic relationship [J. A. Jones and K. G. Munhall, J. Acoust. Soc. Am. **108**, 1246–1251 (2000)]. However, obtaining a precise measurement of the size of the effect was difficult because the speakers' ''habitual vocal pitch,'' evaluated in a control condition, was used as a reference. Although an individual's habitual pitch is reliable over long periods, there may be

considerable short-term variability. In the present study, this noise was significantly reduced by asking speakers to imitate the pitch of a synthetic vowel. This paradigm allowed more precise measurements of adaptation as well as tests of generalization. In a control condition, subjects were able to reliably produce the frequency of the vowel. However, when speakers heard their vocal pitch shifted either up or down during the experimental trials, they relied on the altered auditory feedback to guide their productions. Short-term exposure to the altered feedback generated measurable aftereffects. The data are discussed in terms of a model proposed for fundamental frequency control.

FRIDAY MORNING, 6 DECEMBER 2002 CORAL GARDEN 2 AND 3, 8:00 TO 11:30 A.M.

Session 5aSP

Signal Processing in Acoustics: Underwater Signal Processing

Charles F. Gaumond, Chair
Acoustics Division, Naval Research Laboratory, Code 7142, 4555 Overlook Avenue, SW, Washington, D.C. 20375-5320

Contributed Papers

8:00

5aSP1. A new multichannel spatial diversity technique for long-range acoustic communications in shallow water. Pierre-Philippe J. Beaujean and Emmanuel P. Bernault (Dept. of Ocean Eng., Florida Atlantic Univ., 101 N. Beach Rd., Dania Beach, FL 33004)

A new multichannel spatial diversity technique has been developed for underwater acoustic communications in very shallow waters. This technique combines a novel synchronization method with maximum-likelihood symbol estimation. It was tested with a multichannel Mills-Cross receiver using various numbers of elements. The FAU general purpose acoustic modem source transmitted messages using four types of frequency-hopped multiple-frequency-chirp-keying (FH-MFSK) modulation: 4 hops at 221 coded bits per second (cps), 2 hops at 442 cps, or no hopping at 1172 cps. These types of modulation allowed for robust data transmission in adverse environment. The modem operated between 16 and 32 kHz at 192-dB source level, at ranges from 1 to 5 km in 40 ft of water. Using only four channels of the Mills-Cross receiver array, messages coded at 1172 cps were received with a frame error rate (FER) of 0% at a range of 3 km. In the same four-channel configuration, messages coded at 221 cps were received with no frame error at 5 km. The experimental results matched the performance predictions from the Crepeaus model. This reliable and computation-efficient method can be implemented on new generations of embedded acoustic modems, such as the four-channel FAU acoustic modem, and can provide significant improvements in communication performance.

8:15

5aSP2. Wideband link-budget analysis for undersea acoustic signaling. Joseph A. Rice and Joseph T. Hansen (Phys. Dept., Naval Postgrad. School, Monterey, CA 93943)

Link-budget analysis is commonly applied to satellite and wireless communications for estimating the signal-to-noise ratio (SNR) at the receiver. Link-budget analysis considers transmitter power, transmitter antenna gain, channel losses, channel noise, and receiver antenna gain. For underwater signaling, the terms of the sonar equation readily translate to a formulation of the link budget. However, the strong frequency dependence of underwater acoustic propagation requires special consideration, and is represented as an intermediate result called the channel SNR. The channel SNR includes ambient-noise and transmission-loss components. Several acoustic communication and navigation problems are addressed through wideband link-budget analyses. [Work sponsored by ONR 321.]

8:30

5aSP3. Further investigations into performance metrics for underwater communications. Scott L. Whitney, Geoffrey S. Edelson, Ned B. Thammakhoune, and Michael S. Richman (BAE Systems, M/S MER15-2651, P.O. Box 868, Nashua, NH 03061-0868)

Determining the relationship between the performance of an underwater acoustic data communications system and the operating environmental conditions is a problem that continues to plague researchers. The complexity of the time-varying channel is difficult to measure and model. Therefore an approach that uses metrics measured from data collected at sea to characterize the channel is attractive. As expected, preliminary assessments on limited data have shown that performance depends not only on environmental conditions, but also on system implementation. By extracting a variety of metrics, a better understanding of the subset that discriminate between good and bad performance can be developed. Also by analyzing the relationship between certain metrics and performance, system limitations can be identified for re-evaluation. For example, a surprising result of the initial assessment of performance using a multichannel decision feedback equalizer on real data showed that sparseness of multipath arrivals may be an arbiter of performance [Richman *et al.*, J. Acoust. Soc. Am. **110**, 2619 (2001)]. Therefore changes to the algorithm that allows for sparse arrivals may improve performance. In this paper, a larger number of metrics from greater quantities of real-data and system configurations are measured and evaluated against equalizer results.

8:45

5aSP4. Coherent equalization for underwater acoustic communications: A performance evaluation of adaptive MMSE and time-reversal techniques. James C. Preisig (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, jpreisig@whoi.edu)

A simple analytical expression is presented for evaluating and comparing the performance of channel-estimate-based coherent equalizers. A channel-estimate-based equalizer is one whose filter coefficients are calculated as a function of an explicit estimate of the acoustic channel impulse response. This is in contrast to a direct adaptation equalizer whose coefficients are calculated directly as the solution to a data estimation error minimization problem. The equalizer performance expression is different from other commonly used expressions because it is formulated from the viewpoint of the estimate of the channel impulse response rather than the actual channel impulse response. The expression allows the dependence of the performance of coherent equalizers on channel complexity and the quality of the channel estimate to be separately evaluated. The expression is used to evaluate and compare the performance of adaptive MMSE and time-reversal equalizers. [Work supported by ONR Ocean Acoustics.]

9:00

5aSP5. Application of matching pursuit algorithm for dispersive acoustic signals. Chuen-Song Chen, James H. Miller, G. Faye Boudreaux-Bartels, Gopu R. Potty, and Colin J. Lazauski (Univ. of Rhode Island, Narragansett, RI 02882, ahsong@ele.uri.edu)

Dispersive acoustic signals contain significant amounts of information concerning the shallow water environment structure. It is possible to unravel the modal propagation dispersion characteristics using a proper time-frequency analysis. Recent experiments such as ShelfBreak Primer and ASIAEX collected long range propagation data in shallow water on vertical line arrays from SUS charge explosive sources. Signals collected by the vertical arrays were processed using Morlet wavelets to observe the modal arrival structure due to group speed dispersion. Although the wavelet transform can extract the time-frequency information, it is not sufficient to get high resolution for high frequencies. The matching pursuit algorithm is an alternative approach that provides excellent spectral localization. We present an improved processing algorithm based upon the matching pursuit algorithm to analyze the data from ASIAEX-2001. This analysis yields the dispersion characteristics of the acoustic signals and clearly shows the modal arrivals, especially for higher modes. These modal dispersion diagrams from the results form the basis for long range sediment tomography. [Work supported by ONR.]

9:15

5aSP6. Scale adaptive channel estimation. Paul J. Gendron and T. C. Yang (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375)

Channel (linear operator) estimation requires either *a-priori* knowledge or joint estimation of channel temporal coherence. This information serves to condition the estimation process reducing estimator variance. Kalman-like conditional expectation algorithms have exhibited success by modeling the channel response via time evolution either assuming coherence time *a priori* or jointly estimating it. With coherence time viewed as a ''time resolution'' property of the channel response rather than a ''time evolution'' property a new computational structure results. From this viewpoint we introduce a new channel estimation algorithm ''in-scale'' vice ''in-time.'' In such a framework detail is added, or synonymously coherence time is successively reduced, to meet a maximum *a posteriori* (MAP) criteria. The algorithm starts with a ''mean'' (time invariant) estimate and gradually adds detail evolving in scale until the MAP is attained. Computationally the algorithm exploits conjugate gradient directions recursing from low resolution to higher resolution until the requisite prior complexity (estimated via an empirical Bayes method) and data fidelity are met. Scale projections are computed efficiently and dimension reduction is aided via a wavelet decomposition. Explanations in terms of the minimum description length principle are also provided.

9:30

5aSP7. Extracting the Green's function between receivers using underwater acoustic noise. Philippe Roux, Steve Lynch, W. A. Kuperman, and NPAL group[a] (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238)

Recent experimental and theoretical works in ultrasonics show that the Green's function between transducers fastened to an aluminum sample can be measured from the correlation of thermal noise [R. L. Weaver and O. J. Lobkis, ''Ultrasonics without a source. Thermal fluctuation correlations at MHz frequencies,'' Phys. Rev. Lett. **87**, 134301 (2001)]. Similar results have been obtained in geophysics using seismic noise data [A. Paul and M. Campillo, ''Extracting the Green's function between two stations from coda waves,'' Trans. Am. Geophys. Union **82–47**, F842 (2001)]. Sources of noise in underwater acoustics range from ship noise at low frequency to surface noise and even thermal noise at very high frequencies. We theoretically demonstrate that at least an approximate Green's function can be obtained from surface noise. This result is confirmed by noise data recorded on arrays of receivers during the NPAL98 experiment. [Work supported by ONR.] [a] The NPAL group is composed of 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.

9:45

5aSP8. Robustness of null broadening against source motion. Heechun Song, William Kuperman, William Hodgkiss, Peter Gerstoft (MPL/SIO, 9500 Gilman Dr., La Jolla, CA 92093-0238), and Jaesoo Kim (Korea Maritime Univ., Pusan 606-791, South Korea)

Adaptive array beamforming achieves high resolution and sidelobe suppression by producing sharp nulls in the adaptive beampattern when using sample matrix inversion (SMI). Large aperture sonar arrays with many elements have small resolution cells and interferers may move through many resolution cells in the time required for accumulating a full-rank sample covariance matrix. In addition, longer duration snapshots are required due to the longer transit time across the array. This leads to ''snapshot deficient'' processing. In this paper, the null broadening technique originally developed for an ideal stationary problem is extended to the snapshot deficient problem arising from nonstationarity of the background interference. Null broadening allows the strong interferers to move through resolution cells and increases the number of degrees of freedom, thereby improving the detection of weak stationary signals. Numerical simulations demonstrate the robustness of the null broadening approach.

10:00–10:15 Break

10:15

5aSP9. Experimental results for a prototype 3-D acoustic imaging system using an ultra-sparse planar array. John M. Impagliazzo (Naval Underwater Warfare Ctr., Newport, RI 02841), Alice M. Chiang, and Steven R. Broadstone (Teratech Corp., Burlington, MA 01803 and Naval Underwater Warfare Ctr., Newport, RI 02841)

A handheld high resolution sonar has been under development to provide Navy Divers with a 3-D acoustic imaging system for mine reconnaissance. An ultra-sparse planar array, consisting of 121 1 mm×1 mm, 2 MHz elements, was fabricated to provide 3-D acoustic images. The array was 10 cm×10 cm. A full array at this frequency with elements at half-wavelength spacing would consist of 16 384 elements. The first phase of testing of the planar array was completed in September 2001 with the characterization of the array in the NUWC Acoustic Test Facility (ATF). The center frequency was 2 MHz with a 667 kHz bandwidth. A system-level technology demonstration will be conducted in July 2002 with a real-time beamformer and near real-time 3-D imaging software. The demonstration phase consists of imaging simple targets at a range of 3 m in the ATF. Experimental results obtained will be reported on. [Work supported by the Defense Applied Research Project Agency, Advance Technology Office, Dr. Theo Kooij, Program Manager.]

10:30

5aSP10. Detection of a weak target in the presence of strong interferers. E. J. Sullivan (OASIS, Portsmouth, RI 02971-1032), J. V. Candy (Lawrence Livermore Natl. Lab., Livermore, CA 94550), and W. M. Carey (Boston Univ., Boston, MA 02215)

The detection of a weak target in a hostile environment consisting of ambient noise, distant shipping, and ''known'' interferers, such as passing ships, is treated by applying a parametrically adaptive signal-processing technique. In this approach, the oceanic environment and noise sources are represented by mathematical models and directly incorporated into the processor, thereby allowing their effects to be removed from the decision function prior to target detection. It is then shown that the weak target dominates the resulting decision function. The approach recognizes the variability of the ocean environment and allows the processor to adaptively adjust its inherent parameters and underlying statistics to the changing ocean environment.

10:45

5aSP11. Detector analysis for shallow water active sonar. Thomas J. Pastore and Michael E. Phillips (SPAWAR Systems Ctr.–San Diego, 53560 Hull St., San Diego, CA 92152)

SPAWAR Systems Center–San Diego, in concert with the Office of Naval Research (ONR) and Defense Advanced Research Projects Agency (DARPA) has designed and built a proof-of-concept broadband biomimetic sonar. This proof-of-concept sonar emulates a dolphin biosonar system; emitted broadband signals approximate the frequency and time domain characteristics of signals produced by echolocating dolphins, the receive system is spatially modeled after the binaural geometry of the dolphin, and signal processing algorithms incorporate sequential integration of aspect varying returns. As with any sonar, object detection in shallow water while maintaining an acceptable false alarm rate is an important problem. A comprehensive parametric analysis of detection algorithms is presented, focusing primarily on two detector strategies: a matched filter and a spectral detector. The spectral detector compares the ratio of in-band to out-of-band power, and thus functions something like a phase-incoherent matched filter. This computationally efficient detector is shown to perform well with high proportional bandwidth signals. The detector (either matched filter or spectral) is coupled with an alpha–beta tracker which maintains a running noise estimate and calculates signal excess above the estimated noise level which is compared to a fixed threshold.

11:00

5aSP12. A theoretical analysis of the AWSUM fluctuation-based processors. Kenneth E. Gilbert and Carrick L. Talmadge (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Coliseum Dr., University, MS 38677, kgilbert@olemiss.edu)

Nonlinear signal processing schemes developed originally by R. A. Wagstaff for ocean acoustics and named ''AWSUM,'' have been applied in this laboratory to atmospheric acoustic signals in the presence of wind noise. It has been found that the AWSUM processors can provide dramatic gains in the signal-to-noise ratio for narrow-band acoustic signals strongly degraded by wind noise. This paper presents a theoretical analysis of the AWSUM processors for narrow-band acoustic signals in the presence of Raleigh and non-Raleigh (intermittent) noise. The theory quantitatively explains the observed systematic behaviors of the AWSUM processors. In particular, it is shown that in the limit of large sample number, the signal-to-noise gain for AWSUM processors approaches the ratio N(0)/S(0), where N and S are the probability density distributions for the acoustic pressure squared, for noise alone and for noise-plus-signal, respectively. For a steady narrow-band signal in Raleigh noise, the ratio N(0)/S(0) is given by exp(SNR), where SNR is the usual linear signal-to-noise ratio (e.g., SNR=1 corresponds to 0 dB). For wind noise, which is intermittent noise, the ratio N(0)/P(0) is greater than exp(SNR), so that greater gain is predicted. The predictions are in accord with both simulated data and field data.

11:15

5aSP13. Use of amplitude without phase angle to achieve pseudo-coherent processing gain. Ronald A. Wagstaff (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Oxford, MS 38677)

Processing gain as much as 3 dB per average doubling can be achieved by coherent temporal averaging (i.e., with phase), compared to 1.5 dB per doubling for incoherent averaging (i.e., without phase). However, it is not uncommon for signals to be coherently cancelled as if they never existed, or be severely attenuated. For this reason, coherent temporal averaging is seldom used. A method is presented that utilizes amplitudes to generate ''pseudo-phase angles'' (not real angles). Pseudo-phase angles are combined with the corresponding amplitudes to form pseudo-vectors for vector and coherent averaging. Thus, the large gains that are normally associated only with coherent processing can be achieved without using the real phase angles. This ''pseudo-coherent'' gain is about the same as coherent averaging gain, but signals are not cancelled. Furthermore, the pseudo-phase angles can be used to automatically detect signals. Results from measured data will be presented to illustrate pseudo-coherent gains in signal-to-noise ratio of 10 dB, minimum detectable level of 3 dB, resolution of 2×, and automatic signal detection. [Work supported by U.S. Army Space and Missile Defence Command.]

FRIDAY MORNING, 6 DECEMBER 2002 GRAND CORAL 1, 8:00 A.M. TO 12:00 NOON

Session 5aUW

Underwater Acoustics: General Topics in Underwater Acoustics II

Irene Veljkovic, Cochair
Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Richard S. Keiffer, Cochair
Naval Research Laboratory, Code 7181, Stennis Space Center, Mississippi 39529-5004

Contributed Papers

8:00

5aUW1. Estimation of spatial coherence in shallow water waveguides. Ilya Rozenfeld (1847 W. Main Rd., Portsmouth, RI 02871, Ilya_Rozenfeld@Raytheon.com), William M. Carey (Boston Univ., Boston, MA 02215), Peter G. Cable (BBN Technologies, Arlington, VA 22209), and William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

For various underwater acoustic applications such as beamforming and matched field processing, it is important to understand the spatial characteristics of shallow water acoustic fields. Oceanic and geoacoustic variabilities introduce random fluctuations into the fields, for which an important metric is the spatial coherence that is the correlation function of the field. The theory developed here is based on a perturbation approach to the modal solution of the Helmholtz equation and treats scattering from sound speed fluctuations as well as from rough interfaces. The results here extend previous work to media where fluctuations can have arbitrary correlation functions and are not necessarily horizontally isotropic or homogeneous. This allows for modeling of coherence degradation due to physically realistic mechanisms. The behavior of the coherence function with respect to various parameters is illustrated using a Pekeris waveguide with superimposed randomness. The effects of several different scattering mechanisms on the coherence will be compared and rank ordered for a range of frequencies. [Work partially supported by ONR.]

8:15

5aUW2. Arrival time determined by vertical array: Experimental results in shallow water. Gee-Pinn J. Too, Chih-Chung K. Wang (Dept. of Naval Architecture and Marine Eng., Natl. Cheng Kung Univ., Tainan, Taiwan 70101, ROC), Chifang Chen (Natl. Taiwan Univ., Taipei, Taiwan, ROC), Jim Lynch, and Arthur Newhall (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The arrival time determined by vertical array is a useful parameter to apply matched field processing (MFP) methods for source localization. In this paper, the procedure including wavelet transform for extracting the signal, adaptive filter for retrieving the signal, and wigner-vill distribution for determining the arrival time between sensors is presented. During the South China Sea ASIAEX Experiment, the 103 m deep source was towed by a ship. The data receiving system was a satellite–buoy underwater measurement system with a 16-element vertical line array. The experiment showed that the data was collected with very good quality. Results of this procedure are more efficient, and faster than earlier results based on power spectrum for MFP.

8:30

5aUW3. Toward a practical cross section for rough surface scattering. Shira L. Broschat (School of Elec. Eng. & Computer Sci., Washington State Univ., P.O. Box 642752, Pullman, WA 99164-2752)

The lowest-order small slope approximation (SSA) provides a practical means of obtaining the scattering strength for rough surfaces. While results are accurate over a wide range of scattering angles, as surface slopes increase and for forward scattering beyond the specular angle, even higher-order SSA results become inaccurate. This is due, in part, to the lack of nonlocal interactions in the SSA. Voronovich introduced the nonlocal small slope approximation (NLSSA) as a generalization of the SSA to explicitly include nonlocal interactions. It has been shown that the NLSSA is more accurate than the SSA both for larger surface slopes and at low forward grazing angles. However, the computational cost precludes use of the NLSSA in practice. The NLSSA expression for the scattering strength can be separated into two terms. The first term is identical to the lowest-order SSA and, hence, the second term of the NLSSA can be considered a correction to the lowest-order SSA. In this paper, we make an approximation to the second term at forward scattering angles that reduces it to a simple algebraic expression. We present numerical results and discuss the region of validity for the reduced form of the equation. [Work supported by ONR.]

8:45

5aUW4. Experimental multipath time spread in a bathymetrically complex coastal environment. Frederick R. Di Napoli, Jason S. Shelton, and Ilya Rozenfeld (Raytheon MS C12, 1847 W. Main Rd., Portsmouth, RI 02871)

An acoustic survey was conducted that covered both the continental shelf and slope regions in water depths from 40 m to greater than 1000 m. The types of bathymetric conditions encountered included those likely to be found in any coastal environment. At the benign end of the spectrum a relatively smooth bottom with small slopes connected the shelf and slope. In another area the shelf was relatively short and extended to water depths of approximately 100 m, followed by very steep continental slopes bordered by canyon walls. Intermediate conditions consisted of a short shelf and large continental slopes without canyon walls. Broadband (12 Hz to 4 kHz) acoustic measurements were made under strongly downward-refracting sound-speed conditions using impulsive sources at two different depths. The source at 18 m was in a deep (approximately 40 m) weak surface duct. A strong sound-speed gradient separated this source from the 86-m source depth. The impulsive data have been processed to extract a measure of the effective multipath time spread. Results will be presented showing the sensitivity to source depth, frequency range from the receiver, and bathymetric conditions.

9:00

5aUW5. Coherent backscatter simulation for synthetic aperture sonar. Samantha Dugelay, J. Richard Brothers, and A. Sean Chapman (QinetiQ, Newton's Rd., Weymouth DT4 8UR, UK, sdugelay@qinetiq.com)

COSIM is a coherent backscatter simulator for generic high frequency sonars developed by QinetiQ. This model simulates acoustic backscattering in the time domain from various types of seafloor and targets, in various environmental conditions. The parameters of the model include pulse length and type, sonar characteristics, environmental properties concerning water column and sea surface and physical seafloor properties. In this paper we present descriptions of the various simulation stages, from the seabed generation to the signal simulation. Our main focus in thi paper is the introduction of spatial and temporal fluctuations within the water column and on the seabed and their detrimental effect on synthetic aperture sonar processing. Prior work on seabed clutter (Dugelay and Brothers, *Proceedings Oceans 2002*) and statistical analysis (Dugelay, Brothers and Heald, 142nd ASA Proceedings) is introduced into the seabed simulation module and initial results on the effect of seabed clutter density are presented. More particularly, the introduction of water column fluctuations is studied: the simulation of realistic fluctuations will be described and the effects on the backscattered signals will be analyzed. Finally, synthetic aperture sonar imaging will be shown in a variety of environmental cases and the image degradation will be compared to the environmental fluctuations.

9:15

5aUW6. Models of acoustic scattering from continuous and discrete heterogeneity in marine sediments. Anatoliy N. Ivakin (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, ivakin@apl.washington.edu)

A description of the scattering processes in marine sediments for different types of scatterers, continuous and discrete, is given in terms of the scattering cross section per unit sediment volume and the scattering cross section per unit seabed area. However, there is a difference in the way these volume and surface cross sections are expressed through the characteristics of the scatterers. In the case of a continuous medium, they are expressed in terms of the spatial spectra of inhomogeneity (fluctuations of acoustical parameters) and the interface roughness. In the case of discrete inclusions, the cross sections are described in terms of the individual scattering functions of discrete targets and statistical distributions of their parameters (size, shape, orientation, material, etc). Some simplified expressions for the individual scattering functions of such bodies as shell fragments, rocks, and others are discussed. Examples for the size distribution functions are presented obtained from analysis of published data. Two types of models for discrete scattering are considered. In the first one, the discrete inclusions are uniformly distributed in the sediment volume. In the second model, the scatterers are placed on or near the water-sediment interface or buried at a certain depth below (a top or buried monolayer of scatterers). Note that, rigorously, the individual scattering functions should take into account the influence of the water-sediment interface. Effects which result from this influence are discussed. The frequency-angular dependencies of the seabed scattering strength for the two types of models of discrete scattering are calculated and compared to those for a corresponding continuous seabed medium with volume heterogeneity or interface roughness. [Work supported by ONR, Ocean Acoustics.]

9:30

5aUW7. On the role of the frozen surface approximation in small wave-height perturbation theory for moving surfaces. Richard Keiffer (Naval Res. Lab., Code 7181, Stennis Space Center, MS 39529-5004), Jorge Novarini (Planning Systems, Inc., Long Beach, MS 39560-9702), and Robert Scharstein (Univ. of Alabama, Tuscaloosa, AL 35487-0286)

In the standard development of the small wave-height approximation (SWHA) perturbation theory for scattering from moving rough surfaces [e.g., E. Y. Harper and F. M. Labianca, J. Acoust. Soc. Am. **58**, 349–364

(1975)] the necessity for any sort of frozen surface approximation is avoided by the replacement of the rough boundary by a flat (and static) boundary. In this paper, this seemingly fortuitous byproduct of the small wave-height approximation is examined and found to fail to fully agree with an analysis based on the kinematics of the problem. Specifically, the first-order correction term from standard perturbation approach predicts a scattered amplitude that depends on the source frequency, whereas the kinematics of the problem point to a scattered amplitude that depends on the scattered frequency. It is shown that a perturbation approach in which an explicit frozen surface approximation is made before the SWHA is invoked predicts (first-order) scattered amplitudes that are in agreement with the kinematic analysis. [Work supported by ONR/NRL (PE 61153N-32) and by grants of computer time DoD HPC Shared Resource Center at Stennis Space Center, MS.]

9:45

5aUW8. Impulse response of a two-dimensional rough surface overlying an inhomogeneous, nondispersive medium: A hybrid model. Richard Keiffer (Naval Res. Lab., Code 7181, Stennis Space Center, MS 39529-5004), Jorge Novarini (Planning Systems, Inc., Long Beach, MS 39560-9702), and Guy Norton (Naval Res. Lab., Code 7181, Stennis Space Center, MS 39529-5004)

A numerical model to calculate the impulse response of a two-dimensional, impenetrable, rough surface directly in the time domain has been recently introduced [R. S. Keiffer and J. C. Novarini, J. Acoust. Soc. Am. **107**, 27–39 (2000)]. This model is based on wedge diffraction theory and assumes the half-space containing the source and receiver is homogeneous. In this work, the model is extended to handle media where the index of refraction varies with depth by merging the scattering model with a ray-based propagation model. The resulting hybrid model is tested against a finite-difference time-domain (FDTD) method for backscattering from a corrugated surface in the presence of a refractive layer. This new model can be applied, for example, to calculate acoustic reverberation from the sea surface in cases where the water mass is inhomogeneous and dispersion is negligible. [Work supported by ONR/NRL (PE 61153N-32) and by grants of computer time DoD HPC Shared Resource Center at Stennis Space Center, MS.]

10:00–10:15 Break

10:15

5aUW9. Sub-critical insonification of buried elastic shells. Irena Veljkovic and Henrik Schmidt (MIT, 77 Massachusetts Ave., 5-204, Cambridge, MA 02139)

In a shallow water environment a high frequency high grazing angle mine-hunting sonar approach is vastly limited by the coverage rate, making the detection and classification of buried objects using subcritical grazing incidence an attractive alternative. One of the central issues in mine countermeasurements regarding the physics of scattering from spherical shells is the isolation and the analysis of the resonant excitations of the system distinguishing the manmade elastic targets from rocks or other clutter. Burial of an elastic target in the seabed results in a variety of modifications to the scattered response caused by different physical mechanisms, geometric constrains, and intrinsic sediment properties. The aim of this research is to identify, analyze, and explain the fundamental effects of the sediment and the proximity of the seabed interface on the scattering of sound from elastic spherical shells insonified using low frequencies at subcritical incident angles. A new, comprehensive understanding of the GOATS98 experimental data was obtained distinguishing the effects of the acoustics environment from the resonant signature of a buried elastic target. To achieve this and to further investigate the more intricate details of the scattering process, a numerically improved, OASES-3D modeling framework was used. [Work supported by ONR.]

10:30

5aUW10. Temporal coherence and time spread of ocean surface scattered high frequency acoustic signals. Steven Lutz, David Bradley, and Jon Reeves (Appl. Res. Lab. and Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804)

The influence of the ocean environment on the temporal coherence and time spread of high frequency, surface scattered and fully refracted acoustic signals over intervals ranging from a few seconds to tens of hours have been investigated experimentally. Acoustic data were collected during the broadband coherence experiment conducted in July 2002 off the coast of San Diego in a water depth of 1000 m. Signals transmitted include short pure tone pulses and broadband (up to 22 kHz) linear frequency modulated (LFM) sweeps centered at 20, 40, and 60 kHz. The projectors were suspended at two depths from a moored buoy at two different ranges. The receiver hydrophones were suspended from a subsurface float at depths from 40 m to 300 m. Environmental measurements included a towed CTD string, water temperature at a fixed point, near surface bubble layer using a multibeam sonar, current speed and direction, directional wave height, and wind speed. [Work supported by ONR Code 321US under the ARL Program.]

10:45

5aUW11. Temporal coherence of acoustic signals propagating through the near surface bubble layer. R. Lee Culver, David Bradley, and Jon Reeves (Appl. Res. Lab. and Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804)

The influence of the near surface bubble layer on the temporal coherence of high frequency acoustic signals over intervals ranging from a few seconds to tens of hours have been investigated experimentally. Acoustic data were collected during the broadband coherence experiment conducted in July 2002 off the coast of San Diego in a water depth of 1000 m. Signals transmitted include short pure tone pulses and broadband (up to 22 kHz) linear frequency modulated (LFM) sweeps centered at 20, 40, and 60 kHz. The projector and receiver hydrophone were suspended from FLIP. Environmental measurements included a towed CTD string, water temperature at a fixed point, near surface bubble layer using a multibeam sonar, current speed and direction, directional wave height and wind speed. [Work supported by ONR Code 321US under the ARL Program.]

11:00

5aUW12. Sound speed and attenuation in unconsolidated sand with viscous pore fluids. Brian T. Hefner and Kevin L. Williams (Appl. Phys. Lab., Univ. of Washington, 1013 N.E. 40th St., Seattle, WA 98105)

During the Sediment Acoustics Experiment 1999 (SAX99) several researchers measured sound speed and attenuation in the sediment over a broad range of frequencies (125 Hz to 400 kHz). While the variation of sound speed with frequency was well modeled by Biot theory, the variation of attenuation with frequency deviates from Biot theory at high frequencies. In an effort to understand this deviation, attenuation and sound speed measurements were made in unconsolidated glass beads saturated with various high-viscosity fluids using the ''attenuation array,'' which was used during SAX99. According to Biot theory, changing the viscosity of the pore fluid alters the frequency range where the most significant dispersion occurs. Thus by increasing the viscosity of the fluid, it was possible to observe significant dispersion in the frequency range of the attenuation array (80 to 300 kHz). The results of these measurements were compared to both Biot theory and Buckingham's grain-to-grain shearing model. [Work supported by ONR.]

11:15

5aUW13. Investigation of the relative significance of individual environmental parameters to sonar performance prediction uncertainty. Ding Wang, Wen Xu, and Henrik Schmidt (Ocean Eng. Dept., Acoust. Group, MIT, Cambridge, MA 02139)

A large part of sonar performance prediction uncertainty is associated with the uncertain ocean acoustic environment. Optimal *in situ* measurement strategy, i.e. adaptively capturing the most critical uncertain environmental parameters within operational constrains can minimize the sonar performance prediction uncertainty. Understanding the relative significance of individual environmental parameters to sonar performance prediction uncertainty is fundamental to the heuristics to determine the most critical environmental parameters. Based on this understanding, the optimal parametrization of ocean acoustic environments can be defined, which will significantly simplify the adaptively sampling pattern. As an example, the matched-field processing is used to localize an unknown sound source position in a realistic ocean environment. Typical shallow water environmental models are used with some of the properties being stochastic variables. The ratio of the main lobe peak to the maximum side lobe peak of the ambiguity function and the main lobe peak displacement due to mismatch are chosen as performance metrics, respectively, in two different scenarios. The relative significance of some environmental parameters such as sediment thickness, weights of empirical orthogonal functions (EOFs) has been computed. Some preliminary results are discussed.

11:30

5aUW14. Directional characteristics of conformal transducers on fluid loaded elastic shells. Peter R. Stepanishen (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02892) and Don Cox (Naval Undersea Warfare Ctr. Div. Newport, RI 02841)

A general analytic approach is presented to evaluate the directional characteristics of conformal velocity sources representing transducers on fluid loaded elastic shells. The sources/transducers are assumed to be structurally decoupled from the shell but are coupled to the shell via the fluid. The pressure field is expressed via superposition as the sum of the pressure field from the sources on a rigid shell plus the pressure field from the elastic shell assuming zero normal velocity for the sources. An *in vacuo* eigenvector expansion is used to describe the velocity field of the shell. The modal coefficients of the expansion for the shell are related shells via a set of linear algebraic equations to the modal coefficients of a different expansion for the surface pressure due to the sources on the rigid shell. Shadowing effects of the sources on the response of the shell are included in the formulation. Numerical results for the directional characteristics of specified circular velocity sources on rigid and elastic spherical shells are presented as a special case. The results illustrate the effects of diffraction and shell resonances on the directional characteristics of the sources in the low, mid, and high frequency regions. [Work supported by ONR.]

11:45

5aUW15. Monostatic and bistatic scattering from a density contrast surface. C. Feuillade (Naval Res. Lab., Stennis Space Center, MS 39529-5004) and C. S. Clay (Univ. of Wisconsin-Madison, Madison, WI 53706)

Chu's formalism for the impulse response of a point source to a density contrast isovelocity wedge [J. Acoust. Soc. Am. **86**, 1883–1896 (1989)] is of significant practical importance, since it facilitates the extension of wedge diffraction boundary scattering models to rough penetrable surfaces. The method is used here to study scattering from an experimental surface described by Li *et al.* [J. Acoust. Soc. Am. **96**, 3715–3720 (1994)]. The geometry contains both interior and exterior wedges, and allows the diffracted scattering component to be studied independently of the reflected component. For density contrast values which emulate rigid boundary conditions, the formalism correctly models the experimental results, and duplicates predictions of the Biot–Tolstoy solution for a rigid wedge [J. Acoust. Soc. Am. **29**, 381–391 (1957)]. The formalism is then used to investigate an assumption commonly made in boundary scattering work, i.e., that scattering levels can be effectively estimated by scaling the predictions of a rigid surface scattering model by the reflectivity of the surface. The results suggest that the assymption is most valid for monostatic (backscattering) configurations, and less good for bistatic cases, with increasing differences occurring between interior and exterior wedge geometries. [Work supported by ONR.]

FRIDAY AFTERNOON, 6 DECEMBER 2002 CORAL GALLERY 1, 1:00 TO 5:05 P.M.

Session 5pAAa

Architectural Acoustics, Noise and Structural Acoustics and Vibration: Duct Acoustics and Vibration

Jorge P. Arenas, Cochair
Institute of Acoustics, Universidad Austral de Chile, P.O. Box 567, Valdivia, Chile

Jerry G. Lilly, Cochair
JGL Acoustics, Inc., 5266 N. W. Village Park Drive, Issaquah, Washington 98027

Chair's Introduction—1:00

Invited Papers

1:05

5pAAa1. Low-frequency models to compare acoustical performance of ducts of gradual area change. Jorge P. Arenas and Victor Poblete (Inst. of Acoust., Univ. Austral de Chile, P.O. Box 567, Valdivia, Chile, jparenas@uach.cl)

When two ducts of different circular cross-sectional area are joined together to form an acoustical transmission system, an additional impedance is introduced. This particular problem is found commonly in air conditioning systems among others. Simple models are used to solve the sound wave equation, and they assume that the plane wave mode is responsible for the most part of the transmission/insertion loss calculation. From previous works, it appears that, although higher-order modes can propagate at certain

frequencies, their transmission efficiency is low and they do not cause a critical deviation from the assumption of one-dimensional propagation. The models are based on numerical/semi-analytical and asymptotic approximations to transform the sound pressure wave equation into an acoustic impedance equation with proper boundary conditions. The final aim is to calculate the transfer impedance of the duct section which can be used to calculate the sound field radiated. The model should be extended to ducts with incompressible mean flow for low flow Mach numbers, and further work is currently under progress on this subject. However, for low mean flow, the effect on the transmission loss has been found to be only marginal for some particular geometries. [Work supported by Fondecyt, Grant No. 1020196.]

1:25

5pAAa2. Acoustical performance testing of duct silencers. Jerry G. Lilly (JGL Acoustics, Inc., 5266 NW Village Park Dr., Issaquah, WA 98027, Jerrylilly@aol.com)

The test method used in North America to evaluate the acoustical performance of duct silencers is ASTM E477. The ASTM standard provides an approved method for measuring the aerodynamic pressure drop, dynamic insertion loss, and self-generated noise of duct silencers. Unfortunately, restrictions on the construction of the test duct in the current standard cause erroneous results in the measurement of insertion loss at low frequencies. These errors are due to acoustic resonances that occur naturally in the empty test duct. It is recommended that the standard be modified to require anechoic terminations at both ends of the test duct. It is also recommended that more specific design criteria for the source chamber be established, and that a new section be added to provide for the measurement of the radiated noise from the casing of the test specimen. Specific recommendations relating to all of these issues will be provided.

1:45

5pAAa3. The prediction of noise generated by the interaction of airflow and duct discontinuities. David J. Oldham (School of Architecture and Bldg. Eng., Univ. of Liverpool, Liverpool L69 3BX, UK)

There is a paucity of data available relating to the noise generated by ventilation system elements. This paucity arises from the difficulty in obtaining such data as it requires the use of rare and expensive combined acoustic and aerodynamic measurement facilities. In this paper a prediction technique is proposed which overcomes the need for these expensive test facilities. Discontinuities in ducts result both in the generation of flow noise and a loss of static pressure. The greater the discontinuity the greater is the loss in static pressure and the greater is the sound power generated. Over a number of years there has been considerable research aimed at establishing the correlation between the drop in static pressure across a flow spoiler and the noise generated. In this paper it is shown that for a typical duct discontinuity it is possible to predict the sound power generated from knowledge of its pressure loss characteristics. A prediction method is proposed based upon a normalized sound-power level which is a function of the Strouhal number. Such a prediction technique has enormous potential value in the design of high-velocity air distribution systems.

2:05

5pAAa4. The location and sizing of holes and obstructions in pipes and ducts. David J. Oldham and Max H. F. de Salis (School of Architecture and Bldg. Eng., Univ. of Liverpool, Liverpool L69 3BX, UK)

A change in the cross-sectional area of a duct or pipe due to the presence of an obstruction will modify the eigenfrequencies of that system. A technique is presented for determining the blockage area function of a pipe or duct based upon the concept of these eigenfrequency shifts but using resonance and antiresonance frequencies determined from a single measurement of sound pressure within a duct. The technique is shown to be capable of revealing the location and size of single and multiple obstructions to a high degree of accuracy. The technique has also been successfully applied to the detection, location and sizing of small holes in duct walls. The position of a hole in the duct wall is revealed as the beginning of an apparent gradual expansion in the reconstructed blockage area function emanating from an initial positive dc shift. Once the hole has been located, an impedance model of the duct incorporating the size and location of the hole and the measured wave number at the first-order duct resonance allows the determination of the area of the side hole from a simple quadratic equation.

2:25

5pAAa5. Sound power level measurements of the noise produced by a residential split unit air conditioner. Malcolm J. Crocker (Mech. Eng. Dept., Auburn Univ., Auburn, AL 36849, mcrocker@eng.auburn.edu) and Jorge P. Arenas (Univ. Austral de Chile, Valdivia, Chile)

The noise produced by an air conditioner may be caused by several mechanical and aerodynamic sources. This abstract presents an experimental attempt to characterize and compare the noise sources of a residential split unit air conditioner. In the split unit studied in this research, the compressor and condenser and associated fans are normally located outdoors. Therefore, they were removed from the unit and did not form part of the studies. Care was taken with the unit to separate the inlet and exhaust noise from the noise radiated from the cabinet. The measurements were made with a two-microphone sound intensity probe and these resulted in sound power level data. The sound power levels produced by radiation from the inlet, exhaust, and cabinet were obtained for five different volume flow rates. The effect on the sound power generated by removing the coil and adding a damper in the exhaust was investigated. Measurements and subjective studies show that the low-frequency sound is predominantly radiated from the exhaust and inlet. At high frequency, the cabinet noise dominates. The measuring procedure and data obtained are relevant to industry involved in development and manufacture. [Work supported by both Trane Corporation and Bruel & Kjaer.]

2:45

5pAAa6. Estimation of increased attenuation of turbofan inlet noise with optimized and active liners. William P. Patrick and Larry W. Hardin (United Technologies Res. Ctr., 411 Silver Ln., East Hartford, CT 06108, patricwp@utrc.utc.com)

A modified version of the Eversman finite element code is used to estimate the maximum attenuation benefit obtainable in a typical aircraft nacelle lined with a perfect absorber, having the optimum impedance at each frequency and for each mode. The radiated acoustic wave pattern is computed from the fan face to the far field. Calculations are presented for the first 4 radial modes of the first cut-on circumferential mode at 2 BPF (blade passing frequency) for an Advanced Ducted Propeller model. For each mode the program performed a parametric variation of the nacelle wall liner impedance to determine the optimum impedance which minimized the peak sound pressure level at the far-field boundary. Results for an inlet having a hard wall, an optimal passive liner, a uniform (single element) active liner having an optimized impedance, a liner having linearly varying impedance from fan face to the inlet, and a fully optimized (30 liner segments) active liner are presented. Attenuation of the optimized liner having uniform impedance compared to the hard-wall case is 11.2 dB, 4.9 dB, 4.2 dB, and 7.4 dB for radial modes 0, 1, 2, and 3, respectively. Linearly varying the impedance increased the attenuation slightly. The ability of the liner to significantly alter the pressure pattern within the nacelle is shown quite clearly in the visualized pressure field.

3:05–3:20 Break

Contributed Papers

3:20

5pAAa7. Ducts with spatially varying properties. José Luis Barros (Instituto de Acustica, Univ. Austral de Chile, Casilla 567, Valdivia, Chile, jbarros@uach.cl)

A duct with a varying cross-section and/or acoustic impedance of its internal surface is represented by a series of small duct sections. Each section has parallel walls and a locally reacting internal surface with constant impedance. The effect of the transition between two ducts of a different cross-section is studied. The ducts are considered bidimensional and the characteristic equation is solved numerically, in this way a set of orthogonal modes are obtained with which the sound field can be synthesized. The amplitudes corresponding to each mode can be obtained from the boundary conditions for pressure and velocity between duct sections. Considering a number N of modes, a system of linear equations with $2N$ unknowns is obtained (corresponding to the modal amplitudes of the incident and reflected waves). The acoustic behavior of duct transitions is analyzed considering the three fundamental types of impedance (mass character, stiffness character, and damping). (To be presented in Spanish.)

3:35

5pAAa8. Modeling the aeroacoustics of axial fans from CFD calculations. Alexandre Salesky, Vincent Hennemand, Smaine Kouidri (Ecole Nationale d'Arts et Metiers, ENSAM–LEMFI, 75013 Paris, France), and Yves Berthelot (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

The main source of aeroacoustic noise in axial fans is the distribution of the fluctuating, unsteady, aerodynamic forces on the blades. Numerical simulations were carried out with the CFD code (NUMECA), first with steady flow conditions to validate the aerolic performances (pressure drop as a function of flow rate) of the simulated six-bladed axial fans. Simulations were then made with unsteady flows to compute the fluctuating force distributions on the blades. The turbulence was modeled either with the Baldwin–Lomax model or with the K-epsilon model (extended wall function). The numerical results were satisfactory both in terms of numerical convergence and in terms of the physical characteristic of the forces acting on the blades. The numerical results were then coupled into an in-house aeroacoustics code that computes the farfield radiated noise spectrum and directivity, based on the Ffowcs-Williams Hawkings formulation, or alternatively, on the simpler Lowson model. Results compared favorably with data obtained under nonanechoic conditions, based upon ISO 5801 and ISO 5136 standards.

3:50

5pAAa9. Simulation of acoustic wave behavior in ducts and a plenum using finite elements package ANSYS. Jose de Jesus Villalobos-Luna, Pedro Lopez-Cruz, Ricardo Ramirez-Valencia, and Fernando J. Elizondo-Garza (Acoust. Lab., FIME, Univ. A. de Nuevo Leon, P.O. Box 28 F, Cd. Univ., San Nicolas, 66450, N.L., Mexico, villalobosluna@hotmail.com)

This paper shows and discusses the use of the finite element software ANSYS for simulation of the behavior of sound waves in a ducts-plenum system. The modeling is intended for didactic purposes, beginning with simple three-dimensional geometries of ducts and plenum. Two conditions can be included: (1) with and without acoustical absorbent materials on the inside walls of the ducts and (2) including or not labyrinths. The analysis allows students to see the behavior of the waves inside the duct-plenum system and they can manipulate the parameters of the model to improve the attenuation characteristics of the system.

4:05

5pAAa10. Silencing using flexible plate in a duct. Sripriya Ramamoorthy, Karl Grosh, and Tony G. Nawar (Dept. of Mech. Eng., Univ. of Michigan, Ann Arbor, MI 48105)

A flexible plate interacting with air in a duct can provide passive means for low frequency broadband transmission loss. The sensitivity of the system to various parameters including unintentionally applied tension, effect of external fluid loading, lateral plate cross modes, structural and acoustic boundary condition are analyzed through experimental measurements and theoretical predictions. In order to avoid breakout noise, a backing cavity can be introduced below the plate. This introduces differences in filtering characteristics. Compared to plate in a single duct, the two-duct system will have higher plate resonance frequencies due to cavity loading on the plate. Means to achieve low frequency broadband transmission loss using two-duct silencers will be discussed. Significance of three dimensionality of the problem will be brought out by comparing the results of three-dimensional finite-element analysis with experimental data. Successful designs and experiments for low frequency attenuation will be presented. [Work supported by NSF and ONR.]

4:20

5pAAa11. Active noise control technique for diesel train locomotor exhaust noise abatement. Franco Cotana and Federico Rossi (Dipartimento di Ingegneria Industriale, Univ. of Perugia, Via G. Duranti 1/A-4, 06125 Perugia, Italy, frossi@unipg.it)

An original prototype for train locomotor exhaust gas pipe noise reduction (electronic muffler) is proposed: the system is based on an active noise control technique. An acoustical measurement campaign has shown that locomotor exhaust noise is characterized by very low frequency components (less than 80 Hz) and very high acoustic power (up to 110 dB). A

peculiar electronic muffler characterized by high acoustical efficiency at very low frequencies has been designed and realized at Perugia University Acoustic Laboratory; it has been installed on an Italian D.245 train locomotor, equipped with a 500-kW diesel engine. The electronic muffler has been added to the traditional passive muffler. Very low transmission losses are introduced by the electronic muffler because of its particular shape; thus, engine efficiency does not further decrease. Canceling noise is generated by means of DSP-based numerical algorithm. Disturbing noise and canceling noise destructively interfere at the exhaust duct outlet section; outgoing noise is thus reduced. The control system reduces exhaust noise both in the steady and unsteady engine regime. Measurement results have shown that electronic muffler introduces up to 15 dB noise abatement in the low-frequency components.

4:35

5pAAa12. Active control of automotive fan noise. Anthony Gerard, Alain Berry, and Patrice Masson (GAUS Mech. Eng. Dept., Univ. of Sherbrooke, QC J1J 2R1, Canada, anthony.gerard@gaus.gme.usherb.ca)

Active control for globally reducing the noise radiated by automotive axial engine cooling fans is investigated. First, an aeroacoutic model of the fan is combined with acoustic directivity measurements to derive a distribution of equivalent dipole sources on the fan surface. The results reveal that the fan behaves like a distributed dipole at blade passage tones when the upstream flow through the fan is spatially nonuniform. Numerical simulations of active noise control in the free field have been carried out using the previous aeroacoustic model of the fan and a dipole secondary source in front of the fan. The numerical results show that a single dipole control source is effective in globally controlling the sound radiation of the fan at the blade passage frequency and its first harmonic. Last, an experimental investigation of active control is presented. It consists of a SISO feedforward configuration with either a LMS algorithm (for FIR filters) or a back-retropopagation algorithm (for neural networks) using the Simulink/Dspace environment for real-time implementation.

4:50

5pAAa13. Developing active noise control systems for noise attenuation in ducts. Rosely V. Campos, Rodrigo C. Ivo (Dept. of Electron. and Telecomunication Eng., PUC-Minas, Av. Dom Jose Gaspar 500, 30535-610, Belo Horizonte, MG, Brazil), and Eduardo B. Medeiros (UFMG, 31270-901, Belo Horizonte, MG, Brazil)

The present work describes some of the research effort on Active Noise Control (ANC) being jointly developed by the Catholic University of Minas Gerais (PUC-MINAS) and the Federal University of Minas Gerais (UFMG). Considerations about the implementation of Digital Signal Processing for noise control in ducts has been presented. The objective is to establish a study on Active Noise Control in ducts combining geometry and acoustic parameters modification together with adaptive digital filtering implementation. Both algorithm and digital signal processing details are also discussed. The main results for a typical application where real attenuation has been obtained are presented and considered according to their use in developing real applications. The authors also believe that the present text should provide an interesting overview for both designers and students concerned about Active Noise Control in ducts. (To be presented in Portuguese.)

FRIDAY AFTERNOON, 6 DECEMBER 2002 GRAND CORAL 2, 1:15 TO 4:35 P.M.

Session 5pAAb

Architectural Acoustics and Noise: Acoustic Comfort in Architecture

Elvira Viverios, Cochair
Department Arquititura (ARQ)-CTC, Federal University of Santa Catarina, Campa Universitario Trindade, Florianopolis SC 88040 900, Brazil

Fausto Rodriguez, Cochair
Condor 260, Casa 8, Col. Las Aguilas 01710, D. F. Mexico

Chair's Introduction—1:15

Invited Papers

1:20

5pAAb1. Parametric analysis of open plan offices. Flavia F. Nogueira (Dept. of Civil Eng., Federal Univ. of Santa Catarina, CxP 476, Florianopolis, SC, 88040-900, Brazil) and Elvira B. Viveiros (Federal Univ. of Santa Catarina, Florianopolis, SC, 88040-900, Brazil)

The workspace has been undergoing many changes. Open plan offices are being favored instead of ones of traditional design. In such offices, workstations are separated by partial height barriers, which allow a certain degree of visual privacy and some sound insulation. The challenge in these offices is to provide acoustic privacy for the workstations. Computer simulation was used as a tool for this investigation. Two simple models were generated and their results compared to experimental data measured in two real offices. After validating the approach, models with increasing complexity were generated. Lastly, an ideal office with 64 workstations was created and a parametric survey performed. Nine design parameters were taken as variables and the results are discussed in terms of sound pressure level, in octave bands, and intelligibility index.

1:50

5pAAb2. Acheiving speech intelligibility at Paddington Station, London, UK. Helen M. Goddard (AMS Acoust. Ltd., Rayleigh House, 21 Queen Anne's Pl., Bush Hill Park, London EN1 2QB, UK, info@amsacoustics.co.uk)

Paddington Station in London, UK is a large rail terminus for long distance electric and diesel powered trains. This magnificent train shed has four arched spans and is one of the remaining structural testaments to the architect Brunel. Given the current British and European legislative requirements for intelligible speech in public buildings AMS Acoustics were engaged to design an electroacoustic solution. In this paper we will outline how the significant problems of lively natural acoustics, the high operational noise levels and the strict aesthetic constraints were addressed. The resultant design is radical, using the most recent dsp controlled line array loudspeakers. In the paper we detail the acoustic modeling undertaken to predict both even direct sound pressure level coverage and STI. Further it presents the speech intelligibility measured upon handover of the new system. The design has proved to be successful and given the nature of the space, outstanding speech intelligibility is achieved.

2:20

5pAAb3. Acoustic confort at home: Noise emitted by house installations. Recommendations in order to avoid such noise. Santiago Jimenez (Lab. of Acoust. Eng. Mech., Tech. Univ. of Catalonia, C/Colom, 11 08222, Terrassa, Spain, santiago.jimenez@upc.es)

The present survey consists of the analysis and the study of the solutions used at present in the installations of water supply and elevators. It has been carried out from the acoustic point of view. In order to achieve a thorough study a pilot plant was built in the Laboratory of Acoustics of the School of Industrial Engineering of Terrassa. This pilot plant reproduced different kinds of installations of the water supply in houses. And it has allowed us to systematize the measures and also to determine the optimum solutions from the acoustic perspective. In accordance with the objectives and the process of the survey, the solutions regularly employed in the facilities of water supply and elevators in houses have been analyzed, and levels of noise associated to these facilities have been also presented. A summary of the results obtained in the plant has been included, according to diverse variables. Both the conclusions of the analysis of the data obtained in the laboratory and those of the installations of the houses have been also compared, which has allowed us to describe a series of suggestions with the purpose of reducing the acoustic emission of this type of installations, and increase the acoustic comfort at home. (To be presented in Spanish.)

2:50

5pAAb4. Acoustic insulation of recent flat buildings in Pamplona, Spain. Miguel A. Arana, Ricardo San Martin, Antonio Vela, and Maria L. San Martin (Public Univ. of Navarra, 31006 Pamplona, Spain, marana@unavarra.es)

The Municipal Government of Pamplona and the Housing Department of the Government of Navarra (Spain) decided to evaluate acoustic insulation of new flat buildings in Pamplona. Such study was ordered to the Acoustic Laboratory of the Public University of Navarra. During the last year, measurements relating to acoustic insulation were carried out in most of the new blocks of flats made in Pamplona. The results obtained show that acoustic insulation is insufficient in most buildings both for airborne sound reduction and for impact sound. The figures required by Spanish legislation do not fulfil in more than 80 percent of the cases. On the other hand, a great correlation between sound reduction figures obtained from index of ISO 140-4 and from R (in dBA) of the Spanish norm for *in situ* measurements was carried out. A new Spanish legislation will be approved shortly.

3:20

5pAAb5. The need for standards and codes to ensure an acoustically comfortable environment in multifamily housing buildings in Mexico City. Eduardo Kotasek Gonzalez and Fausto Rodriguez Manzo (Div. de Ciencias y Artes para el Diseno, Universidad Autonoma Metropolitana, Azcapotzalco, Mexico, D.F., Mexico, rfme@correo.azc.uam.mx)

It is clear that almost all kinds of buildings require protection against noise and undesirable sounds, however, there are some countries where this is not yet regulated, such is the case of Mexico. Mexico City, the biggest city in the world could also be the noisiest. This is a problem which is today being debated; in fact there is no doubt that this has an important influence on the acoustic comfort conditions of dwellings, besides the habits and culture of the inhabitants, which are very different from those in the Anglo-Saxon countries. These are all details that must be taken into account in the design of an acoustic comfort standard for buildings in cities like Mexico. In this paper we deal with this problem and it suggests some recommendations to consider in a proposed acoustic comfort standard or code to be applied in the design of multifamily housing buildings in Mexico City.

3:50

5pAAb6. The architecture of shopping centers: An acoustical perspective. Marisa S. Kusakawa (Dept. of Civil Eng., Federal Univ. of Santa Catarina, CxP 476, Florianopolis, SC, 88040-900, Brazil) and Elvira B. Viveiros (Federal Univ. of Santa Catarina, Florianopolis, SC, 88040-900, Brazil)

This paper analyzes the current state-of-the-art of the architecture of shopping centers from an acoustical point of view. There has been a worldwide spread of such buildings, whose characteristics do not differ much across cultural or climatic conditions. Despite the success achieved as a commercial product for big and medium cities, the acoustical comfort of employees and consumers has not been of concern. The food court and recreational areas tend to be reverberant spaces and usually very noisy. The research aims to correlate modern architectural solutions and typical sound fields generated in these buildings, taking a Brazilian case study. Also intended is to establish guidelines for architects and planners in order to provide acoustical comfort.

4:05

5pAAb7. Criteria for acoustic discomfort from major internal noises in residential and office spaces. Jin Yong Jeon (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea)

Although acoustic discomfort is difficult to qualify (in association with the acoustical characteristics of building spaces), there are some cases where noise limits related to activity disturbance can be clearly defined. In the experiment, the limits of perceptual loudness of noise were investigated in designed rooms similar to residential and office environment. As a result of the floor impact noise evaluations, the limit of lightweight impact noise was around 60 dBA, whereas the limit of heavyweight impact noise was lower. It was revealed that the subjects perceived the loudness of noise from a vacuum cleaner as not irritating. Noises implicated with humans, e.g., walking or talking sounds, are the most annoying in office environment. This paper also presents some definable parameters in the psychoacoustical sense such as Zwicker parameters and ACF/IACF factors along with subjective evaluation.

4:20

5pAAb8. Classroom acoustics in Omaha, Nebraska: Measurements and outreach. Erica E. Bowden, Lily M. Wang, and David T. Bradley (Arch. Eng. Prog., Univ. of Nebraska-Lincoln, Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0681, lwang@unl.edu)

This project collected data detailing the current status of classroom acoustics while simultaneously increasing awareness of educational acoustics and exposure to acoustical engineering research in local schools. In 8 Omaha-area schools, 2 acoustical aspects which directly influence speech intelligibility were measured and studied: (1) background noise levels resulting from a building's mechanical systems; and (2) reverberation times within the classroom. Prior to each measurement session, a presentation on good classroom acoustic principles was presented to the teacher and students in that classroom. Subsequently, the teacher and students assisted the principal investigator and research assistants in the measurement of the acoustic data from their classroom. In total, 13 classrooms were studied, with the assistance of 425 middle and high school students in 14 different classes. The background noise results have been associated with commonly used noise criterion curves. Results are compared to recommended background noise levels and reverberation times suggested by ANSI standard S12.60-2002. The resulting database provides knowledge on the current status of classrooms in Omaha, and may be indicative of classroom situations across the midwestern United States.

FRIDAY AFTERNOON, 6 DECEMBER 2002 CORAL KINGDOM 2 AND 3, 1:00 TO 3:45 P.M.

Session 5pAB

Animal Bioacoustics: Bioacoustics as a Conservation Tool II

Renata de Sousa Lima, Chair
Lab de Mamiferos Aquaticos LMA, INPA C P 478, Manaus AM CEP 69083 000, Brazil

Contributed Papers

1:00

5pAB1. Ship strikes and whales: Shadows, mirrors and paradoxes. Edmund R. Gerstein, Joseph E. Blue (Leviathan Legacy, Inc., 1318 SW 14th St., Boca Raton, FL 33486, Gerstein2@aol.com), and Steven E. Forsythe (Naval Undersea Warfare Ctr., Newport, RI 02841)

Ship collisions with whales and other marine animals occur near the surface. Here Snells law applies and can play a crucial role in the survival and behavioral ecology of marine mammals. In this study we investigated some theoretical constructs of near surface sound propagation that can render whales vulnerable to collisions. Empirical measurements support predictions and demonstrate that propeller noise directly ahead of many large vessels can become indistinguishable from the ambient noise. Large acoustic shadows in front of approaching ships and the confluence of near surface acoustic propagation effects seriously challenge the whales' ability to detect and locate approaching ships. At the surface, ship noise profiles are loudest at the stern and off the port and starboard sides while remaining remarkably quiet at the bow. Whales with calves swimming near the surface alerted to ship sounds may seek refuge by actively swimming into the quieter shadow zones directly ahead of vessels. Once here, hydrodynamic forces could overcome calves and adults. While this behavioral scenario has not been investigated, acoustic measurements reveal a disturbing paradox, suggesting the circumstance is plausible. Further investigations may result in directional acoustic solutions that light-up these acoustic shadows and can alert whales in time to avoid injury.

1:15

5pAB2. Impact of boats on the vocal behavior of humpback whales off Brazil. Renata S. Sousa-Lima (Projeto Baleia Jubarte (PBJ), Praia do Kitongo s/no, Caravelas, Bahia, Brazil 45900-000, pboi@inpa.gov.br), Maria E. Morete (PBJ/USP, Brazil), Roberto C. Fortes, Ana C. Freitas, and Marcia H. Engel (PBJ, Brazil)

Experiments were performed to evaluate the impact of boat approaches on the vocal behavior of humpback whales that breed around the Abrolhos archipelago off the Brazilian coast. Vocalizing whales were located by monitoring the amplitude of each individual's sounds. Silent approaches to focal singers were performed using a small zodiac with an improvised sail. Songs were recorded with one hydrophone connected to a portable DAT recorder before and during the approach of a vessel. Spec-

trograms were correlated to each whale's activity and position with respect to the vessel. Differences in song variables of two individuals were tested when the motor was ON and OFF. Both whales sang shorter versions of their songs when exposed to engine noise. No alteration in mean phrase duration was detected, but the number of phrases in each theme decreased. Three individuals interrupted their songs after a motor boat switched gears within 300 m, and resumed singing when the engine was returned to neutral. The limited number of observables calls for further investigation to evaluate the consequences of boat impact. More reliable information will allow regulation to achieve sustainable tourism activity in the area. [Work supported by FBPN, MacArthur Foundation, Society for Marine Mammalogy, Instituto Baleia Jubarte.]

1:30

5pAB3. Effect of boat traffic on the vocal and surfacing behavior of river dolphins: A key for abundance estimation methods. Maria Claudia Diazgranados, Paola Meja, and Adriana Acosta (Omacha Foundation, Carrera 5 #26a-47 (506) Bogot, Colombia, omacha@cable.net.co)

To evaluate the effect of boat traffic on vocal and surfacing behavior of river dolphins, *Inia geoffrensis* and *Sotalia fluviatilis*, field recordings and scanners were carried out in the Amazon River during 2001. Changes on surface dolphin activities and vocalization rates were analyzed. Twenty complete slot interferences between boats and dolphin vocal behavior were selected from 2400 min of total recording. Significant differences between signal types (squeals for *Inia* and whistles for *Sotalia*) between pre-exposure, exposure, and post-exposure to the interference were found, for both *Inia* and *Sotalia* species. Six hundred scanners were conducted from a 6-m-high base observation platform. Interferences with pre-exposure, exposure, and post-exposure times were selected with only one boat crossing the dolphin area. Significant differences in frequency, type of event, and immediate response were observed for both *Inia* and *Sotalia* species, especially when a speedboat was crossing the area at high speed. This information shows that boat traffic did affect vocal and surface behavior of river dolphin species inhabiting the Amazon River in Colombia, and that abundance estimations using boat transect along the river may change depending upon the type and speed of the vessel used.

1:45

5pAB4. Effects of boat noise on vocalizations of gray whales (*Eschrichtius robustus*) in Bahia Magdalena, Baja California Sur, Mexico. Francisco Ollervides (RARE Ctr. for Tropical Conservation, 1840 Wilson Blvd., Ste. 402, Arlington, VA 22201)

Enhanced propagation characteristic of sound underwater has led to development of excellent acoustic sensory systems among cetaceans, which use sound for individual recognition, predator detection, communication, facilitation of social cohesion, and coordination during courtship and foraging behaviors. Increased input of noise from human activities in the oceans has sparked public and scientific concern about potential effects on these animals. In Bahia Magdalena, anthropogenic noise originates mostly from boat traffic. Boat noise is similar at frequencies below 2 kHz to gray whale vocalizations; thus, it has the potential to interfere with these animals' communication. Significant differences in ambient noise levels (chi squared=44.1, $df=3$, $p=0.000\,01$) were observed depending on boat traffic, biological and nonbiological contributors, and water depth at various regions throughout Bahia Magdalena. This study contributes to baseline acoustic profiles in representative habitats where whale watching occurs for future management considerations. It was evaluated whether boat noise affected duration, bandwidth, emphasized frequency, high and low frequencies, and repetition rate of the 11 different gray whale signal types by measuring and contrasting such characteristics in the presence and absence of boat noise. Monitoring these changes in vocal behavior will aid in determining if these animals are being disturbed by noise. (To be presented in Spanish.)

2:00

5pAB5. Impact of anthropogenic sounds on fishes. Arthur N. Popper (Dept. of Biol. and Neurosci. and Cognit. Sci. Prog., Univ. of Maryland, College Park, MD 20742, apopper@umd.edu), Robert D. McCauley, and Jane Fewtrell (Curtin Univ., Perth 6845, Western Australia)

It is widely suspected that anthropogenic sound could impact the life and survival of marine mammals. One of the major sources of anthropogenic sounds in many parts of the world are seismic air guns used in geologic exploration. Yet there are few data on the physiological and behavioral impacts of air-guns on fishes or marine mammals. This study tested the effects of exposure to air guns on the ears of marine fishes. Pink snapper (*Chrysophrys auratus*) held in cages were exposed to signals from an air gun towed toward and away from the cages, mimicking the stimulus from an approaching/passing seismic vessel. Results from scanning electron microscopic examination of the ears showed extensive damage occurring weeks after exposure. Damage included deep holes in the saccular epithelium in places where sensory hair cells would be located, and/or ''blebbing'' on the surface of the epithelium suggesting dying sensory cells. While it is recognized that fishes outside of a cage might swim away from a loud sound, the results reported here suggest that very intense sounds could have a significant impact on the auditory system of fishes, and perhaps other marine organisms.

2:15–2:30 Break

2:30

5pAB6. Fish, mollusks and other sea animals' use of sound, and the impact of anthropogenic noise in the marine acoustic environment. Michael Stocker (Michael Stocker Assoc., P.O. Box 559, Lagunitas, CA 94938)

Many marine animals use sound and acoustic energy sensors to adapt to their environment. Most biological studies closely examine a particular species' relationship to a specific stimulus. This report examines the fields of research on marine biological adaptations to sound since 1950, assembling an overview of the biological importance of sound in the ocean. It also examines the various sources of anthropogenic noise in the sea with a focus on the potential impacts of that noise on the marine acoustic environment. [Work sponsored by Earth Island Institute.]

2:45

5pAB7. Assessment of training noise impacts on the Red-cockaded woodpecker. Larry Pater and David Delaney (U.S. Army Engineer Res. and Development Ctr., 2902 Farber Dr., Champaign, IL 61821)

The Endangered Species Act requires that federal agencies carry out programs to conserve Threatened and Endangered Species (TES) and to evaluate the impacts of agency activities on listed species. Few scientific studies have investigated the effects of military training activity, including noise, on TES. Regulatory agencies, when faced with a lack of definitive impact information, tend to be conservative in their findings. This can lead to land use restrictions that can affect the military's ability to train to standard and maintain operational readiness. The primary objectives for the multiyear study reported here were to determine the effects of military training noise on the reproductive fitness of the endangered Red-cockaded Woodpecker (RCW) and to provide noise dose-response data to guide species conservation decisions. The research was conducted at Fort Stewart, an Army installation in SE Georgia. Overall results from this three year field study are presented. [Support for this research was provided by the Strategic Environmental Research and Development Program (SERDP).]

3:00

5pAB8. Far-field auditory evoked potentials in small and large felids. Edward J. Walsh (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131), Douglas L. Armstrong, Thomas Curro (Henry Doorly Zoo, Omaha, NE 68108), and JoAnn McGee (Boys Town Natl. Res. Hospital, Omaha, NE 68131)

The survival of all tigers existing today is seriously threatened, at least partially, as a consequence of the encroachment of civilization. As a prelude to a proposed long-term study intended to determine the impact of urban noise as an extinction pressure for indigenous tiger populations, auditory brainstem responses (ABRs) were recorded from representatives of three of the four surviving tiger substrains, *Panthera tigris altacia*, *Panthera tigris tigris*, and *Pantera tigris sumatrae*, along with an adult male lion (*Panthera leo*). Standard procedures were employed to record ABRs on location at the Henry Doorly Zoo and response features were measured along with thresholds to tone bursts spanning the empirically determined frequency response range. Threshold estimates were corrected to compensate for background masking noise by applying a correction factor generated by subtracting frequency-threshold estimates derived from gerbil measurements made on location at the zoo from frequency-threshold estimates derived from measurements made under controlled conditions in the lab from the same animal. While overall response features and input–output characteristics were similar to those observed in domestic cats and other mammals, large felids responded preferentially to very low-frequency tone bursts and poorly to high-frequency tone bursts when compared with small fields.

3:15

5pAB9. The development and use of a high frequency horizontal sonar to monitor channel catfish in shallow water commercial aquaculture ponds. James Chambers, Daniel Wren, Brian Carpenter (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677), C. Douglas Minchew, and Rachel Beecham (Natl. Warmwater Aquaculture Ctr., Mississippi State Univ., Stoneville, MS 38776)

With increasing seafood demand and a declining capture in offshore fisheries, aquaculture is poised to become a major growth industry in the United States in the 21st century. A major portion of the costs associated with raising channel catfish is related to the cost of feed and aeration. These costs are directly related to the total number of fish being raised in each pond. Unfortunately, at the present time a pond would have to be totally harvested to get an accurate inventory. An alternate technique, discussed here, is the use of a high frequency (420 kHz) horizontally scanning sonar to size and count the fish in a pond. Commercial catfish ponds range in size 0.04 to 7.2 hectares. The most desirable size is about 4 hectares. These ponds will typically be about 100 m in length and 1 to 2 m in depth. This sonar is capable of 1 cm range resolution combined with an approximately 100 m maximum range. Results from field tests will be presented to show the systems effectiveness in determining the number and possibly size distribution of commercial catfish ponds. In addition design considerations and operating parameters for the shallow water environment will be discussed.

3:30

5pAB10. Biological responses to acoustical stress in fishes. Arthur N. Popper (Dept. of Biol. and Neurosci. & Cognit. Sci. Prog., Univ. of Maryland, College Park, MD 20742, apopper@umd.edu), Andrew S. Kane, and Michael S. Smith (Univ. of Maryland, College Park, MD 20742)

High level sound exposure, particularly for long periods, have significant effects on humans and other animals. Many species of fish are being exposed to increased ambient sounds associated with multiple sources including shipping, seismic experiments, sonars, and recreational boating. Previous efforts have focused on how such marine anthropogenic sounds affect marine mammals. However, the effects of such sounds on fishes have rarely been examined. Acoustical stress to fish is also likely to be found in many aquaculture facilities, yet little concern has been directed at the appropriate acoustical environment for optimal growth and development. We examined the effects of increased ambient sound on the hearing and overall health of fish species differing in hearing capabilities. Fish were reared in 600-L aquaria under quiet (110 dB *re*: 1 μPa) or noisy (170 dB *re*: 1 μPa) conditions and individuals were removed after different durations of noise exposure. We then measured alterations in noise-exposed fish relative to hearing capabilities (using auditory brainstem responses), pathology of the inner ear (using SEM), and stress (using plasma cortisol levels). Preliminary data suggest that susceptibility to noise-induced stress may be species-specific.

FRIDAY AFTERNOON, 6 DECEMBER 2002 CORAL GARDEN 1, 1:00 TO 3:30 P.M.

Session 5pBB

Biomedical Ultrasound/Bioresponse to Vibration: Ultrasound in Blood and Bone, Physiological Acoustics

Gregory Clement, Cochair
Department of Radiology, Harvard Medical School, Brigham and Womens Hospital, Boston, Massachusetts 02115

Itzala Rabadan, Cochair
ESME, IPN, Av. Juan de Dios, Botiz s/n, Col. Lindavista 07738, D. F. Mexico

Contributed Papers

1:00

5pBB1. System to measure heart performance. Armando Andrade, Heriberto Rios, Pablo R. Lizana, Ernestina Puente, and Diego Mendoza (Laboratorio de Desarrollo Tecnologico en Bioingenieria)

Systems to measure heart condition are applied to patients with early or chronic cardiac problems with the aim of diagnosing and exactly locating the problem. Two very important factors exist that are taken into account in order to obtain a reliable diagnosis and to be able to give suitable medical treatment. One of them is the volume of blood that the heart pumps, the other is the temperature gradient. In our system we measure both parameters at the same time with the purpose of determining how the heart is working from the amount of blood pumped per unit time. (To be presented in Spanish.)

1:15

5pBB2. Variation of the "Black Hole" phenomenon with speed and stroke rate. Dong-Guk Paeng and K. Kirk Shung (205 Hallowell Bldg., University Park, PA 16802, paeng@psu.edu)

The Black Hole (BH) phenomenon, a central hypoechoic hole surrounded by a hyperechoic zone, has been observed and investigated only under steady flow. The phenomenon was recently observed under pulsatile flow, but its variation during a cycle has never been reported yet. In order to better understand the BH phenomenon and its variation, a GE LOGIQ 700 Expert system was used to collect B mode images from porcine blood in a rigid tube (ID: 9.5 mm). The BH became weaker or disappeared during the systole in terms of contrast and size. The variation of the BH over a cycle was investigated with the peak speed levels (10–25 cm/s), stroke rates (20, 40, 60 beats/min), mean speed levels (10, 20 cm/s), and hematocrits (12%–46%). Under steady flow, the BH phenomenon started to appear at an asymmetrical radial position at very low speeds when speed was increased. The BH shifted to the center of the tube at higher speeds and finally disappeared at ~100 cm/s. However, the asymmetrical position at lower speeds was not observed when speed was decreased. The variation in the BH is presumably due to the red cell aggregation, which is affected by the shear rate and acceleration during a pulsatile cycle.

1:30

5pBB3. Time reversal focusing through a human skull optimized by skull surface detection algorithm. Jean-Francois Aubry, Didier Cassereau, and Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, 12 rue Vauquelin, 75231 Paris Cedex 05, France, jean-francois.aubry@loa.espci.fr)

High-intensity focused ultrasounds are mainly envisioned to cure liver and prostate cancer. In order to cure brain tumors, it has been shown that time reversal coupled with amplitude compensation is an interesting technique to correct the aberrations induced by a human skull. This focusing technique should be improved by compensating the amplitude at the exact location of the skull. To measure the geometrical shape of the skull, two different techniques are investigated and compared in terms of precision and duration of the measurement. The first method is based on the measurement of the propagation time from the array to the skull and back of pulsed signals emitted and received by each pair of transducers of an echographic array of 128 transducers. For each pair, the reflection point on the skull is located on an ellipse whose centers correspond to the positions of the emitting and receiving elements. The envelope of the complete set of ellipses is then extracted, giving the shape of the skull. This technique is compared to echographic imaging of the skull using the same array. Comparisons between these techniques will be presented using the Phased Array Simulation software and the corresponding experimental focusing.

1:45

5pBB4. Focusing ultrasound through bone and tissue layers in the wave vector frequency domain. Greg T. Clement, P. Jason White, and Kullervo Hynynen (Dept. of Radiol., Harvard Med. School, Brigham and Women's Hospital, Boston, MA 02115)

A planar projection algorithm is combined with ray theory to propagate ultrasound through an arbitrary number of randomly oriented isotropic tissue layers. The propagation information can then be used for beam aberration correction. This method, which is intended for applications in therapeutic ultrasound, calculates or measures the space–time pressure field in a plane and uses wave vector frequency-domain methods to project the field through the media. The approach requires information obtained *a priori* from MRI or CT images and is valid for longitudinal propagation through tissue and bone layers at low incident angles. The algorithm is verified by propagating fields created by a 1.5 MHz, 104-element therapy array through a combination of layered materials, including plastic phantoms, fresh porcine fat-muscle layers, and *ex vivo* human bone samples.

2:00

5pBB5. Transmission of a large unfocused 120-kHz and 1-MHz ultrasound beam through the human skull. Constantin-C. Coussios, Christy K. Holland, and George J. Shaw (Dept. of Biomed. Eng., Univ. of Cincinnati, 231 A. Sabin Way, Cincinnati, OH 45267-0586)

A three-dimensional characterization of the acoustic beam produced by two 6-cm-diameter unfocused 120-kHz and 1-MHz transducers was performed both in the free field and inside a human skull penetrated through the temporal bone. In the context of transcranial ultrasound-enhanced thrombolysis, intracranial measurements aim at determining the effect of the skull on the beam structure and peak-to-peak pressure amplitude delivered to the middle cerebral artery, which has been identified as a common site for ischemic stroke. Placement of the 120-kHz transducer at least 3.5 cm away from the temporal bone yielded a beam pattern very similar to that in the free field with less than 4.5-dB attenuation in peak-to-peak pressure amplitude. Furthermore, short pulse durations reduced the effect of reflections from the far side of the skull and small rotations of the skull relative to the beam axis hardly affected the intracranial field. By contrast, the presence of the skull was found to strongly disturb the 1-MHz beam for all configurations and to cause as much as 20-dB attenuation in peak-to-peak pressure amplitude. Therefore, 120-kHz ultrasound requires less phase aberration correction and attenuation compensation than 1 MHz for therapeutic transcranial applications.

2:15–2:30 Break

2:30

5pBB6. Frequency and porosity dependence of acoustic properties in bovine cancellous bone. S. W. Yoon, Kang Il Lee, and Heui-Seol Roh (Acoust. Res. Lab., Dept. of Phys., SungKyunKwan Univ., Suwon 440-746, Republic of Korea)

Acoustic properties of bovine cancellous bone were experimentally and theoretically investigated in the frequency range of 0.5–1 MHz. For theoretical estimation the modified Biot–Attenborough (MBA) model was employed. Phase velocity, attenuation coefficient, and BUA of cancellous bone were described as the functions of frequency and porosity. The relative orientation between the ultrasonic beam and the bone specimens was the mediolateral (ML) direction of the bovine tibia. Since the fast and slow waves were completely overlapped, only one longitudinal wave was observed for propagation in the ML direction. The MBA model for nonrigid porous medium with circular cylindrical pores well described the experimental results of acoustic wave propagation in cancellous bone. [Work supported by BK21 Program, KRF(KRF-2000-015-DP0178), and KOSEF(2000-1-11100-002-2) in Korea.]

2:45

5pBB7. Fetal electrocardiograph. Heriberto Rios, Armando Andrade, Ernestina Puente, Pablo R. Lizana, and Diego Mendoza (Laboratorio de Desarrollo Tecnologico en Bioingenieria, Calle "H" No. 11, Col. Jardines de Santa, Ecatepec Edo, Mexico C.P. 55450, Mexico)

The high intra-uterine death rate is due to failure in appropriately diagnosing some problems in the cardiobreathing system of the fetus during pregnancy. The electrocardiograph is one apparatus which might detect problems at an early stage. With electrodes located near the womb and uterus, in a way similar to the normal technique, the detection of so-called biopotential differences, caused by concentrations of ions, can be achieved. The fetal electrocardiograph is based on an ultrasound technique aimed at detecting intrauterine problems in pregnant women, because it is a noninvasive technique due to the very low level of ultrasound power used. With this system, the following tests can be done: Heart movements from the ninth week onwards; Rapid and safe diagnosis of intrauterine fetal death; Location and size of the placenta. The construction of the fetal electrocardiograph requires instrument level components directly mounted on the printed circuit board, in order to avoid stray capacitance in the

cabling which prevents the detection of the E.C.G. activity. The low cost of the system makes it affordable to low budget institutions; in contrast, available commercial systems are priced in U.S. Dollars. (To be presented in Spanish.)

3:00

5pBB8. Unique gel-coupled acoustic sensor array monitors human voice and physiology. Michael Scanlon (US Army Res. Lab., AMSRL-SE-SA, 2800 Powder Mill Rd., Adelphi, MD 20783-1197, mscanlon@arl.army.mil)

The health and performance of soldiers, firefighters, and other first responders in strenuous and hazardous environments can be continuously and remotely monitored with body-worn acoustic sensors. The Army Research Laboratory's gel-coupled acoustic physiological monitoring sensor has acoustic impedance properties similar to the skin that facilitate the transmission of body sounds into the sensor pad, yet significantly repel ambient airborne noises due to an impedance mismatch. Acoustic signal processing detects physiological events such as heartbeats, breaths, wheezes, coughs, blood pressure, activity, motion, and voice for communication and automatic speech recognition. Acoustic sensors can be in a helmet or in a strap around the neck, chest, and wrist. Although the physiological sounds have high SNR, the acoustic sensor also responds to motion-induced artifacts that sometimes obscure meaningful physiology. A noise-canceling sensor array configuration helps remove motion noise by using two acoustic sensors on the front sides of the neck and 2 additional acoustic sensors on each wrist. The motion noise detected on all 4 sensors will be dissimilar and out of phase, yet the physiology on all 4 sensors is covariant. Pulse wave transit time between neck and wrist will indicate systolic blood pressure. Data from a firefighter experiment will be presented.

3:15

5pBB9. Thermodynamical analysis of acoustical perturbations in the bronchial tree. Margarita Puente (Lab de Fenomenos Colectivos, Departamento de Fisica, Facultad de Ciencias, UNAM, Mexico), Armando Perez-Guerrero (UAM-Iztapalapa, Mexico), and Manuel Alvarado (UNAM, Mexico)

In the airways, very complex flows occur because of different conditions and the existence of a lot of complications: constantly changing temperature and pressure during the respiration process, a normally turbulent flow in the trachea which, in heavy breathing, remains so in the first three or four generations of airways, changes of the direction of the flow over the breathing cycle, from inspiration to expiration, etc. We also know the air that flows in the bronchial tree is perturbed by several sources such as the heart and the circulatory system, the diaphragm and stomach movements, etc., which produce sound waves. Thus an acoustical analysis of the phenomenon can lead us to a physical model which could help us to better understand the phenomena and to demonstrate the importance to clinical applications such as the pneumocardiograms. To this purpose we use a thermodynamical model that originally was developed to analyze supersonic air jets to explain the production of shock waves in the bronchial tree.

FRIDAY AFTERNOON, 6 DECEMBER 2002 CORAL KINGDOM 1, 1:00 TO 4:35 P.M.

Session 5pNSa

Noise and Architectural Acoustics: Soundscapes and Community Noise II

Brigitte Schulte-Fortkamp, Cochair
Institute of Technical Acoustics, Technical University Berlin, Secr TA 7, Einsteinufer 25, 10587 Berlin, Germany

Paul D. Schomer, Cochair
Schomer and Associates, Inc., 2117 Robert Drive, Champaign, Illinois 61821

Chair's Introduction—1:00

Invited Papers

1:05

5pNSa1. How to measure soundscapes. A theoretical and practical approach. Brigitte Schulte-Fortkamp (Inst. of Tech. Acoust., TU-Berlin, Einsteinufer 25, D-10587 Berlin, Germany)

Noise sources interact with the specific acoustic and environmental makeup, topography, meteorology, land use pattern, and lifestyle. The evaluation of soundscapes needs subject-related methodological procedures. With such suitable measurements a way has to be found that allows us to rely on different dimensions on reaction to noise. Improving the soundscape of an urban environment imposes to account for the qualitative appreciation as a cognitive judgment given by listeners and, particularly, for the interaction between acoustic dimensions and other sensory modalities in qualitative judgments of an urban environment (Maffiolo). The structure of the residential area that is, the combination of noise sources are important for the judgment of a soundscope and are also important as subjective parameters which are relevant in people's point of view. Moreover, the relationship of both define the background for assessments. Studies are needed on the subject and its capability in perception and interpretation; studies on the subject inside the society, studies on the social and cultural context, and field studies including physical measurements. Soundscapes may be defined in its effects on man and *vice versa* and probably acoustical ecology will serve to understand the function of soundscapes.

1:25

5pNSa2. Exploit the soundscape and reformulate traffic noise goals. Tor Kihlman (Chalmers Univ. of Technol., Appl. Acoust., 412 96 Gothenburg, Sweden)

''Soundscape support to health'' (www.soundscape.nu) is a 4 year research program including studies of traffic noise propagation in cities and different health effects of traffic noise. Special attention is given to the effect of access to a quiet side of dwellings. One goal of the program is to find more relevant goals than the present longterm general goal LAeq,24h$<$55 dB for all residents which on one hand is unrealistic and on the other hand does not imply a good environment—it is a compromise; at this level a substantial fraction of an exposed population is highly annoyed. However, the big spatial SPL variations within the city typically ranging from 40–70 dB can be systematically exploited. Traffic noise goals should therefore be reformulated to protect existing quietness and promote that all residents get access to a quiet environment within reach even though a typical level at facades facing the traffic will exceed 60 or 65 dB during a foreseeable future. In the paper how the results from the program may be used to formulate goals that are more credible than the present and that lead to decreasing negative health effects of the traffic noise are described.

1:45

5pNSa3. Semantic differential analysis on the soundscape of urban open public spaces. Jian Kang and Mei Zhang (School of Architecture, Sheffield Univ., Western Bank, Sheffield S10 2TN, UK)

Physical comfort in urban open public spaces, including acoustic comfort, has been paid increasing attention. The objective of this research is to identify factors that affect acoustic comfort and soundscape in urban open public spaces. In January 2002 a soundscape walk was organized with 50 university students and evaluation on acoustic comfort and sound quality was carried out in four typical urban open public spaces in Sheffield, UK. Both overall acoustic comfort and individual sounds were evaluated, using 28 indices. Objective measurements were made simultaneously. In March–April 2002, 500 interviews were carried out with various users in two selected open public spaces. Semantic differential analysis of the results shows that the description and evaluation of acoustic comfort and soundscape in urban open public spaces is more complicated than that for product sound quality, but it is still possible to identify several major factors including pleasant, timber, content, and evaluation. Based on the analysis, methods for evaluating acoustic comfort and soundscape in urban open public spaces are proposed. [Work supported by Key Action 4 ''City of Tomorrow and Cultural Heritage'' from the programme ''Energy, Environment and Sustainable Development'' within the Fifth Framework Programme of the European Commission.]

2:05

5pNSa4. A human-hearing-related prediction tool for soundscapes and community noise. Klaus Genuit (HEAD Acoustics GmbH, Ebertstr. 30a, 52134 Herzogenrath, Germany, klaus.genuit@head-acoustics.de)

There are several methods of calculation available for the prediction of the A-weighted sound-pressure level of environmental noise, which are, however, not suitable for a qualified prediction of the residents' annoyance and physiological strain. The subjectively felt noise quality does not only depend on the A-weighted sound-pressure level, but also on other psychoacoustical parameters, such as loudness, roughness, sharpness, etc. In addition to these physical and psychoacoustical aspects of noise, the so-called psychological or cognitive aspects have to be considered, too, which means that the listeners' expectations, their mental attitude, as well as the information content of the noise finally influence the noise quality perceived by the individual persons. Within the scope of a research project SVEN (Sound Quality of Vehicle Exterior Noise), which is promoted by the EC, a new tool has been developed which allows a binaural simulation and prediction of the environmental noise to evaluate the influence of different contributions by the sound events with respect to the psychoacoustical parameters, the spatial distribution, movement, and frequency. By means of this tool it is now possible to consider completely new aspects regarding the audible perception of noise when establishing a soundscape or when planning community noise.

2:25

5pNSa5. The contribution of noise to perception of neighborhood satisfaction, and the probability of leaving the area. RF Soames Job, Julie Hatfield (Dept. Psych., Univ. of Sydney, Sydney, Australia), Norman L. Carter (Univ. of Sydney, Sydney, Australia), Peter Peploe (Natl. Acoust. Labs., Chatswood, Australia), Richard Taylor, and Stephen Morrell (Univ. of Sydney, Sydney, Australia)

When the runaways at Sydney Airport were reconfigured, noise increased in some low-noise areas and decreased in some high-noise areas (while other areas remained highly exposed and unchanged, or minimally exposed to aircraft noise). Residents were surveyed before ($N=1015$), immediately after, and several years after the reconfiguration. This data set allows consideration of the extent to which ratings of the neighborhood are determined by noise versus other factors (parks, air pollution, safety). Residents were also asked whether they would consider moving because of the noise. From the follow-up data we examined which features of the residents satisfaction with the area and the noise, as well as features of the residents themselves (noise sensitivity, general sensitivity, personality, home ownership) predicted considering moving and moving. These results address the issue of the importance of soundscape in residential satisfaction, relative to other factors.

2:45

5pNSa6. Predicting habituation/adaptation to aircraft noise. Julie Hatfield, RF Soames Job (Dept. of Psych., Univ. of Sydney, Sydney, Australia), Stephen Morrell, Norman L. Carter (Univ. of Sydney, Sydney, Australia), Peter Peploe (Natl. Acoust. Labs., Chatswood, Australia), and Richard Taylor (Univ. of Sydney, Sydney, Australia)

Whether or not people habituate and/or adapt to a negative soundscape remains a contentious issue. The Sydney Airport Health Study has involved interviews with residents before, soon after, and several years after the runways were reconfigured. Noise increased in some low-noise areas and decreased in some high-noise areas (while other areas remained highly exposed and unchanged, or minimally exposed to aircraft noise). Of the pre-reconfiguration sample, 23.6% reported getting used to the noise. This paper examines the factors which predict who will get used to the noise and who will not, by regression against a number of predictors: noise exposure, sensitivity, attitude, personality, age, gender. The post-reconfiguration studies also allow examination of the speed with which reaction changes with changed exposure, and the factors which predict changes in reaction. Underlying mechanisms other than adaptation and habituation are considered.

3:05–3:20 Break

3:20

5pNSa7. 1/*f* dynamics in the urban soundscape. Dick Botteldooren, Bert De Coensel, and Tom De Muer (Dept. of Information Technol., Ghent Univ., St. Pietersnieuwstraat 41, 9000 Gent, Belgium)

Urban soundscapes are characterized by much more than loudness alone. Subjective description has been used by many researchers to grasp these additional dimensions, but very little objective criteria are found in literature. In this paper the dynamics of the soundscape is proposed as a potential indicator. By analyzing loudness and pitch fluctuations in longer sound fragments recorded in urban environment, $1/f$ spectral behavior at frequencies ranging from 0.01 to 10 Hz was discovered. Such behavior is typical for complex systems and was found to be very common in (classical) music, already in the seventies. Some hypotheses will be given on the origin of this frequency dependence of urban-soundscape dynamics. The possibility of discriminating soundscapes on the basis of the frequency dependence of loudness and pitch fluctuation will be illustrated. For this purpose the urban soundscapes are compared to ''extreme'' soundscapes that have a very distinct character that is clear to every observer.

3:40

5pNSa8. Socioeconomic status and the experience of noise. Aslak Fyhri (Inst. of Transport Economics, P.O. Box 6110, N0602 Oslo, Norway)

In this study how road traffic noise is experienced in an affluent (western) and a more deprived (eastern) area of Oslo is compared. The results are based on interviews with 1400 people in Oslo, together with calculations of 24 h equivalent noise levels at each individual apartment's most exposed side (the Nordic calculation method). People in the western part of town report far more positive qualities in their neighborhood than those in the east. They are more satisfied with the visual aesthetics of the neighborhood, the quality of outdoor areas, the level of air pollution, and the central location of the city area. The only exception to this somewhat rosy picture is the exposure to road traffic noise. There are more people (31% of the respondents) exposed to noise levels above 65 dBA in the western area, than in the eastern area (11%). The respondents in the west are also more annoyed by road traffic noise, compared with those in the east. When controlling for individual noise levels, the results do not support the hypothesis.

4:00

5pNSa9. The neighborhood soundscape and the residents' perceptions of its quality. Ronny Klaeboe, Aslak Fyhri (Inst. of Transport Economics, P.O. Box 6110, N-0602 Oslo, Norway), and Sigurd Solberg (KILDE Akustikk, N-5701 Voss, Norway)

In autumn of 2001 a small socio-acoustic survey was undertaken in an affluent Oslo city area. A telephone survey provided data on neighborhood quality and environmental annoyances from 400 respondents after a response rate of 38%. The Nordic calculation method was employed to estimate 24h LA_{eq} noise exposure values in front of the most exposed facade of each resident's apartment, and for various neighborhood locations. The quality of the neighborhood soundscape of each dwelling (and respondent) was thereafter characterized by a simple indicator of the maximum of these exposure values within a radius of 75 meters. Ordinal logit models were used to model people's noise annoyance when at home as a function of the noise level outside the most exposed facade and the modifying impact of the neighborhood soundscape. The analyses using data from the Oslo West city area thus replicate previous analyses undertaken in Oslo East and Drammen. However, the Oslo West questionnaire also included simple questions about the residents' perceptions of their neighborhood soundscape, and their use of their neighborhood area. The inclusion of these questions made it possible to analyze the correspondence between the estimated neighborhood soundscape quality and people's assessment of their noise neighborhood.

Contributed Paper

4:20

5pNSa10. The sagrada familia: A daily whisper of the spirit. Arturo Campos (Sagrada Familia of Barcelona. Malorca 401, Barcelona, Spain joseartca@yahoo.com)

Sounds from the top of the Temple; songs of loyals going beyond the city, sounds of the rustic forms of the spaces for the soul; reality and miracle in just one voice, voice of religion for participation; mystery, mysticism, meditation, happiness, music and songs, bells and church towers, choirs and pulpits, religion and city. The religious architecture of Gaudi is involved by mystic sounds, real or imaginary sounds, whispers almost imperceptibles and sonorous surprises in the journeys destined to peace and spiritual calm. The Sagrada Familia in Barcelona is a Temple where the sonorous itineraries will exist and where the acoustic design of choirs, bells and bell towers was an essential idea of Antoni Gaudi. A living building with its voice singing with the people a song which will take in the whole city; ringing bells and bell towers crossing the sky and sending music, view and hearing of all the people to the new cathedral, a symbol to guide pilgrimages, ceremonies, songs and festivals. The Sagrada Familia and Barcelona, a union that Gaudi wanted to turn in total and where the acoustic design will be transformed in a daily murmur, highly spiritual. (To be presented in Spanish.)

FRIDAY AFTERNOON, 6 DECEMBER 2002 CORAL SEA 1 AND 2, 1:00 TO 2:30 P.M.

Session 5pNSb

Noise: Noise in Commercial and Other Indoor Spaces

Ralph T. Muehleisen, Chair

Department of Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder, Colorado 80309

Contributed Papers

1:00

5pNSb1. A comparison of radiosity with current methods of sound level prediction in commercial spaces. C. Walter BeamerIV and Ralph T. Muehleisen (Dept. of Civil, Environ., and Architectural Eng., Univ. of Colorado, Boulder, CO 80309)

The ray tracing and image methods (and variations thereof) are widely used for the computation of sound fields in architectural spaces. The ray tracing and image methods are best suited for spaces with mostly specular reflecting surfaces. The radiosity method, a method based on solving a system of energy balance equations, is best applied to spaces with mainly diffusely reflective surfaces. Because very few spaces are either purely specular or purely diffuse, all methods must deal with both types of reflecting surfaces. A comparison of the radiosity method to other methods for the prediction of sound levels in commercial environments is presented. [Work supported by NSF.]

1:15

5pNSb2. Characterization of the occupational noise environment of the Portuguese textile sector. Pedro M. Arezes and A. Sergio Miguel (Dept. of Production Eng., School of Eng., Univ. of Minho, 4800-058 Guimaraes, Portugal, parezes@dps.uminho.pt)

In this study we aim to characterize the occupational noise environment of five large industrial companies of the textile and apparel sector, as well as the audiometric profile of the exposed workers. All individual exposures were classified into three categories: namely (1) acceptable risk: personal noise exposure level below 85 dB(A); (2) moderate risk: personal noise exposure level between 85 and 90 dB(A); and (3) high risk: personal noise exposure level greater than 90 dB(A). It was observed that the majority of the workplaces presented personal exposure levels greater than 90 dB(A), thus involving a high risk of bilateral hypoacusia. In some areas, such as weaving, that percentage exceeded 90%. Workers' audiometric profiles were analyzed according to the professional disease criteria presented by the Portuguese legislation. An effort should be made for developing Hearing Conservation Programs leading to the implementation of collective and personal protective measures. [Work supported by the Portuguese Institute for Development and Inspection of Working Conditions (IDICT).] (To be presented in Portuguese.)

1:30

5pNSb3. The propagation of sound in tunnels. Kai Ming Li and King Kwong Iu (Dept. of Mech. Eng., The Hong Kong Polytechnic Univ., Hung Hom, Kowloon, Hong Kong)

The sound propagation in tunnels is addressed theoretically and experimentally. In many previous studies, the image source method is frequently used. However, these early theoretical models are somewhat inadequate because the effect of multiple reflections in long enclosures is often modeled by the incoherent summation of contributions from all image sources. Ignoring the phase effect, these numerical models are unlikely to be satisfactory for predicting the intricate interference patterns due to contributions from each image source. In the present paper, the interference effect is incorporated by summing the contributions from the image sources coherently. To develop a simple numerical model, tunnels are represented by long rectangular enclosures with either geometrically reflecting or impedance boundaries. Scale model experiments are conducted for the validation of the numerical model. In some of the scale model experiments, the enclosure walls are lined with a carpet for simulating the impedance boundary condition. Large-scale outdoor measurements have also been conducted in two tunnels designed originally for road traffic use. It has been shown that the proposed numerical model agrees reasonably well with experimental data. [Work supported by the Research Grants Council, The Industry Department, NAP Acoustics (Far East) Ltd., and The Hong Kong Polytechnic University.]

1:45

5pNSb4. Dancing the aerobics "hearing loss" choreography. Beatriz M. Pinto, Antonio P. O. Carvalho, and Sergio Gallagher (Lab. of Acoust., Dept. of Civil Eng., Univ. of Porto, P-4200-465 Porto, Portugal, beaporumtriz@hotmail.com)

This paper presents an overview of gymnasiums' acoustic problems when used for aerobics exercises classes (and similar) with loud noise levels of amplified music. This type of gymnasium is usually a highly reverberant space, which is a consequence of a large volume surrounded by hard surfaces. A sample of five schools in Portugal was chosen for this survey. Noise levels in each room were measured using a precision sound level meter, and analyzed to calculate the standardized daily personal noise exposure levels (LEP,d). LEP,d values from 79 to 91 dB(A) were

found to be typical values in this type of room, inducing a health risk for its occupants. The reverberation time (RT) values were also measured and compared with some European legal requirements (Portugal, France, and Belgium) for nearly similar situations. RT values (1 kHz) from 0.9 s to 2.8 s were found. These reverberation time values clearly differentiate between good and acoustically inadequate rooms. Some noise level and RT limits for this type of environment are given and suggestions for the improvement of the acoustical environment are shown. Significant reductions in reverberation time values and noise levels can be obtained by simple measures.

2:00

5pNSb5. Noise levels in the learning–teaching activities in a dental medicine school. Andreia Matos, Antonio P. O. Carvalho (Lab. of Acoust., Dept. of Civil Eng., Faculty of Eng., Univ. of Porto, P-4200-465 Porto, Portugal, andreiabmmatos@hotmail.com), and Joao C. S. Fernandes (Univ. of Porto, P-4200-392 Porto, Portugal)

The noise levels made by different clinical handpieces and laboratory engines are considered to be the main descriptors of acoustical comfort in learning spaces in a dental medicine school. Sound levels were measured in five types of classrooms and teaching laboratories at the University of Porto Dental Medicine School. Handpiece noise measurements were made while instruments were running free and during operations with cutting tools (tooth, metal, and acrylic). Noise levels were determined using a precision sound level meter, which was positioned at ear level and also at one-meter distance from the operator. Some of the handpieces were brand new and the others had a few years of use. The sound levels encountered were between 60 and 99 dB(A) and were compared with the noise limits in A-weighted sound pressure level for mechanical equipments installed in educational buildings included in the Portuguese Noise Code and in other European countries codes. The daily personal noise exposure levels (LEP,d) of the students and professors were calculated to be between 85 and 90 dB(A) and were compared with the European legal limits. Some noise limits for this type of environment are proposed and suggestions for the improvement of the acoustical environment are given.

2:15

5pNSb6. Hearing loss in shoemakers exposed to noise and solvent in Porto Alegre, Brazil. Vera Steffen, Cynthia Ponte, Aleteia Alano, Renata Limberger (Universidade Federal do Rio Grande do Sul. Av. Ipiranga, 2752, Porto Alegre, Brazil, vsteffen@terra.com.br), Flavio Simoes (FAU-Ritter dos Reis, R. Orfanotrofio, 555, Porto Alegre, Brazil), Hudson Abella, Viviane Sebben, Vera Marques, Denise Pereira (Fundacao Estadual de Producao e Pesquisa em Saude, Porto Alegre, Brazil), and Samir Gerges (Universidade Federal de Santa Catarina, Florianopolis, Brazil)

Clinical and experimental studies indicate a possible harmful effect of chemicals, especially organic solvents, on the hearing system. In combined exposure to noise and solvents, very common in industry, it is most likely that a synergetic action of these factors enhances the traumatizing effect of exposure to noise. The aim of this study was to access the incidence and the risk of hearing impairment in 80 shoemakers. In this work we studied the effects of occupational exposure to solvent and noise in the hearing of workers of a factory in Porto Alegre, Brazil. In the studied group we analyzed solvents concentration in the air, biological marker for toluene in urine (hippuric acid), and levels of environmental noises. We analyze too the flow sheet of industrial production. The analysis of these results showed the bilateral hearing loss of 22.45%, and 18.37% in unilateral loss, in the workers of occupational exposure to toluene and noise, with levels allowed by the Brazilian legislation. The findings suggest that exposure to toluene and noise has a toxic effect on the auditory system. Further research is needed for the better understanding of the interaction of these and other agents in the auditory systems. [Work supported by CNq/FEPPS.]

FRIDAY AFTERNOON, 6 DECEMBER 2002 CORAL ISLAND 1 AND 2, 1:00 TO 5:00 P.M.

Session 5pPA

Physical Acoustics: Resonant Ultrasound Spectroscopy and the Acoustic Properties of Solids, Liquids and Gases

Vera A. Khokhlova, Cochair
Department of Acoustics, Physics Faculty, Moscow State University, Lenin Hills, Moscow 119899, Russia

Sergio Suarez, Cochair
Centro de Investigacion en Coputacion, Av. Juan de Dios Batiz, s/n, Col. Lindavista, CP07070 Mexico, D.F. Mexico

Contributed Papers

1:00

5pPA1. Resonance acoustic spectroscopy of lossy materials. Lev A. Ostrovsky (Zel Technologies/NOAA ETL, 325 Broadway, Boulder, CO 80305), Andrey V. Lebedev, Irina A. Soustova (Inst. of Appl. Phys., Nizhny Novgorod 603950, Russia), Alexander M. Sutin (Stevens Inst. of Technol., Hoboken, NJ 07030), and Paul A. Johnson (Los Alamos Natl. Lab., Los Alamos, NM 87501)

Resonance acoustic spectroscopy (RAS) is known as an efficient tool for determining the elastic and dissipative parameters of materials. However, its use for diagnostics of a variety of materials with a low quality factor (including materials with defects) is often impeded by overlapping of resonance responses at different modes. Here, a method of processing experimental data is suggested which enables one to determine resonance frequencies and Q-factors in cases of broad and overlapping resonances. In experiments, both undamaged and cracked samples were studied. A swept-frequency excitation was used for polycarbonate samples, whereas an impulse method with impact excitation was used for concrete samples. Signal processing was performed with the use of the suggested algorithm. In particular, the decrease of the Q-factor and splitting of resonance frequencies due to crack formation were registered. This method, which demonstrates a high efficiency even at a significant resonance overlapping, can be used for nondestructive testing of a broad class of materials.

1:15

5pPA2. Investigation of clamping force in resonant ultrasound spectroscopy. Hang Guo and Amit Lal (Dept. of Elec. and Computer Eng., Univ. of Wisconsin–Madison, 1415 Engineering Dr., Madison, WI 53706)

In this paper we studied the effect of clamping force on using resonant ultrasound spectroscopy (RUS) to determine the elasticity of solids. The foundation of RUS is the free vibration of solids under free boundary conditions. In the practical RUS measurement, the sample of solids is required to contact piezoelectric transducers for excitation and detection of its vibrations, and an amount of clamping force is needed to obtain high signal to noise ratio for the useful measured spectrum. However, the clamping force will shift natural frequencies of the sample. Thus, the accuracy of RUS is affected. In our study, finite element method is used to analyze the effect of the clamping force on the vibration of the sample. A static structural analysis for the sample under the clamping force was first done to turn the effect of the clamping force into the pre-strain of the sample, and then the natural frequencies of the pre-strained sample were calculated. Experimentally, a model to control the clamping force was set up from the bending of an equivalent cantilever. In the end, silicon samples of different sizes had been measured under different clamping forces to verify the presented method for the study of clamping force in RUS.

1:30

5pPA3. Resonant ultrasound spectroscopy applied to misoriented crystals of low symmetry: Corundum. J. R. Gladden, Jin H. So, and J. D. Maynard (Phys. Dept., Penn State Univ., University Park, PA 16802)

The vibrational resonance spectrum of a single crystal depends on the mass density, geometry, and elastic tensor of the material. Resonant ultrasound spectroscopy (RUS) derives the elastic tensor from a measured set of sample resonances. Any tensor quantity must be defined relative to some coordinate system which, for single crystals, is often coordinated with the crystal lattice vectors. If the elastic tensor coordinate system does not coincide with the principal axes of the sample, the tensor must be transformed via a rotation matrix creating an effective elastic tensor for the sample. Misorientation may occur by design, sample preparation constraints, or errors. This problem becomes particularly important for crystals of low symmetry, such as corundum, a trigonal crystal with six independent elastic constants, or when highly accurate results for derived elastic constants are desired. In this talk, techniques for analyzing such misoriented samples with RUS will be presented along with experimental data for corundum and a correction for the sign of C14.

1:45

5pPA4. Application of acoustic spectroscopy to rocks. Brian J. Zadler and John A. Scales (Ctr. for Wave Phenomena, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401)

Resonant ultrasound spectroscopy techniques were used on dried Elberton granite to infer elastic moduli and obtain experimental uncertainties for the Q values of the eigenmodes. A suite of independent measurements was performed on a 2.26-cm-diameter, 7.15-cm-long cylinder mounted between piezoelectric acoustic transducers. An orthorhombic model was fit resulting in an inferred 10% compressional wave anisotropy perpendicular to the symmetry axis. This is consistent with ultrasonic noncontacting tomography measurements made by Malcolm [A. Malcolm and J. A. Scales, preprint, 2001]. Observed Q values ranged from 80 to 206 $pm41$ over the range 20 to 90 kHz, showing a significant variation due mostly to mounting. Theory is being developed for simple isotropic pseudo-rocks to link dry moduli M_{dry} and single monolayer moduli M_{mono} to pore properties to better understand measured Q. It was observed by Tittman [B. R. Tittman, V. A. Clark, and J. M. Richardson, J. Geophys. Res. **85**, 5199–5208 (1980)] that the effects of a single monolayer of water coating the pore space can significantly impact measured Q. [Work supported by Center for Wave Phenomena and Center for Rock Abuse, Colorado School of Mines.]

2:00

5pPA5. Mapping out fiber directions in composite laminates using ultrasound. David K. Hsu (Ctr. for Nondestruct. Eval., Iowa State Univ., 1915 Scholl Rd., Ames, IA 50011, dhsu@cnde.iastate.edu)

A method for ultrasonically mapping out the fiber orientation of the plies in a composite laminate is presented. The method employs normal-incidence longitudinal ultrasound but exploits the presence of small structural imperfections that are always present at the interfaces between plies. *C*-scan images of ply interfaces contain streaking features that contain fiber orientation information; however, such images can often be quite noisy and the fiber orientations were difficult to discern visually. To overcome such difficulties, 2-D FFT was applied to the *C*-scan image of an interface, and the resultant amplitude-versus-angle plot was used to determine the fiber orientation. Since the ply thickness was an unknown and the interface echoes were usually not distinct, the placement of time windows was impractical. Therefore, the entirety of full-waveform *B*-scan data were used. The 2-D FFT was performed on *C*-scan slices at small time (depth) increments regardless of the actual location of the ply interfaces. This method was more robust and worked successfully on a large variety of composite laminates. [Work supported by NSF IUCRC.]

2:15

5pPA6. Simulations of elastic wave propagation through Voronoi polycrystals. Joseph A. Turner and Goutam Ghoshal (Dept. of Eng. Mech., W317.4 NH, Univ. of Nebraska-Lincoln, Lincoln, NE 68588, jaturner@unl.edu)

The scattering of elastic waves in polycrystalline media is relevant for ultrasonic materials characterization and nondestructive evaluation. Ultrasonic attenuation and backscatter are routinely used for extracting microstructural parameters such as grain size and grain texture. The inversion of experimental data requires robust ultrasonic scattering models. Such models are often idealizations of real media through assumptions such as constant density, single grain size, and randomness hypotheses. The accuracy and limits of applicability of these models cannot be fully tested due to practical limits of real materials processing. Here, this problem is examined in terms of numerical simulations of elastic waves through two-dimensional polycrystals. The numerical models are based on the Voronoi polycrystal. Voronoi tessellations have been shown to model accurately the microstructure of polycrystalline metals and ceramics. The Voronoi cells are discretized using finite elements and integrated directly in time. The material properties of the individual Voronoi cells are chosen according to appropriate distributions here, cubic crystals that are statistically isotropic. Results are presented and compared with scattering theories. Issues relevant to spatial/ensemble averaging will also be discussed. These simulations will provide insight into the attenuation models relevant for polycrystalline materials. [Work supported by DOE.]

2:30

5pPA7. Acoustical properties of finite sonic crystals slabs. Andreas Håkansson, José Sánchez-Dehesa (Autonomous Univ. of Madrid, Cantoblanco, 28049 Madrid, Spain), Lorenzo Sanchis, and Francisco Cervera (Polytech. Univ., 46022 Valencia, Spain)

This work shows that a slab consisting of a periodic distribution of rigid cylinders in air constitutes a sort of acoustic interferometer with properties similar to its lightwave counterpart. The performance of this finite sonic crystal has been analyzed experimentally as well as theoretically as a function of its thickness and the symmetry configuration of the

cylinders. Thus, reflectance spectra have been obtained by using the standing wave ratio technique. Experimental observations are well reproduced by a self-consistent wave theory that incorporates all orders of multiple scattering. In particular, this theory allows obtaining the complex frequencies of internal modes responsible for zero reflectance features observed in the spectra. [Work supported by CEAL-BSCH and CICyT of Spain.]

2:45

5pPA8. Statistical analysis and measurement of thermodynamical properties of a supersonic air jet. Armando Perez-Guerrero (Area de Mecanica Estadistica, Departamento de Fisica, Div. de Ciencias Basicas e Ingenieria, Universidad Autonoma Metropolitana-Iztapalapa, Mexico), Manuel Alvarado, Catalina Stern (Laboratorio de Acustica, Departamento de Fisica, Facultad de Ciencias, UNAM), and Margarita Puente (Laboratorio de Fenomenos Colectivos, Departamento de Fisica, Facultad de Ciencias, UNAM)

In this work, we present a new method for measuring the pressure and the temperature of a supersonic airflow. These quantities are necessary to evaluate mathematical models applied to the study of shock waves. By means of a microphone and a thermocouple we measure these quantities through perturbations produced inside the flow by shock waves. We applied Fourier transform and several statistic methods in order to analyze and evaluate the signals obtained.

3:00–3:15 Break

3:15

5pPA9. Evaluation of periodgrams as a tool to study spectral densities of heterodyne signals. Jose Manuel Alvarado and Catalina Stern (Lab. de Acustica, Dept. Fisica, Fac. Ciencias UNAM, Mexico)

The use of periodgrams as a tool to study heterodyne signals is evaluated. The heterodyne signal carries information about density fluctuations from light scattered by a turbulent jet. The goal of the analysis is to be able to separate the information about density fluctuations of acoustics origin (peaked at about 2 MHz) from fluctuations in the density due to entropy irregularities (wide peak centered at the origin). There are frequency regions where both overlap. Various methods are compared: Bartlett, Welch, Yule–Walker, Burg and others. (To be presented in Spanish.)

3:30

5pPA10. Rayleigh scattering to study acoustic waves in a turbulent jet flow. Carlos Azpeitia and Cesar Aguilar (Lab. de Acustica, Depto Fisica, Fac. Ciencias UNAM, Mexico)

Rayleigh scattering is used as a visualization technique to study density fluctuations due to acoustical waves produced at the exit of a turbulent air jet flow. Scattering is produced by sending a 532-nm laser beam through the flow. The scattered light is collected by a lens and projected to a screen to achieve an image. Pictures showing the crossover between expansion and compression waves in the jet have been obtained. The technique allows the mapping of the pattern of shock waves along the jet. These pictures can provide information of the exit velocity by analyzing the angle formed by the waves closest to the nozzle. Rayleigh scattering can also be used in an alternative technique that detects the scattered light by heterodyning. The current that comes out of the photodetector is proportional to the spatial Fourier transform (for a particular wave vector) of the density fluctuations as a function of time. [Work supported by UNAM through the DGAPA Project No. IN107599.]

3:45

5pPA11. Photoacoustic study of gaseous carbon disulfide with a pulsed nitrogen laser. Stanley A. Cheyne and Walter C. McDermottIII (Dept. of Phys. and Astron., Hampden-Sydney College, Hampden-Sydney, VA 23943, wcmdermott@hsc.edu)

The photoacoustic effect has been used to study gaseous Carbon Disulfide. A 1.25-MW nitrogen laser having a wavelength of 337 nm and pulse duration of 800 ps was used as the photoacoustic source when typically, continuous-wave lasers modulated at the resonant frequency of the photoacoustic test cell are used to detect rarified gases. For this work, several photoacoustic test cells, designed and constructed to optimize the acoustic signal, were used in order to detect the presence of a gas at low concentration. A variety of detectors were used including a condenser microphone, PVDF tranducer, optical probe beam deflection, and a fiber optic sensor. Results will also be presented showing the relationship between gas concentration and pressure.

4:00

5pPA12. Single-transducer continuous-wave noninvasive ultrasonic characterization of fluids. Christopher S. Kwiatkowski and Dipen N. Sinha (Los Alamos Natl. Lab., Electron. and Electrochemical Mater. and Devices Group, M.S. D-429, Los Alamos, NM 87545)

Ultrasonic fluid characterization studies that utilize a continuous wave (cw) frequency-sweep technique typically employ a source transducer and a receiver transducer. This arrangement can lead to problems such as transducer misalignment, diffraction effects, and mechanical and electrical coupling. In contrast, the traditional tone-burst technique that uses a single transducer for both transmit and receive functions can alleviate some of the problems associated with the cw technique. However, the frequency-sweep technique allows for much higher signal-to-noise ratios and observation of relevant phenomena that occur early in time that can be missed by the tone-burst technique. Reflection-mode single-transducer techniques have been developed in the past, but these usually require high-precision bridge circuits and exact balancing circuitry to provide satisfactory results over a narrow frequency range. Due to impedance-matching issues, the dynamic range of such measurements can be problematic. We describe a single-transducer cw frequency-sweep technique that takes advantage of the high dynamic range of current commerically available instrumentation, which allows noninvasive fluid characterization. The advantages of using such a device, as well as experimental results and a theoretical model, will be presented. Comparisons with two transducer designs will also be discussed.

4:15

5pPA13. Ultrasonic investigations of water mixtures with polyethylene glycol 1000. Bogumil B. J. Linde (Univ. of Gdansk, Inst. of Exp. Phys., ul. Wita Stwosza 57, 80-952 Gdansk, Poland), Christian Schmelzer, and Eike Rosenfeld (Fachhochschule Merseburg, Fachbereich Informatik und Angewandte Naturwissenschaften, Merseburg, Germany)

The velocity of ultrasound and density were measured in the mixture of polyethylene glycol 1000 and water in the temperature ranges 283–308 K. These results allowed us to calculate the adiabatic compressibilities from the Laplace equation. The compressibility dependencies on the temperature and concentration are presented as group of isotherms with one common point of intersection. Such a relation between adiabatic compressibility and concentration and temperature indicates the formation of a liquid pseudostable structure. It was compared with the solid clathrate. This structure were suggested by Baumgarten & Atkinson as well as by Endo, although the interpretation was different. The first authors relate the formation of clathrate structure to the minima of the compressibility isotherms but according to Endo the intersection of the compressibility isotherms are related to such a structure. None of these were applicable to the results.

4:30

5pPA14. An experimental study for scale prevention in boiler by use of ultrasonic waves. Pil Woo Heo, Yang Lae Lee, Eui Su Lim (Korea Inst. of Machinery and Mater., Taejon, Korea), and Kwang Sik Koh (Kyungpook Natl. Univ.)

In case of a boiler, scale is made on the surface of tube by the chemical reactions of Ca and Mg ions contained in the water and heat transfer rate is reduced because of the increment of heat resistance. Thus, it brings about the reduced energy transfer efficiency and also environmental pollution due to the use of chemicals for scale removing. In this paper, we have first investigated the effects of irradiated ultrasonic wave on water with impurities in a beaker. The experiments show that exposed water is less transparent and has finer particles as compared to unexposed water. This means that the ultrasound shakes water in the beaker and breaks out particles, and so broken fine particles are suspend in the state of less precipitation. Second, the laboratory experiment with an exposed sample in the similar condition to the boiler shows better scale prevention effects as compared to the unexposed one. And also 20 kHz ultrasound represents about 3 times better scale prevention effects than 40 kHz. Finally, scale prevention experiments in the real small boiler tell us exposed sample results in 3659% less scale formation than an unexposed one.

4:45

5pPA15. Comparative study of acetate, chloride and nitrate ions in aqueous and methanol media. Rajarathinam Ghandhimathi (Urumu Dhanalakshmi College, Trichirapalli, Tamilnadu, India)

A good deal of work has been carried out in aqueous and methanol solutions of various electrolytes at various concentrations and different temperatures ranging from 15 °C to 55 °C through ultrasonic measurements which provide qualitative information regarding the elastic forces between molecules on the basis of internal pressure and free volume. The variation of internal pressure and free volume of nitrate, chloride, and acetate ions in aqueous and methanol solutions is compared. It is found that the internal pressure increases rapidly with concentration in chloride systems and decreases with concentration in solutions of nitrates. It increases very slightly with concentration in the case of acetates. The difference in internal pressure predicts the nature of solute. If the difference in internal pressure is positive, the solute has a tendency to enhance the internal pressure of the solvent when added to it. Hence, there is an increase in internal pressure in the case of chloride ions. If the difference in internal pressure is negative, the internal pressure will decrease, which is observed in the case of nitrate ions.

FRIDAY AFTERNOON, 6 DECEMBER 2002 GRAND CORAL 3, 1:00 TO 4:00 P.M.

Session 5pSC

Speech Communication: Prosody and Phonation (Poster Session)

Mark K. Tiede, Cochair
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

Jorge C. Lucero, Cochair
Department de Matematica, University de Brasilia, Campus Universitario, Brasilia DF 70910-900, Brazil

Contributed Papers

All posters will be on display from 1:00 p.m. to 4:00 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 2:30 p.m. and contributors of even-numbered papers will be at their posters from 2:30 p.m. to 4:00 p.m.

5pSC1. Visual correlates to focal accent and their relation to the fundamental frequency contour. Sonia Sangari (Dept. of Computer and Information Sci., Linkoping Univ., SE-581 83, Linkoping, Sweden, sonsa@ida.liu.se)

Focal accent in speech is signaled by a fundamental frequency manifestation and by prolonged segment duration in connection with the manifestation. In the present investigation, the visual correlates to focal accent in read speech are studied and a computational model is hypothesized. The acoustics is synchronized with the head movements in order to relate the movements to the manifestations of the fundamental frequency contour. It is found that head movements are correlated to the primary stressed syllable of the word-carrying focal accent. The obtained data show that it is possible to model the movements by means of a second-order linear system.

5pSC2. The effects of variations in voice type on the acquisition of second language vocabularly. Joe Barcroft (Dept. of Romance Lang. and Lit., Washington Univ., Campus Box 1077, St. Louis, MO 63130) and Mitchell S. Sommers (Washington Univ., St. Louis, MO 63130)

The present study was designed to examine whether variations in voice type affect second language (L2) vocabulary acquisition relative to conditions with constant acoustic-phonetic mappings. Listeners with no previous formal training in Spanish attempted to learn 24 Spanish words in one of three conditions: (1) No Variability: all items were presented in a neutral or conversational voice type; (2) Moderate Variability: items were presented in three different voice types (neutral, nasal, and time-expanded); and (3) High Variability: items were presented in six different voice types (the three voice types used in the moderate variability condition along with whispered, excited, and pitch-shifted). All stimuli were recorded by a native speaker of Spanish and pilot testing indicated nearly

perfect identification of the test stimuli by native speakers. Accuracy and latency measures were obtained for both production and translation of L2 vocabulary. No significant differences between the three learning conditions were observed, indicating that acoustic variation had little effect on the obtained measures of L2 vocabulary acquisition. The results are discussed from modular and interactive perspectives, focusing on how diverse acoustic examples of words may become normalized during second language acquisition.

5pSC3. An acoustic model of normal speech prosody. Maureen A. Casper (The College of New Rochelle Grad. School, 29 Castle Pl., New Rochelle, NY 10805), Lawrence J. Raphael (Adelphi Univ., Long Island, NY), and Katherine S. Harris (Haskins Labs., New Haven, CT)

Investigations of the mechanisms underlying durational effects have shown that the displacements and velocities of articulator movement for the same acoustic-phonetic segments are variable. Kinematic measures reveal more extreme displacements and increased velocity in stressed syllables as compared to weakly stressed syllables [Cohen *et al.,* J. Acoust. Soc. Am. **98**, 2894(A) (1995)], except for those in phrase-final position. To further our understanding of the acoustic consequences of prosodic variations, six normal speakers read sentences containing the syllable [pap] in response to a series of questions. Pairwise comparisons of the prosodic conditions showed statistically significant differences for four prosodic contrasts. For three of the four contrast, the speakers showed longer syllable durations and higher formant and $f0$ frequency values for stressed syllables. In the phrase-final accented versus non-phrase-final accented contrast, however, the phrase-final accented syllables, although showing longer syllable duration, displayed lower frequency values for $F1$, $F2$, and $f0$. These data suggest a model in which time emerges as a function of articulatory events rather than a determiner of those events.

5pSC4. Acoustic effects of speaking rate changes in articulatory models. Jeff Berry (Univ. of Wisconsin, Waisman Ctr., Rm. 493, 1500 Highland Ave., Madison, WI 53705)

Speaking rate effects on articulatory phasing, velocity, and magnitude have been reported in the speech-kinematic literature. Yet little is known about the acoustic consequences of manipulating these dimensions. Using an articulatory modeling method developed by Milenkovic, simulations of the acoustic effects of changes in articulatory phasing, velocity, and magnitude have been explored. Articulatory models have been optimized for speakers from the University of Wisconsin X-Ray Microbeam Speech Production Database. Preliminary work has revealed that models of rate change which use only a single dimension (e.g., phasing) may reveal less plausible acoustic solutions than models involving multiple dimensions. Using sample productions of ''a boy'' to determine articulatory parameter time histories, the acoustic consequences of rate change simulations have been described by the F2 value at onset (F2on), the maximum F2 value (F2tar), and the shape of the F2 trajectory throughout the diphthong. Speaking rate models that combine velocity and phase changes to articulatory parameters result in a reduction in the magnitude of the F2on effect, near elimination of any effect on F2tar, and a relative increase in the homogeneity of the F2 trajectory shapes across speaking rates when compared with phase-only models. [Work supported by NIH (NIDCD) P60 DC1409-10.]

5pSC5. Continuous temporal alignment of an $F0$ peak produces categorical behavior. Laura C. Redi (Speech Commun. Group, RLE, MIT, 77 Massachusetts Ave., Cambridge, MA 02139, redi@mit.edu)

Assessing which aspects of fundamental frequency ($F0$) contours speakers are capable of reliably reproducing provides insight into how $F0$ patterns should be modeled for synthesis purposes, as well as how speakers represent such patterns. This experiment investigated the reproducibility of the temporal alignment of an $F0$ peak with respect to a stressed syllable. Subjects were required to imitate a series of synthetic stimuli in which an $F0$ peak was shifted through the stressed and following weak syllable of the word *minglingly*. Preliminary results showed that $F0$ peaks in the subjects' reproductions of the stimuli clustered about the midpoint of the stressed or following weak syllable when the corresponding stimulus had a peak in the stressed or following weak syllable, respectively. Moreover, relatively large variation within a syllable in stimulus peak timing resulted in little variation in peak timing in subjects' corresponding reproductions, while stimuli with peaks near the boundary between the stressed and weak syllables induced a large degree of variation in peak timing in the subjects' reproductions. This suggests that listeners may be making categorical distinctions between some intonation patterns based on the temporal location of an $F0$ peak with respect to a stressed syllable. [Work supported by NSF.]

5pSC6. Prosodic evaluation of speech. Juan A. Hernandez and Rene O. Arechiga (UAM-A, Av. San Pablo #180 Col Reynosa Tamaulipas, Delegacion Azcapotzalco, Cp. 02200 Mexico, D. F., arechiga@correo.azc.uam.mx)

This work presents English prosodic evaluation advances: specifically intonation and accent as spoken by hispanics. A widely used algorithm Hirst (Momel) [D. J. Hirst and R. Espesser, Travaux de l'Institut de Phontique d'Aix **15**, 71–85 (1993)] to determine the Fundamental Frequency contour ($F0$), based on quadratic regression and splines is reproduced, and an alternative method, with comparable results, based on signal processing is proposed. We are presently experimenting with Hidden Markov Modes for accent evaluation. Intonation is the melody of speech and helps in transmitting emotions and intentions. The feature most commonly associated with intonation is the Fundamental Frequency of the speech signal. The accent refers to the emphasis given to a particular syllable within a word. The features associated with accent are: duration, pitch, and intensity of vocalic nuclei. The Hirst (Momel) algorithm is divided into five steps: preprocessing of Fundamental Frequency, estimation, partition and reduction of target candidates, and interpolation of the curve. The alternative method consists of: Preprocessing of Fundamental Frequency; Detection of voiceless regions; Fill-up of voiceless regions; Digital filtering, downsampling and upsampling. (To be presented in Spanish.)

5pSC7. Strategies for the timing of the fundamental frequency contour. Bertil Lyberg (Telia Res. AB, SE-123 86 Farsta, Sweden)

The acoustic manifestation of the prosodic features are heavily influenced by the segmental composition of the utterances. In investigations about the fundamental frequency contour utterances built of only sonorants are often used and thereby the influence of constrictions of the vocal tract is avoided or at least diminished. In order to arrive at a detailed model of the fundamental frequency it is necessary to exactly know what is happening in different segmental environments. In the present study we are examining the effects which voiceless consonants have upon the fundamental frequency contour. Three different hypotheses are considered: the truncation, the timing adjustment and the rate adjustment hypothesis. The results show that the truncation hypothesis seems to be of limited validity at least from a detailed description of the fundamental frequency contour.

5pSC8. Intonational phonology and the domain of liaison in French. Rebecca Brown and Sun-Ah Jun (Dept. of Linguist., UCLA, Los Angeles, CA 90095-1543, rab@humnet.ucla.edu)

French liaison is a phenomenon where certain word final consonants are pronounced as the onset of a following vowel-initial word. The application of liaison varies depending on a number of factors including morphological, phonological, syntactic, and sociolinguistic factors. For the cases where liaison applies postlexically, prosodic phonologists have claimed that liaison applies within a phonological phrase (PhP), a prosodic

domain indirectly defined by syntax. However, many of the empirical studies (e.g., Morin and Kaye, 1982; de Jong, 1994; Post, 2000) have found data that cannot be accounted for by this view, suggesting that this characterization is inadequate. In the current study, liaison data were examined to see if the domain of liaison matches a prosodic unit defined by intonation under the framework of intonational phonology. Data were taken from Post (2000): 27 sentences [nine each in three conditions: the target sequence (C#V) being inside a PhP, across optional PhPs and obligatory PhPs]. Sentences were randomized and printed with no internal punctuation. Intonation was analyzed using Jun and Fougerons (1995, 2000, in press) model of French intonation. Preliminary results show that the domain of liaison matches an accentual phrase, 91% (including all the optional PhPs), higher than the results predicted by PhP.

5pSC9. On the origins and scope on final lowering. Amalia Arvaniti and Lara Taylor (Dept. of Linguist., UCSD, 9500 Gilman Dr., #0108, La Jolla, CA 92093, amalia@ling.ucsd.edu)

This experiment examines final lowering, the lower than expected scaling of the last pitch accent in downstepping word lists. In word lists of four or five items two parameters were manipulated. First the number of unaccented syllables between accents was either constantly two (''LIma beans, GREEN beans and SOY beans''); alternatively, the presence of ''and'' resulted in an extra syllable between the penultimate and final accent (''LIma beans, NAvy beans and SOY beans''). Second, the number of unaccented syllables after the final accent was either one (''SOY beans'') or three (''HAricot beans''). The results showed evidence of final lowering, in that the final accent in four word lists was lower than the fourth accent in five word lists, but the presence of ''and'' did not produce extra lowering of the final peak, suggesting that final lowering is an independent phenomenon, not an epiphenomenon of declination (contra-Grabe). On the other hand, the final accent was scaled higher when the number of unaccented syllables following it was larger. This suggests that final lowering is not under phonological control but a physiological phenomenon, since it does not target the last accent but a fixed time interval before the end of the utterance.

5pSC10. Modeling prosody: Different approaches. Lesley M. Carmichael (Dept. of Linguist., Univ. of Washington, Seattle, WA 98195-4340)

Prosody pervades all aspects of a speech signal, both in terms of raw acoustic outcomes and linguistically meaningful units, from the phoneme to the discourse unit. It is carried in the suprasegmental features of fundamental frequency, loudness, and duration. Several models have been developed to account for the way prosody organizes speech, and they vary widely in terms of their theoretical assumptions, organizational primitives, actual procedures of application to speech, and intended use (e.g., to generate speech from text vs. to model the prosodic phonology of a language). In many cases, these models overtly contradict one another with regard to their fundamental premises or their identification of the perceptible objects of linguistic prosody. These competing models are directly compared. Each model is applied to the same speech samples. This parallel analysis allows for a critical inspection of each model and its efficacy in assessing the suprasegmental behavior of the speech. The analyses illustrate how different approaches are better equipped to account for different aspects of prosody. Viewing the models and their successes from an objective perspective allows for creative possibilities in terms of combining strengths from models which might otherwise be considered fundamentally incompatible.

5pSC11. Rhythmic organization of read word lists. Fred Cummins (Dept. of Comput. Sci., Univ. College Dublin, Dublin 4, Ireland)

Rhythmic constraints have been notoriously hard to identify in speech timing. Appropriate experimental tasks may help in identifying grouping principles and timing constraints. In a baseline experiment, subjects read wordlists which contained eight trochees. They were instructed to read the words with regular timing. The inter–onset intervals obtained show evidence of a binary grouping into ''feet'' containing two words each, and of two-word feet grouping into 4-word super-feet. Results were consistent with a hierarchical rhythm grouping of units into pairs, with the second element of each such group exhibiting some degree of deceleration or final lengthening. In a second condition, each word list contained 7 trochees and one (list medial) iamb. Inter–onset intervals suggest that a similar grouping takes place, with the stressed syllable of the iamb, rather than the first syllable, marking the onset of a unit. In a third condition, the iamb was preceded by a dactyl, giving a 4-syllable foot. Again, results suggest that the stressed syllables, rather than word onsets, delimit production units. Together the results suggest that the stress foot is a unit which serves to structure speech in time.

5pSC12. The influence of onset-density in speech production. Michael Vitevitch (Dept. of Psych., 1415 Jayhawk Blvd., Lawrence, KS 66045)

Numerous studies have suggested that the initial phoneme (or syllable onset) plays an important role in speech perception and production. Recent findings have also demonstrated the importance that the number of similar sounding words plays in lexical retrieval during speech perception and production. The results from a picture-naming task demonstrate that the number of words that share the onset of the target word (onset-density) also influences speech production as well as speech perception. Specifically, pictures of words with sparse onsets were named more quickly than pictures of words with dense onsets. These results suggest that lexical and sublexical information interact during lexical retrieval in speech production. The implications for feedforward and feedback models of speech production will be discussed.

5pSC13. Effects of prosodic position and phonetic context on Spanish coda stops. Marisol del-Teso-Craviotto (Dept. of Linguist., Cornell Univ., 203 Morrill Hall, Ithaca, NY 14850)

The paper reports on a study of Peninsular Spanish stops in coda position. The voicing and duration of the stops [b p d t g k] was measured in initial (V.CV) and syllable final (VC.C) contexts to study the possible effects of prosodic position and segmental phonetic context on the realization of the stops under study. Results show that in initial position, stops overwhelmingly surface with their underlying voicing, whereas in coda position there is evidence of weakening processes. Thus voiceless stops have voiced outcomes in 56.4% of the cases. However, weakening cannot account for the high percentage of voiceless outcomes for both underlying voiceless and voiced stops in coda position (43.6% and 31.7%, respectively.) The majority of these voiceless outcomes can be accounted for by regressive assimilation. The results of the study also show voicing differences between stops followed by obstruents, which surface as voiced in 52.1% of the cases, and the stops followed by sonorants, which surface as voiced 80.6% of the time. This difference provides limited support for Steriades [manuscript (1999)] proposal that voicing is licensed by particular phonetic cues and not by prosodic position [L. Lombardi, Natl. Lang. Ling. Theory **13**, 39–74 (1995)].

5pSC14. Acoustic profiles of negative emotion. Marc D. Pell (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada)

A study was initiated to acoustically characterize and differentiate discrete categories of negatively valenced emotions conveyed through speech prosody. Utterances elicited from eight encoders (actors) in different emotional tones were perceptually rated by a group of decoders to gauge how strongly each token was associated with the basic emotions of ''anger,'' ''disgust,'' and ''sadness'' using a seven-choice response paradigm. Tokens rated as highly representative of each target emotion by greater than

80% of decoders were examined acoustically. Measures of fundamental frequency (mean, range, sd), amplitude (mean, range, sd), and duration (speech rate, %voiced) were obtained from each token and for utterances spoken in a ''neutral'' tone by the same encoders. Normalized measures were compared among emotional categories to uncover reliable acoustic dimensions that may have contributed to perceptually distinct vocal symbols of negative emotion states. Results pointed to important differences in duration, amplitude, and especially fundamental frequency in discriminating among prosodic signals representing distinct negative emotions. These findings extend work on the acoustic underpinnings of positive and negative vocalizations in speech [M. D. Pell, J. Acoust. Soc. Am. **109**, 1668–1680 (2001)], providing finer specification of these parameters within the family of ''negative'' emotions. [Work supported by NSERC.]

5pSC15. The link between prosody and meaning in the production of emotional homophones. Lynne C. Nygaard, Neelam Patel, and Jennifer S. Queen (Dept. of Psych., Emory Univ., Atlanta, GA 30322)

Recent research has shown that characteristics of a talker's voice are preserved in memory representations of spoken words. These findings suggest that aspects of surface form that co-occur frequently with a word's phonological form and meaning may influence lexical processing. To determine if surface form is linked to word meaning, this study examined the production of emotional homophones including one emotional (happy or sad) and one neutral meaning, each with distinct spellings (e.g., bridal/bridle). It was hypothesized that each meaning of these homophones would be linked to different characteristic surface forms, with emotional prosody linked to the emotional meaning and no systematic prosody linked to the neutral meaning. Participants read aloud either the emotional or the neutral spelling of each homophone and their utterances were recorded and acoustically analyzed. For happy homophones, average fundamental frequency and rms amplitude were higher when participants read aloud the happy spelling than when they read the neutral spelling of the same homophone. For sad homophones, shimmer and jitter measures were greater when reading the sad then the neutral spelling. These results suggest that emotional prosody is linked to emotional meaning in spoken words and this link appears to guide speech production.

5pSC16. The effects of talker characterisics on implicit memory for spoken sentences. Jennifer Queen, Lynne Nygaard, and Lawrence Barsalou (Dept. of Psych., Emory Univ., Atlanta, GA 30322)

Recent research has shown that surface characteristics of speech such as talker identity are retained in long-term memory for individual spoken words, but little research has examined retention of surface form in the processing of larger linguistic units. The present experiment investigated the effects of talker identity and talker-relevant background information on implicit memory for spoken sentences. Listeners heard a conversation between two speakers. The conversation focused on 2 topic areas, football at Emory and lunching at the dining hall. At test, listeners were presented with sentences in noise that were either previously heard in the conversation or new sentences on the same topics. Half of the old test items were spoken by the speaker who said them during the conversation and half were spoken by the other speaker. Results showed that listeners correctly identified more sentences at test when they were repeated by the same speaker than by the other speaker. The results suggest that talker characteristics are retained and used in sentence as well as spoken word recognition.

5pSC17. Difference limen for aspiration noise in perception of breathy voice quality. Rahul Shrivastav and Christine Sapienza (Univ. of Florida, P.O. Box 117420, Dauer Hall, Gainesville, FL 32611)

Despite much research, the relationship between the vocal acoustic signal and the perception of breathy voice quality is debated. Shrivastav (2001) used an auditory model to study the perception of breathy voice quality and found that listeners' perception of breathiness was related to the partial loudness of the periodic energy obtained with the aspiration noise acting as a masker. This measure was found to be a better predictor of perceptual judgments than conventionally used acoustic measures of breathiness. The aim of the present experiment was to determine the minimal change in aspiration noise that led to a change in the perception of breathiness. A continuum of voices, identical in all aspects except aspiration noise, was synthesized using a Klatt synthesizer. These stimuli were presented in pairs and listeners identified each pair as being the same or different in terms of their breathiness. Percent-correct scores for each stimulus pair were calculated and compared with differences in the acoustic and auditory spectral measures. Findings of this experiment may be useful in the development of objective tools for the quantification of breathy voice quality.

5pSC18. A simple hydrodynamic semi-continuum model of vocal-fold motion. Drew LaMar (Dept. of Mathematics, Univ. of Texas, Austin, TX 78712), Yingyong Qi (Qualcomm, Inc., San Diego, CA 92121), and Jack Xin (Univ. of Texas, Austin, TX 78712)

Vocal-fold (VF) motion is a fundamental process in voice production, and is also a challenging problem for direct numerical computation because the VF dynamics depends on nonlinear coupling of airflow with the response of VFs, which undergo opening and closing, and induce internal flow variation. A traditional modeling approach makes use of Bernoulli's law to treat the airflow, which is known to be inaccurate during VF opening. A new hydrodynamic semi-continuum system for VF motion is formulated. The airflow is modeled by a quasi-one-dimensional continuum aerodynamic system, and the VF by a classical lumped two-mass system. The reduced flow system contains the Bernoulli's law as a special case, and is derivable from the two-dimensional compressible Navier–Stokes equations. Since no (quasi-)steady approximation is made, transients and rapid changes of solutions are captured, e.g. the double pressure peaks at opening and closing stages of VF motion in agreement with experimental data. It is demonstrated numerically, via a split-time finite difference method, that the system is simple, robust, and models *in vivo* VF airflows and oscillations efficiently. [Work partially supported by NSF and ARO.]

5pSC19. Observed intraglottal pressures, their analytic representations, and driving forces for two-mass models. Ronald C. Scherer (Dept. of Commun. Disord., Bowling Green State Univ., Bowling Green, OH 43403, ronalds@bgnet.bgsu.edu), Guangnian Zhai, Zipeng Zhu, and Lewis Fulcher (Bowling Green State Univ., Bowling Green, OH 43403)

Multimass modeling of phonation should be based on valid aerodynamic characteristics. Measured glottal wall pressures over a wide range of glottal geometries and transglottal pressures were used to obtain average driving forces on the surfaces of two-mass models of phonation. A Plexiglas model of the larynx was used for this purpose. Predictions from the classic equations of Ishizaka and Flanagan were compared to these driving forces and movies of the calculated motions of the two-mass models were produced. The data show that the magnitudes of the negative pressures on the lower glottal region tend to be considerably less than those predicted by the Ishizaka–Flanagan equations, which leads to an increase in the number of cycles (longer time) for the vocal folds to reach steady state, lower peak volume flows, and larger open quotients. Subsequently, analytic expressions were used to represent the intraglottal pressures, with special attention to the entrance losses and the form used to describe the Poisseuille frictional effects, and their accuracy was assessed by a comparison with the measured flow rates and pressures. Results from the analytic representations were compared with those generated from the Ishizaka and Flanagan formulation. [Work supported by NIH Grant No. 2 R01 DC03577.]

5pSC20. EM sensor measurement of glottal structure versus time. John F. Holzricher (Univ. of California, Davis and Lawrence Livermore Natl. Lab., L-1, P.O. Box 808, Livermore, CA 94550), Lawrence C. Ng, Gerald J. Burke (Lawrence Livermore Natl. Lab., Livermore, CA 94550), James B. Kobler, and John J. Rosowski (Massachusetts Eye and Ear Infirmary, Boston, MA 02114)

EM wave sensors are being used to measure human vocal tract movements during voiced speech. However, when used in the glottal region there remains uncertainty regarding the contributions to the sensor signal from the vocal fold opening and closing versus those from pressure induced trachea–wall movements. Several signal source hypotheses were tested on a subject who had undergone tracheostomy 4 years ago as a consequence of laryngeal paresis. Measurements of vocal fold and tracheal wall motions were made using an EM sensor, a laser-Doppler velocimeter, and an electroglottograph. Simultaneous acoustic data came from a subglottal pressure sensor and a microphone at the lips. Extensive 3-D numerical simulations of EM wave propagation into the neck were performed in order to estimate the amplitude and phase of the reflected EM waves from the 2 different sources. The simulations and experiments show that these sensors measure, depending upon location, both the opening and closing of the vocal folds and the movement of the tracheal walls. When placed over the larynx, the vocal folds are the dominant source. The understanding of the signal sources is important for many potential applications under consideration.

5pSC21. Pressure and velocity profiles in a static mechanical hemilarynx model. Fariborz Alipour (Dept. of Speech Pathol. & Audiol., Univ. of Iowa, 334 WJSHC, Iowa City, IA 52242, alipour@shc.uiowa.edu) and Ronald C. Scherer (Bowling Green State Univ., Bowling Green, OH 43403)

This study examined pressure and velocity profiles in a hemilarynx mechanical model of phonation. The glottal section was fabricated from hard plastic. Twelve pressure taps were mounted in the vocal-fold surface and connected to a differential pressure transducer through a pressure switch. The glottal gap was measured with a filter gauge and the glottal angle was well specified by use of a laser system. Eight pressure transducers were placed in the flat Plexiglas wall opposite the vocal fold. Hot-wire anemometry was used to obtain velocity profiles upstream and downstream of the glottis. The results indicate that the pressures were asymmetric in the glottis: the pressure distribution on the vocal fold was consistent with pressure change along a parallel duct, whereas the pressures on the opposite flat wall typically were lower. The upstream velocity profiles were symmetric regardless of the constriction shape and size. The jet velocity downstream of the glottis was turbulent even for laminar upstream conditions. The narrower the glottis, the closer the jet stayed to the upper wall and traveled further before expanding. The turbulence ratio was greater on the skirt of the jet and reached as high as 25% of the maximum jet velocity.

5pSC22. Analysis of vocal fold vibration on prosodic events by high speed digital imaging systems. Masanobu Kumada (Natl. Rehabilitation Ctr. for the Disabled, 4-1 Namiki, Tokorozawa Saitama 359-8555, Japan, kumada@rehab.go.jp), Noriko Kobayashi, Hajime Hirose (Kitazato Univ., Kanagawa, Japan), Niro Tayama (Natl. Hospital Med. Ctr.), Hiroshi Imagawa (Univ. of Tokyo, Tokyo, Japan), Ken-Ichi Sakakibara (NTT Commun. Sci. Labs., Kanagawa, Japan), Takaharu Nito (Univ. of Tokyo, Tokyo, Japan), Mamiko Wada (Tokyo Metropolitan Rehabilitation Hospital, Tokyo, Japan), Shin'ichi Kakurai, Chieko Kumada (Univ. of Tokyo, Tokyo, Japan), and Seiji Niimi (Intl. Univ. for Health and Welfare, Tochigi, Japan)

A physiological study of prosody is indispensable in terms of the physiological interest as well as of the evaluation and treatment for pathological cases of prosody. We have already reported the study with our high speed digital imaging system [143rd ASA Meeting, 2002] that two major categories (Categories A and B) were found in the ways of changing of vocal fold vibrations at the register change in sustained phonations. In Category A, changes were very smooth in terms of the vibration of the vocal fold vibration. In Category B, changes were not so smooth with some additional events at the register change, such as the anterior–posterior phase difference of the vibration, the abduction of the vocal folds, or the interruption of the phonation. The number of the subtypes for Category B was thought to increase if more subjects with more range of variety are analyzed. In this study, the changes of vocal fold vibration on prosodic events in free talk as the pitch changes were analyzed, using two types of high speed digital imaging system: the one with recording time of two seconds at a rate of 4,500 images of ~256 pixels per second; the other with a recording time of seven seconds at the same rate with the same spacial resolution as the former. This study showed that high speed digital imaging systems were very powerful tools for the study of prosody.

5pSC23. On the acoustic effects of piriform recesses in speech production. Jul Setsu Cha and Abeer Alwan (Dept. of Elec. Eng., UCLA, 66-147 E EIV, 405 Hilgard Ave., Los Angeles, CA 90095, jul@icsl.ucla.edu)

The 3-D vocal tract (VT) geometries for several talkers were reconstructed from consecutive 2-D MR images. Cross-sectional areas were then estimated and the corresponding VT transfer functions were synthesized using an articulatory speech synthesizer. Geometries for eleven English vowels spoken by two males, and seven vowels by one female were obtained. The acoustic effects of the piriform recesses, which extend interiorly on either side of the larynx, were studied by modeling the recesses as side branches. For many vowels, the inclusion of the piriform recesses, which results in lowering the formant frequencies, yields better estimates of the natural formants. The largest effect is noticed on $F1$ where the value of the formant is inversely related to the recesses' lengths. In addition, zeros in the VT transfer function are created and estimated well when the length of the recess is large. The length of a piriform recess seems to have a larger effect than the volume on the location of the zero frequencies in the modeled VT frequency response. Detailed results are presented and compared to other studies that examined the acoustic effects of these recesses [Dang and Honda, J. Acoust. Soc. Am. **101**, 456–465 (1997); Story *et al., ibid.* **104**, 471–487 (1998)].

5pSC24. Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck. Harold A. Cheyne II (Voice and Speech Lab., Massachusetts Eye and Ear Infirmary, 243 Charles St., Boston, MA 02114, Harold_Cheyne@meei.harvard.edu)

Quantifying vocal function is becoming increasingly important both for corroborating clinicians' subjective impressions during a voice evaluation and for assessing the effectiveness of surgery or voice therapy. This work investigated the extent to which measures of the acceleration of the skin on the neck (near the larynx) could be used to estimate important vocal function parameters. Simultaneous acceleration, acoustic and aerodynamic recordings from one male subject with a normal voice were made to examine relationships between the acceleration signal and the acoustic or aerodynamic signals. A vocal system model was also developed to provide insights into these relationships. The model's acceleration-derived measures of Maximum Flow Declination Rate (MFDR) and Sound Pressure Level (SPL) were compared to the same measures obtained from the airflow (MFDR) and acoustic (SPL) signals at the lips. Results showed high correlations between the two MFDR ($r^2=0.89$) and SPL ($r^2=0.75$) measures, suggesting that acceleration-based measures have the

potential to provide quantitative data on vocal function. [Funding for this research was provided by funding from NIDCD Grant No. R21 DC05397 and SBIR Grant No. R44 DC02917.]

5pSC25. Examination of models of mucosal wave propagation. David A. Berry (UCLA Div. of Head & Neck Surgery, 31-24 Rehab. Ctr., Los Angeles, CA 90095-1794)

Over the years, many models have been proposed to capture mucosal wave propagation along the medial surface of the vocal folds. However, historically, little quantitative data have existed with which to evaluate such models. Indeed, standard methods of laryngeal examination have imaged the medial surface vibrations of the vocal folds very poorly. However, new methods of imaging the medial surface dynamics of the vocal folds have recently been reported [D. A. Berry, D. W. Montequin, and N. Tayama, J. Acoust. Soc. Am. **110**, 2539–2547 (2001)]. These data are used to evaluate a variety of models of mucosal wave propagation. [Work supported by NIDCD Grant No. R01 DC03072.]

5pSC26. Histological investigation of the supra-glottal structures in humans for understanding abnormal phonation. Miwako Kimura (ENT Dept., Takeda General Hospital, Yamagamachi 2-27, Aiduwakamatu City, Fukushima, Japan, mkimu-tky@umin.ac.jp), Ken-Ichi Sakakibara, (NTT Commun., Sci. Labs., Japan), Hiroshi Imagawa (Univ. of Tokyo, Tokyo, Japan), Roger Chan (Purdue Univ., West Lafayette, IN 47907), Seijii Niimi (Intl. Univ. of Health and Welfare, Japan), and Niro Tayama (Intl. Medical Ctr. of Japan, Tokyo, Japan)

Phonation is the vocal fold vibration on normal voice. But sometimes we can observe the other phonation styles like as the pressed voice or some throat singings like as ''kargyraa'' or ''drone'' in Khoomei in Mongolian music. Also, clinically, we know that some patients who have the wide glottal slit in phonation because of the recurrence nerve palsy or after partial laryngectomy, could make the ''supra-glottal phonation.'' The ''supra-glottal phonation'' would be made from the vibration of ''supra-glottal structures'' such as the false vocal folds, the arytenoids and the epiglottis, etc. Endoscopic examination suggests the existence of some contractile functions in supra-glottal space. However, these phonation systems have not been clear to explain their neuromuscular mechanism in histology. This study aimed to find the basis for making the supra-glottal phonation from the points of view of the histological structures. We tried to investigate if there were any muscles that could contract the supra-glottal structures. The samples are the excised larynx of human beings. They were fixed by formalin after excision. We observed their macroscopic anatomy, and also with the microscopic observation their histological preparations after the process of the embedding in paraffin, slicing for the preparation and HE (hematoxylin-eosin) staining.

5pSC27. Jitter, shimmer, and noise in pathological voice quality perception. Jody Kreiman, Bruce R. Gerratt, and Brian Gabelman (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab Ctr., Los Angeles, CA 90095-1794)

Although jitter, shimmer, and noise characterize all voice signals, their perceptual importance has not been established psychoacoustically. To determine which of these acoustic attributes is important in listeners' perceptions of the inharmonic component in pathologic voices, copies of natural pathological voices were synthesized parametrically using seven different models of the inharmonic part of the voice source: jitter only, shimmer only, noise only, jitter plus shimmer, jitter plus noise, shimmer plus noise, and jitter plus shimmer plus noise. Listeners then compared pairs of these stimuli and judged whether they were the same or different. Listeners also compared synthetic and natural stimuli, to determine the extent to which different aspects of the inharmonic part of the source improved or worsened the quality of the match to the original natural voice sample. Preliminary results suggest that jitter and shimmer may not be perceptually important features of most pathological voices, despite their long history as acoustic measures of voice. In contrast, it appears that listeners perceive spectrally shaped additive noise as the critical inharmonic acoustic element contributing to pathologic voice quality. [Research supported by NIDCD.]

FRIDAY AFTERNOON, 6 DECEMBER 2002 CORAL GARDEN 2 AND 3, 1:00 TO 2:15 P.M.

Session 5pSPa

Signal Processing in Acoustics: Time Reversal Signal Processing

Philippe Roux, Chair

Marine Physical Laboratory, Scripps Institution of Oceanography, 88203 Shellback Way, La Jolla, California 92093-0328

Contributed Papers

1:00

5pSPa1. Real time inverse filter focusing by iterative time reversal. Gabriel Montaldo, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, Universit Paris VII/ESPCI, CNRS UMR 7587, Paris, France)

In order to achieve an optimal focusing through heterogeneous media, we need to build an inverse filter of the propagation operator. Time-reversal achieves such an inverse filter in nondissipative media. However, as soon as losses appear in the medium, time-reversal is no more equivalent to an inverse filter. Consequently, it does no more produce the optimal focusing and beam degradations may appear. In such cases, previous works have shown that the optimal focusing can be recovered by using a spatiotemporal inverse filter technique. This process requires the presence of a set of receivers inside the medium but allows us to reach the optimal focusing even in extreme situations such as ultrasonic focusing through a human skull. To date, it remained time consuming and implied fastidious numerical calculations. We propose here a new way to process this inverse filtering in real time and without any calculation. The new process is based on iterative time-reversal processing. Contrary to the classical inversion, this iteration does not require any computation as it achieves the inverse filter experimentally using waves propagation instead of computational power. The feasibility of this iterative technique is theoretically and experimentally demonstrated for ultrasound applications and its potential for telecommunications is discussed.

1:15

5pSPa2. Broadband-DORT applied to bistatic active sonar. David M. Fromm and Charles F. Gaumond (Acoust. Div. Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5320)

The decomposition of the time reversal operator (DORT) is a single frequency method of spatially isolating scatterers with a multiple-source/multiple-receiver system [C. Prada, S. Manneville, D. Spoliansky, and M. Fink, J. Acoust. Soc. Am. **99**, 2067–2076 (1996)]. This technique has 8 been extended into a coherent, broadband method. Broadband DORT has also been shown to isolate resolvable scatterers at various depths and ranges in a bistatic, active sonar in shallow water. Several limitations of the method are demonstrated through the use of numerical simulations and backpropagations. One limitation is the resolvability of a system in shallow water. The dependence of resolution on environmental parameters, e.g., signal-to-noise ratio, water depth, bottom parameters and sound speed profile—and system parameters—e.g., frequency content and source/receiver configuration—is shown by forming images of a single scatterer, two scatterers and many scatterers. The limits of continuity over frequency of the broadband singular vectors and values caused by frequency fading are also shown through data visualization techniques. [Work supported by ONR.]

1:30

5pSPa3. Detection improvement using multiple time-reversal guide-sources in shallow water. Charles F. Gaumond and David M. Fromm (Acoust. Div. Naval Res. Lab., Washington, DC 20375-5320)

Detection in a monostatic broadband, active sonar system in shallow water is degraded by propagation induced spreading. The method of multiple guide sources has been presented as a method to mitigate this degradation. This paper shows the detection improvement from the use of the separable kernel receiver (SKR) that is based on a statistical behavior of the ensemble of multiple guide-source signals, which provide several one-way propagation signals from guide sources at various ranges at a single depth. A measure of information is derived to quantify the variability of these guide-source signals. The improvement of detection is shown as a function of range and depth for several sets of system and environmental parameters. In particular, two environments are shown. One is the case of farther ranges and lower frequencies in shallow water where the propagation causes dispersion. The other is the case of shorter ranges and higher frequencies in shallow water where the propagation effects are predominantly multipath. [Work supported by ONR.]

1:45

5pSPa4. Combining and comparing time reversal processing and adaptive channel equalization for communication sequences. Geoffrey F. Edelmann, Seongil Kim, W. S. Hodgkiss, W. A. Kuperman, H. C. Song (Scripps Inst. of Oceanogr., 9500 Gilman Dr., La Jolla, CA 92093-0238), and Tuncay Akal (SACLANT Undersea Res. Ctr., La Spezia, Italy)

Time reversal (TR) backpropagates a signal to its origin despite the complexity of the propagation channel. The application of TR to acoustic communications previously has been demonstrated. Here, we discuss the advantages of an implementation of the TR process cascaded with adaptive channel equalization. This combination makes spatial diversity unnecessary in the receiver structure. In addition, comparisons will be drawn between single source and TR transmissions in combination with adaptive channel equalization. [Work supported by ONR.]

2:00

5pSPa5. Multiple focusing with adaptive time-reversal mirror. J. S. Kim and K. C. Shin (Div. of Ocean Systems Eng., Korea Maritime Univ., Busan 606-791, South Korea, jskim@kmaritime.ac.kr)

Recently, adaptivity was introduced to time-reversal mirror (TRM) to steer the nulls, and referred to as an adaptive time-reversal mirror (ATRM) [J. S. Kim, H. C. Song, and W. A. Kuperman, J. Acoust. Soc. Am. **109**, 1817–1825 (2001)]. In this study, ATRM is extended to simultaneous multiple focusing in a free space as well as in a wave guide. The multiple focusing is achieved by imposing a set of constraints in the formulation for finding the weight vectors. The algorithm is applied to the long-range underwater acoustic communication to show that the self-equalizations at multiple receiving locations are achieved. Also the size of resolution cell, where the distortionless response (or, self-equalization) is approximately achieved, is discussed. [Work supported by Grant No. R05-2002-000-01097-0 from the basic Research Program of the Korea Science and Engineering Foundation.]

FRIDAY AFTERNOON, 6 DECEMBER 2002 CORAL GARDEN 2 AND 3, 2:30 TO 3:30 P.M.

Session 5pSPb

Signal Processing in Acoustics: Digital Signal Processing Techniques

Alan W. Meyer, Chair
Lawrence Livermore National Laboratory, L-154, 7000 East Avenue, Livermore, California 94550

Contributed Papers

2:30

5pSPb1. Algorithms and architectures for highly efficient digital beamforming of acoustic arrays. Harvey Woodsum, William Hogan, and John Gaidos (Sonetech Corp., 10 Commerce Park North, Unit 1, Bedford, NH 03110)

The creation of broad bandwidth, directional transmit beams with an ($N\times M$ element) acoustic transmitting array normally requires generation NM time delayed versions of the signal to be transmitted, with precise control required for each signal channel. Similarly, for generation of a directional receive beam, NM time delays are needed to be applied to the receive channels, which are then weighted and summed to create the beam output time series. Recently, an architecture has been developed which allows efficient beamforming with a much smaller number of delayed channels, this being on the order of the square root of the number of elements, for planar geometries, and on the order of the cube root of the total number of elements for volumetric arrays, thus achieving an enormous savings in beamforming complexity. With this approach, an offline

genetic algorithm is utilized to select time delays and a one to many mapping function, to allow minimization of the number of delays to be applied, according to a fitness criterion which optimizes the process toward the ideal case. An actual hardware implementation in which this algorithm and architecture has been successfully utilized is described.

2:45

5pSPb2. A novel beamformer array design for noise source location from intermediate measurement distances. Jorgen Hald and Jacob J. Christensen (Bruel & Kjaer SVM A/S, Skodsborgvej 307, DK-2850 Naerum, Denmark)

Very regular beamformer array geometries will have grating lobes in their directivity pattern when applied above a certain frequency. The grating lobes can be suppressed by removing periodicities, typically by the use of random arrays or spiral arrays. Large irregular arrays are, however, difficult to build due to their complicated support structure and cabling. The present paper describes a novel array design, which maintains the low grating lobe level of irregular arrays, but which has a regularity that allows a much simpler support structure and cabling. The performance is compared with that of comparable regular and irregular arrays, and verified through practical measurements. Examples of application to the localization of source of wind noise and high-frequency engine noise will be presented.

3:00

5pSPb3. Random sampling applied to array beamforming: The aggregate beamformer. David I. Havelock (Natl. Res. Council, 1200 Montreal Rd., Ottawa, ON K1A 0R6, Canada)

It is shown how beamforming can be accomplished by randomly sampling array elements and time-aligning the data samples. The time-aligned samples are referred to as the ''aggregate'' signal. Time-alignment is done using conventional beamforming delays but data values are not summed. Beamforming weights (array shading) can be applied by modifying the sampling probability for each sensor. Simple methods for reconstruction of the original signal from the aggregate signal are discussed. The reconstructed signal of this aggregate beamformer has the same directional response as a conventional beamformer except that off-beam signals are transformed into white noise. This noise can be reduced as required by adjusting the sampling rate of the aggregate signal. Higher beam-steering resolution than that of conventional beamforming is achieved without the need for data interpolation. The hardware and computational complexity of the aggregate beamformer does not increase as the number of array elements is increased. Aggregate beamforming offers a high degree of integration for arrays with many elements, such as 3-D arrays.

3:15

5pSPb4. Improving noise source location using the unwrapped phase method. Yongxin Wang and Corinne Darvennes (Dept. of Mech. Eng., Tennessee Technolog. Univ., Box 5014, Cookeville, TN 38505, Yongxin@tntech.edu)

Time Delay of Arrival (TDOA) is the key to using a microphone array for sound source location. The Unwrapped Phase (UP) Method has been used to estimate the TDOA, which is proportional to the slope of the phase of the cross-power spectrum plotted versus frequency. However, reflecting surfaces cause ''jumps'' in the phase plot. A Zero-Magnitude of Transfer Function (ZMTF) criterion is proposed to automatically pick up the ''jumps'' to better estimate the TDOA. To verify this, tests were performed in three rooms with different acoustical properties. The sound source consisted of a speaker broadcasting a computer-generated broadband random signal. ZMTF works well in an ordinary room with small reverberation and yields accurate TDOA results. An average error of 0.2% is obtained. In a medium-sized live room, ZMTF identifies ''jumps,'' and an average error of 1.6% is obtained. In a large reverberation room, however, ZMTF only marginally improved the UP method. The norm of residues and Delay Error were found helpful to evaluate the best results.

FRIDAY AFTERNOON, 6 DECEMBER 2002 GRAND CORAL 1, 1:00 TO 5:35 P.M.

Session 5pUW

Underwater Acoustics and Acoustical Oceanography: Littoral Environmental Variability and Its Acoustic Effects III

Timothy F. Duda, Cochair

Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Bigelow Laboratory 202, MS 11, Woods Hole, Massachusetts 02543-1053

Bruce H. Pasewark, Cochair

Acoustics Division, Naval Research Laboratory, Code 7127, 4555 Overlook Avenue, Washington, D.C. 20375-5350

Invited Paper

1:00

5pUW1. Measurement and analysis of acoustic intensity fluctuations induced by South China Sea internal solitons. C.-S. Chiu, S. Ramp, C. Miller (Naval Postgrad. School, Monterey, CA 93943), J. Lynch, T. Duda, A. Newhall, K. von der Heydt, and J. Kemp (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

In May 2001, a suite of acoustic and oceanographic sensors was deployed by a team of US, Taiwan, and Singapore scientists in the NE South China Sea to study the effects of ocean variability on low-frequency sound propagation in a shelfbreak environment. The primary acoustic receiver was an L-shaped hydrophone array moored on the continental shelf that monitored a variety of signals transmitted parallel to and across the shelfbreak by fixed and towed sources. This talk discusses and contrasts the amplitude fluctuations in the signals measured by the vertical segment of the listening array in two different days, one with the passage of several huge solitons that depressed the shallow isotherms to the sea bottom, and one with a much less energetic soliton field. In addition to

exhibiting large temporal changes, the acoustic data also show a much more vertically diffused sound intensity field as the huge solitons occupied and passed through the transmission path. Using a space-time continuous empirical sound speed model based on the moored temperature records and geoacoustic parameters estimated using the close-range towed source signals, the observed acoustic intensity fluctuations are explained using coupled-mode physics. [Work supported by ONR under the ASIAEX program.]

Contributed Papers

1:20

5pUW2. Nonlinear internal waves in the South China Sea during ASIAEX. Antony Liu (NASA Goddard Space Flight Ctr., Greenbelt, MD 20771), David Tang (Natl. Taiwan Univ., Taipei, Taiwan), and Steve Ramp (Naval Postgrad. School, Monterey, CA)

Internal wave distribution maps have been compiled from more than 100 ERS-1/2, RADARSAT and Space Shuttle SAR images in the South China Sea (SCS) from 1993 to 2000. Based on these distribution maps, most of the internal waves in the northeast part of the SCS were propagating westward. The wave crest can be as long as 200 km with an amplitude of 100 m, due to strong current from the Kuroshio branching out into the SCS. In the recent Asian Seas International Acoustics Experiment (ASIAEX), moorings have been deployed in April 2000 and May 2001. Simultaneous RADARSAT ScanSAR images have been collected during the field test to integrate with the model and the *in-situ* measurements in the SCS. During ASIAEX in May 2001, many large internal waves were observed at the test area and were the major features for acoustic volume interaction. The evolution and dissipation of huge internal waves on the shelf break, mode-two waves, and wave–wave interaction are very important issues for acoustic propagation. Numerical simulations have been performed to produce the wave evolution on the continental shelf and compared with the field measurements. The implication of internal wave effects on acoustic propagation will also be discussed. [Work supported by ONR.]

1:35

5pUW3. Tracking the generation sites and packet variability of internal solitons in the South China Sea. Steven R. Ramp, Ching Sang Chiu, Frederick L. Bahr (Dept. of Oceanogr., Naval Postgrad. School, Monterey, CA 93943), James Lynch, Tim Duda (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), David Tang (Natl. Taiwan Univ., Taipei, ROC), and Anthony K. Liu (NASA Goddard Space Flight Ctr., Greenbelt, MD 20771)

An array of seven moorings were deployed across the Chinese continental shelf and slope near Dongsha Island as part of the ONR-sponsored Asian Seas International Acoustics Experiment (ASIAEX). The moorings spanned 82- to 350-m depth and sampled temperature, salinity, and velocity at 1- or 2-min intervals from 21 April to 19 May 2001. The most energetic features observed in the data set were the highly nonlinear internal solitary waves. This paper attempts to identify the generation sites for these waves and determine their predictability. A statistical analysis was performed on the wave orbital velocities to determine their direction of propagation, and inversely, their sources. The bimodal directional histograms point to two generation sites: The largest waves came from around the Batan Islands in the Luzon Strait, and the smaller ones from near Dong-sha Island. The large wave arrivals were semidiurnal with a very strong diurnal inequality, such that the strongest waves arrived every 24 h. The wave packets arrived in two groups separated by a 5-day period in between when weak or no waves arrived. The wave groups had an unusual and counterintuitive relationship to the barotropic tide, and two hypotheses will be advanced to explain this phenomenon.

1:50

5pUW4. Nonlinear internal wave energy dissipation and evolution on the South China Sea Slope. Marshall H. Orr and Peter C. Mignerey (The Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375, orr@wave.nrl.navy.mil)

Nonlinear internal waves have been observed propagating up the slope of the South China Sea during the recent AsiaEX experiment. Energy dissipation rates have been extracted. The location of the initiation of the depression to elevation conversion has been identified. Scaling parameters have been extracted and used to initialize a two-layer evolution models simulation. The calculations are compared to the *in situ* observations. [Work supported by the ONR Ocean Acoustic Program and ONRs NRL base funding.]

2:05

5pUW5. Correlation of acoustic scattering variables and physical oceanographic parameter time series in the 2001 ASIAEX South China Sea experiment. T. F. Duda, J. F. Lynch, A. E. Newhall (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), C.-S. Chiu, S. R. Ramp (Naval Postgrad. School, Monterey, CA 93943), and C.-F. Chen (Natl. Taiwan Univ., Taipei, Taiwan, ROC)

During May 2001, 210- to 550-Hz signals were transmitted from five moored sources to moored array receivers in 120-m water on the continental slope in the northern portion of the South China Sea. Transmissions were made for three weeks over distances of 20 km along slope and 32 km across slope, in an area with highly irregular topography. The oceanography in the transmission area exhibited large variations, with the largest being due to the nonlinear internal wave field. In this study, we examine the cross-correlation of the measured oceanographic variability with the measured acoustic variability, in particular with the travel time wander and spread and the acoustic intensity. We emphasize the variability within specific oceanographic bands: subtidal, tidal, and high frequency. [Research sponsored by ONR.]

2:20

5pUW6. Environmental correlate of acoustic impulse response and arrivals with solibore propagation in ASIAEX SCS experiment. Chi-Fang Chen, Ying-Tsong Lin, Wen-Fei Lin, David Tang (Natl. Taiwan Univ., No. 1 Roosevelt Rd., Sec. #4, Taipei, Taiwan 106, ROC), Yiing-Jang Yang (Naval Acad., Kaohsiung, Taiwan, ROC), James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA), Chin-Sang Chiu, and Steven Ramp (Naval Postgrad. School, Monterey, CA)

An ASIAEX South China Sea experiment was conducted in April–May 2001. The experiment site is characterized as an active internal tide/solibore propagation region along the northwestern shelf break of the South China Sea. On May 4–5, one package of solibore propagated through the sources and receiver arrays as clearly seen from the satellite image and the temperature time series recorded in the environmental moorings. The propagation path of the solibore is reconstructed using the data from the environmental moorings. In this paper we also present results of the acoustic impulse responses and the arrivals of the same time as the solibore passing through. Correlations of the acoustic impulse responses and the arrivals with the solibore propagation show the solibore's

effect on the acoustic propagation when the solibore enters the acoustic propagation paths. [Work supported by the National Science Council of Republic of China for Taiwanese authors. Work supported by ONR, for the U.S. authors.]

2:35

5pUW7. Soliton-induced matched-field array-gain variability in the South China Sea. Peter C. Mignerey, Marshall H. Orr, Bruce H. Pasewark (Acoust. Div. 7120, Naval Res. Lab., Washington, DC 20375), Timothy F. Duda, James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Ching-Sang Chiu (Naval Postgrad. School, Monterey, CA 93943)

Large amplitude nonlinear internal-wave packets propagated across the acoustic signal propagation path at the ASIAEX experimental site. The wave packets perturbed the sound-speed field and caused temporal variability in the phase of narrow-band (224, 300, 400, and 500 Hz) acoustic signals propagating along upslope and cross slope paths. The acoustic signals were received on a vertical hydrophone array. The signal phase variability caused degradation and temporal variability in the signal gain of a matched-field processor applied to the data. The temporal variability of the sound-speed field along the signal propagation paths has been simulated using field CTD and flow visualization measurements to initialize the sound-speed field. A matched field processor has been applied to the modeled acoustic field. The calculated gain degradation will be compared to field data and inferences will be drawn concerning our ability to simulate the sound speed field variability. [Work supported by ONR.]

2:50

5pUW8. Acoustic beam broadening by solitary bores during ASIAEx 2001 (South China Sea). Stephen N. Wolf, Bruce H. Pasewark, Marshall H. Orr (Acoust. Div., Naval Res. Lab., Washington, DC 20375), James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Ching-Sang Chiu (Naval Postgrad. School, Monterey, CA 93943)

Previous theoretical and modeling studies have predicted that acoustic interaction with anisotropic internal wave fields will introduce mode-order-dependent horizontal refraction, thereby limiting the resolution of horizontal receiving arrays. An experiment to observe this phenomenon was performed on the continental shelf southwest of Taiwan. A nonlinear, bottom-laid horizontal array received signals over an 18-day period from two sound sources (300 and 500 Hz) located 18 km to the northeast. Water depth along the propagation path was in the range 110–180 m. The prevailing downward-refracting sound speed profile was modulated by internal waves. At semidiurnal intervals strong internal bores originating in the Luzon Strait entered the propagation path at the source position, introducing strongly anisotropic sound speed perturbations. The wavecrests of the internal bores' surface expression were oriented at typical angles 20–30 deg with respect to the acoustic propagation path. Aperture-limited signal beam widths (FWHM), nominally 1.5 deg, were broadened while the semidiurnal bores were in the acoustic propagation path. Details of the acoustic and oceanographic measurements and model calculations of beam broadening by the sound speed perturbations will be presented. [Work supported by ONR.]

3:05–3:20 Break

3:20

5pUW9. Remote estimation of seabed attenuation using a chirp sonar during ASIAEX. Steven G. Schock (Dept. of Ocean Eng., Florida Atlantic Univ., Boca Raton, FL 33431)

Acoustic remote sensing of the seabed using a chirp sonar generates estimates of the attenuation coefficient of sediments in the South and East China Seas. The chirp sonar transmits FM pulses over the band of 1 to 15 kHz at normal incidence to the seabed. After cross correlating the acoustic returns with the transmission wave form, a wideband reflection profile is generated showing sediment layering. The time series data is filtered into 11 pass bands to generate 11 images with center frequencies of 1 kHz spacing starting at 2 kHz and ending at 12 kHz. The attenuation coefficient of the surficial layer is estimated by dividing the average reflection amplitude of the sediment layer interface echo from the base of the surficial layer by the average amplitude of the sediment-water interface echo for each bandpass image. The amplitude ratios are plotted as a function of frequency. The slope of a line fitted to the frequency function in a least squares sense is called the estimated attenuation coefficient. The estimates are useful for sediment classification and aocustic propagation modeling.

3:35

5pUW10. Acoustic signal spatial and temporal correlation variability in the ASIAEx 2001 South China Sea experiment. Bruce H. Pasewark, Stephen N. Wolf, Marshall H. Orr (Naval Res. Lab., Code 7120, Washington, DC 20375-5350), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Broadband acoustic signal intensity fluctuations, temporal correlations, and horizontal spatial correlations have been extracted from a horizontal array data set. The relationship between the acoustic signal variability and the fluid processes that perturb the sound-speed field will be discussed. The 2.048-s linear frequency-modulated signals were received on a 450-m horizontal array. The signals were centered at 300 and 500 Hz and had respective bandwidths of 60 and 90 Hz. The signal intensity was found to fluctuate by 10 to 20 dB as a function of time and space. Horizontal correlation length varied over time from 5 to 100 acoustic wavelengths and the signal correlation time ranged from 1 to 50 min. The data were acquired in May 2001 during the Asian Seas International Acoustics Experiment (AsiaEx). The 18-km acoustic propagation path lay along a 125-m isobath at the South China Sea shelf break. The experiment site was oceanographically dynamic, with energetic internal gravity wave fields and currents that strongly impacted the acoustic propagation. [Work supported by ONR.]

3:50

5pUW11. Ambient noise level on vertical array in ASIAEX South China Sea experiment. Ruey-Chang Wei and Yung-Chien Chen (Inst. of Undersea Technol., Natl. Sun Yat-Sen Univ., Kaohsiung City, Taiwan)

Previous literatures show that ambient noise in deep water should decrease with depth, but would it be the same in the shallow? The ASIAEX 2001 South China Sea experiment is on the continental shelf-break, with water depth from 100 m to 400 m, so the depth effect of ambient noise in shallow water can be explored. In this study, acoustic signal of ambient noise was collected on vertical line array, which has 16 elements, from 700 to 1200 Hz. From the results of this study, ambient noise has insignificant change with depth, however it is much affected by the variation of temperature, frequency, and occurrence of internal wave. Higher frequency tends to have lower SPL and the decay is of 5 dB/octave approximately. The temperature profile controls the vertical distribution of ambient noise level, so the internal wave effect can be significant. Moreover, the serious weather condition, like the typhoon during the experiment, can greatly increase the ambient noise level and change its distribution.

4:05

5pUW12. Statistical analysis of low-frequency noise recorded in ASIAEX by a PANDA system. John R. Potter, Koay T. Beng, and Venugopalan Pallayil (ARL, TMSI, NUS, 12a Kent Ridge Rd., Singapore 119223)

As part of the ASIAEX experiment, the Acoustic Research Laboratory in the Tropical Marine Science Institute at the National University of Singapore deployed a Pop-up Ambient Noise Data Acquisition (PANDA)

system. The PANDA was recovered 18 days later with over 9 days of continuous data recorded from a single hydrophone at 2 kSa/s. The data show the various sources that were deployed as part of the experiment, but also provide interesting statistical information on low-frequency ambient noise in the region and the passage of numerous ships. This deployment was in a heavy shipping traffic area, hostile both in terms of potential snagging by fishing activity and in terms of the high levels of noise encountered, both of which are of interest for the deployment and successful use of autonomous acoustic systems in busy littoral waters. We present some statistical results from the 3+ GByte of data. [Work supported by the Defence Science and Technology Agency, Singapore and the US ONR.]

4:20

5pUW13. Applications of source localization by waveguide invariant theory during the ASIAEX sea tests. Altan Turgut, Stephen Wolf, Bruce Pasewark, Marshall Orr (Naval Res. Lab., Acoust. Div., Washington, DC 20375), and James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Applications of source localization in range and azimuth with minimum environmental information have been investigated by using a sample data set from the ASIAEX vertical and horizontal arrays. Source localization by waveguide invariant theory has been successfully demonstrated at the ASIAEX experimental site by using broadband (50–200 Hz and 240–260 Hz) signals transmitted from a towed J-15 source. Range and bearing of both J-15 transducer and other surface ships in the vicinity have been successfully identified using a reference signal from the J-15 source at a known range. Accuracy of the localization has been also investigated for both adiabatic and strongly range-dependent environments. It has been shown that, when nonlinear internal waves are not present, the assumption of adiabatic propagation is sufficient for the accurate localization of sources from 0–10 km ranges at the experimental site. [Work supported by ONR.]

4:35

5pUW14. Three-dimensional acoustic effects in the ASIAEX SCS experiment. Chi-Fang Chen, Li-Wen Hsieh (Natl. Taiwan Univ., No. 1 Roosevelt Rd. Sec #4, Taipei, Taiwan 106, ROC), Roger Oba, and Steven Finette (Naval Res. Lab., Washington, DC)

A three-dimensional ocean environment is generated using measured ocean data from the ASIAEX South China Sea experiment. The experiment site is characterized as an active internal tide/solibore propagation region along the Northwestern shelf break of the South China Sea. The three-dimensional acoustic effects are studied using a wide-angle version of the parabolic equation code FOR3D. Mode coupling matrices and horizontal refractions are shown to demonstrate the 3-D effects. Variations in topography of the shelf break and in the water column due to the internal solibore propagation cause significant mode coupling and 3-D effects in the acoustic field. [Work supported by National Science Council of Republic of China for first and second authors. Work supported by ONR for the third and fourth authors.]

4:50

5pUW15. Initial reference level of the absolute reverberation intensity in shallow water. Ji-Xun Zhou and Xue-Zhen Zhang (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332 and Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, jixun.zhou@me.gatech.edu)

A quality database of reverberation is absolutely essential if one is to understand the shallow water reverberation problem. To develop a practical prediction model, invert seabottom scattering strength (and acoustic parameters), or to develop a seabottom scattering model, all require a reliable reverberation database either for comparison or justification. It is relatively easy to get nonabsolute intensity curves from at-sea measurements. However, to get wideband absolute reverberation level (ARL) for both short and long range is a very delicate task that can be prone to errors. It often requires a measurement system that can cover a large dynamic range up to 150 dB and avoid signal overflow and saturation. It may need to precisely measure source level (SL) and calibrate measurement system. In this paper, a traditional method to get ARL frequently used by Chinese scientists is introduced by showing the results from the recent ASIAEX and 1996 Yellow Sea experiments. The ARL (relative to SL) at 1 second (after an explosive source is detonated) is defined as the initial reference level (IRL). The IRL from the ASIAEX and Yellow Sea experiments as a function of frequency will be given to compare with the data from other ''natural laboratories'' with different sediments, and with our theoretical prediction. The advantage of introducing a concept of IRL will be discussed. [Work supported by ONR and CAS.]

5:05

5pUW16. Observations and modeling of acoustic intensity fluctuations seen during the 1996 summer New England shelfbreak PRIMER experiment. A. Fredricks, J. Colosi, J. Lynch, G. Gawarkiewicz, T. Duda (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), C.-S. Chiu (Naval Postgrad. School, Monterey, CA 93943), and P. Abbott (Ocean Acoust. Services and Instrumentation Systems, Inc., Lexington, MA 02421)

In this paper we discuss the intensity fluctuation characteristics of 400-Hz acoustic transmissions made in the Mid-Atlantic Bight during the summer of 1996. The transmissions traversed variable topography, the dynamic shelfbreak front, and strong nonlinear internal waves, all of which affected the acoustic propagation. In studying the intensity fluctuations, we look both at data from the PRIMER vertical line array receiver and at output from a broadband parabolic equation model intended to simulate the experimental conditions with high fidelity. The statistics of the fluctuations, the physical causes of the fluctuations, and the correlation with the oceanographic forcing in different frequency bands (subtidal, tidal, and high frequency) all are examined. [Work supported by ONR.]

5:20

5pUW17. Internal gravity waves and horizontal refraction of sound in the ocean. Oleg A. Godin (CIRES, Univ. of Colorado and NOAA/Environ. Technol. Lab., Boulder, CO 80305), Alexander G. Voronovich, and Valery U. Zavorotny (NOAA/Environ. Technol. Lab., Boulder, CO 80305)

Horizontal refraction, i.e., the deviation of acoustic propagation paths from the source–receiver vertical plane, is usually weak in the ocean. In this paper, a perturbation approach [O. A. Godin, J. Comput. Acoust. **10**, 123–151 (2002)] to describe horizontal refraction within the ray and the ''vertical modes—horizontal rays'' theories is extended to media with a combination of stochastic and deterministic inhomogeneities and subsequently utilized to quantify 3-D acoustic effects due to internal gravity waves. Within the perturbation approach, statistical moments of adiabatic mode amplitude and phase and ray travel time are given explicitly by integrals of respective moments of cross-range derivatives of environmental parameters along the trajectory of the sound wave in an unperturbed (range-dependent) medium. Specific scenarios considered simulate conditions of recent propagation experiments. In shallow water, internal wave soliton-induced horizontal refraction can result in travel time biases $O(10\ \text{ms})$ at propagation ranges as small as 10 km. In deep water, at ranges about 1000 km horizontal refraction due to internal waves with the Garrett–Munk spectrum leads to acoustic travel times being $O(10\ \text{ms})$ less than in 2-D simulations. Without the internal waves, the travel time bias is found to be smaller by an order of magnitude. [Work supported by ONR.]

Session 5eID

Interdisciplinary: Tutorial on Industrial Noise Control

5eID1. Tutorial on industrial noise control. Samir N. Y. Gerges (Mech. Eng. Dept. Federal Univ. of Santa Catarina, Cx. P. 476 Florianopolis, SC Brazil CEP: 88040-900) and Brandon Tinianov (Johns Manville, 10100 W. Ute Ave., P. O. Box 80162, Littleton, CO 80162).

This tutorial is intended to provide a practical overview of applied acoustics and industrial noise control, including the fundamentals of applied acoustics, industrial noise control and case histories. The following topics will be covered: Sound radiation, sound transmission, outdoor sound propagation and indoor sound fields, sound absorption materials, machinery noise and its control (fans, motors, compressors, air jets, pneumatic tools) and the design of sound enclosures.