

CODEN: JASMAN

The Journal of the Acoustical Society of America

ISSN: 0001-4966

Vol. 113, No. 4, Pt. 1

April 2003

ACOUSTICAL NEWS—USA	1767
USA Meetings Calendar	1767
ACOUSTICAL NEWS—INTERNATIONAL	1775
International Meetings Calendar	1775
ABSTRACTS FROM ACOUSTICS RESEARCH LETTERS ONLINE	1779
REVIEWS OF ACOUSTICAL PATENTS	1781

GENERAL LINEAR ACOUSTICS [20]

Acoustic and mechanical response of reservoir rocks under variable saturation and effective pressure	C. L. Ravazzoli, J. E. Santos, J. M. Carcione	1801
Traction-free vibrations of finite trigonal elastic cylinders	Paul R. Heyliger, Ward L. Johnson	1812
Symmetrization of Ritz approximation functions for vibrational analysis of trigonal cylinders	Ward L. Johnson, Paul R. Heyliger	1826
Investigation of the vibrational modes of edge-constrained fibrous samples placed in a standing wave tube	Bryan H. Song, J. Stuart Bolton	1833
Experimental validation of two elastodynamic models for the wave field generated by ultrasonic transducers	Maarten C. M. Bakker, Martin D. Verweij	1850

NONLINEAR ACOUSTICS [25]

Modeling of finite amplitude acoustic waves in closed cavities using the Galerkin method	Robert R. Erickson, Ben T. Zinn	1863
--	---------------------------------	------

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

Acoustic attenuation in gas mixtures with nitrogen: Experimental data and calculations	Sally G. Ejakov, Scott Phillips, Yefim Dain, Richard M. Lueptow, Jacobus H. Visser	1871
Scattering of the fundamental shear horizontal mode from steps and notches in plates	A. Demma, P. Cawley, M. Lowe	1880
Acoustic streaming measurements in annular thermoacoustic engines	Stéphane Job, Vitalyi Gusev, Pierrick Lotton, Michel Bruneau	1892

TRANSDUCTION [38]

Analysis of axially polarized piezoelectric cylinders with arbitrary boundary conditions on flat surfaces	D. D. Ebenezer, R. Ramesh	1900
---	---------------------------	------

(Continued)

CONTENTS—Continued from preceding page

STRUCTURAL ACOUSTICS AND VIBRATION [40]

On the dynamic stiffness of preloaded vibration isolators in the audible frequency range: Modeling and experiments	Leif Kari	1909
A substructure approach for the midfrequency vibration of stochastic systems	Abhijit Sarkar, Roger Ghanem	1922
Modal characteristics of in-plane vibration of circular plates clamped at the outer edge	N. H. Farag, J. Pan	1935
The wave transmission coefficients and coupling loss factors of point connected structures	R. S. Langley, P. J. Shorter	1947
Investigation into the validity of two-dimensional models for sound radiation from waves in rails	D. J. Thompson, C. J. C. Jones, N. Turner	1965

NOISE: ITS EFFECTS AND CONTROL [50]

Helmholtz resonator with extended neck	Ahmet Selamet, Iljae Lee	1975
Noise reduction in a launch vehicle fairing using actively tuned loudspeakers	Jonathan D. Kemp, Robert L. Clark	1986

ARCHITECTURAL ACOUSTICS [55]

Experimental investigation of spatial correlation in broadband reverberant sound fields	Ingyu Chun, Boaz Rafaely, Phillip Joseph	1995
Speech transmission index from running speech: A neural network approach	F. F. Li, T. J. Cox	1999

ACOUSTICAL MEASUREMENTS AND INSTRUMENTATION [58]

Correction of ultrasonic wave aberration with a time delay and amplitude filter	Svein-Erik Måsøy, Tonni F. Johansen, Bjørn Angelsen	2009
---	---	------

PHYSIOLOGICAL ACOUSTICS [64]

Modeling individual differences in ferret external ear transfer functions	Jan W. H. Schnupp, John Booth, Andrew J. King	2021
Transient emission suppression tuning curve attributes in relation to psychoacoustic threshold	Erika M. Zettner, Richard C. Folsom	2031

PSYCHOLOGICAL ACOUSTICS [66]

Modulation masking in cochlear implant listeners: envelope versus tonotopic components	Monita Chatterjee	2042
A practical method of predicting the loudness of complex electrical stimuli	Colette M. McKay, Katherine R. Henshall, Rebecca J. Farrell, Hugh J. McDermott	2054
Speech recognition under conditions of frequency-place compression and expansion	Deniz Baskent, Robert V. Shannon	2064

SPEECH PRODUCTION [70]

Numerical simulation of self-sustained oscillation of a voice-producing element based on Navier–Stokes equations and the finite element method	Martinus P. de Vries, Marc C. Hamburg, Harm K. Schutte, Gijsbertus J. Verkerke, Arthur E. P. Veldman	2077
--	--	------

SPEECH PERCEPTION [71]

Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing	Judy R. Dubno, Amy R. Horwitz, Jayne B. Ahlstrom	2084
--	--	------

CONTENTS—Continued from preceding page

SPEECH PROCESSING AND COMMUNICATION SYSTEMS [72]

- | | | |
|---|-------------------------|------|
| Segmental intelligibility of four currently used text-to-speech synthesis methods | Horabail S. Venkatagiri | 2095 |
|---|-------------------------|------|

MUSIC AND MUSICAL INSTRUMENTS [75]

- | | | |
|----------------------------------|------------------|------|
| Modal analysis of a violin octet | George Bissinger | 2105 |
|----------------------------------|------------------|------|

BIOACOUSTICS [80]

- | | | |
|--|---|------|
| Laryngeal biomechanics and vocal communication in the squirrel monkey (<i>Saimiri boliviensis</i>) | Charles H. Brown, Fariborz Alipour, David A. Berry, Douglas Montequin | 2114 |
| Variation in chick-a-dee calls of a Carolina chickadee population, <i>Poecile carolinensis</i> : Identity and redundancy within note types | Todd M. Freeberg, Jeffrey R. Lucas, Barbara Clucas | 2127 |
| Time-frequency model for echo-delay resolution in wideband biosonar | Nicola Neretti, Mark I. Sanderson, Nathan Intrator, James A. Simmons | 2137 |
| Investigations of the precedence effect in budgerigars: Effects of stimulus type, intensity, duration, and location | Micheal L. Dent, Robert J. Dooling | 2146 |
| Investigations of the precedence effect in budgerigars: The perceived location of auditory images | Micheal L. Dent, Robert J. Dooling | 2159 |
| Diversity in noise-induced temporary hearing loss in otophysine fishes | Sonja Amoser, Friedrich Ladich | 2170 |

ERRATA

- | | | |
|---|-------------------------------------|------|
| Erratum: "Benchmark solutions of plane wave bottom reflection loss" [J. Acoust. Soc. Am. 104, 3305–3312 (1998)] | Michael A. Ainslie, Alvin J. Robins | 2180 |
|---|-------------------------------------|------|

CUMULATIVE AUTHOR INDEX

2181

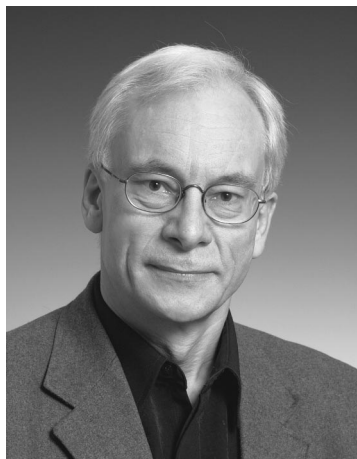
ACOUSTICAL NEWS—USA

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



Jens Holger Rindel—For research and teaching in room acoustics.



Robert S. Schlauch—For contributions to understanding intensity perception and attentional factors in hearing.



Ronald McKay—For contributions to the acoustics of performance spaces and teaching facilities.

Schultz Grant deadline extended

Applications for the seventh Theodore John Schultz Grant to be awarded by the Newman Fund will be received up to 30 June 2003. The Grant will be made in late summer of 2003.

The grant, usually in the amount of \$3000, provides partial support for the development of improved teaching methods, new curricula, or research in architectural acoustics education. Applicants must have teaching experience in acoustics and should manifest a strong desire to develop improved methods and materials for teaching architectural acoustics.

Applicants should send a curriculum vitae, a letter describing how the grant will be used, and an outline budget indicating expenses for which support is being sought. Applications and attachments should be postmarked no later than 30 June 2003 and mailed to: Newman Student Award Fund, Acoustical Society of America, Attn: Schultz Grant, 2 Huntington Quadrangle, Suite 1NO1, Melville, NY 11747-4502.

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.

2003

- | | |
|----------------|--|
| 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, |

12–16 May

23–25 June

27–30 July

5–8 Oct.

10–14 Nov.

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>].

Symposium on Environmental Consequences of Underwater Sound (ECOUS), San Antonio, TX [www.lsr.org/ECOUS].

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].

1st Conference on Acoustic Communication by Animals, University of Maryland, College Park, MD [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: <http://asa.aip.org/communication.html>].

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].

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 1N01, 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 1N01, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2270. Non-U.S. orders add \$11 per index.

Some indexes are out of print as noted below.

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

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

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

- Agarwal, Ramesh K., 1 Clayton Terr., St. Louis, MO 63131
- Agrest, Emmanuil, Arts and Sciences, Johnson and Wales Univ., 701 East Bay St., Charleston, SC 29403
- Alexander, Aaron S., Braden Manufacturing, 5199 North Mingo Rd., Tulsa, OK 74117
- Allen, Corrie B., 13304 West 64th Place, Arvada, CO 80004
- Anderson, Robert M., Geology and Geophysics, Univ. of Hawaii, 1680 East-West Rd., Post Bldg. 814B, Honolulu, HI 96822
- Andrew, Marilee A., APL/UW, Ctr. for Industrial & Medical Ultrasound, Box 355640, 1013 NE 40th St., Seattle, WA 98105-6698

- Aoyama, Katsura, Communication Disorders, Texas Tech. Univ. Health Sci. Ctr., 3601 4th St., Stop 6073, Lubbock, TX 79430-6073
- Arieh, Yaov, John B. Pierce Lab., Inc., 290 Congress Ave., New Haven, CT 06519
- Banchet, Julian, R/D Tech, 505 Boul. Du Parc Technologique, Quebec, PQ G1P 4S9, Canada
- Basinger, David L., Otologics, 5445 Airport Blvd., Boulder, CO 80301
- Bauhs, John A., 13 Red Fox Rd., North Oaks, MN 55127-6323
- Bayley, Antony M., Plantronics Limited, OEM Engineering, Interface Business Park, Bincknoll Ln., Wootton Bassett, Wilshire SN4 8QQ, U.K.
- Beardsley, Reginald H., 11507 Mile Dr., Houston, TX 77065
- Benazzato, Roberto, Via Vicenza 76, Camisano Vic. No. 36043, Italy
- Berger, Ralph, 22 Apple Ridge Rd., Cumberland, RI 02864
- Bistafa, Sylvio R., Mechanical Engineering, Univ. of Sao Paulo, Av. Prof. Mello Morhes, 2231, Sao Paulo SP 05508-900, Brazil
- Boon Siong, Wee, Block 506, Bedok North Ave. 3, #16-329, 460506 Singapore
- Brigham, Jonee Kulman, 678 Co. Rd. B., East, Maplewood, MN 55117
- Brodeur, Pierre, SoniSys LLC, 1734 Cooper Lake Dr., Smyrna, GA 30080-6400
- Buckingham, Christian E., NSTD, Johns Hopkins Univ., Applied Physics Lab., 11100 Johns Hopkins Rd., MS 8-314, Laurel, MD 20723-6099
- Campos Rodriguez, Arturo, Calle 25 No. 419-A X52, Jadines de Merida, Merida, Yucatan 97135, Mexico
- Carkner, Phil, Siplast, 1000 East Rochelle Blvd., Irvine, TX 75062
- Cheyne, Harold A., Voice and Speech Lab., Massachusetts Eye and Ear Infirmary, 243 Charles St., Boston, MA 02114
- Christensen, James L., Sunwest Technologies, Inc., 1050 Pioneer Way, Suite H, El Cajon, CA 92020
- Chulani, Hareh M., Electronics Dept., Inst. de Astrofisica de Canarias, c/via Lactea, S/N, La Laguna 38200, Spain
- Colby, Leider, School of Music, Univ. of Miami, P.O. Box 248165, Coral Gables, FL 33124-7610
- Cox, Benjamin T., Arup Acoustics, 8 St. Thomas St., Winchester, Hants SO23 9HE, U.K.
- Dance, Stephen, School of Engineering, South Bank Univ., Borough Rd., London SE1 0AA, U.K.
- Demain, Seth E., Wilson, Ihrig & Associates, Inc., 5776 Broadway, Oakland, CA 94618
- Drotleff, Horst A., Endersbacher Str. 54, Stuttgart 70374, Germany
- Duerr, Bernd, Kohlweisslingsweg 36, Dortmund NRW 44269, Germany
- Fajinni, Olatunji, K., O-Five Studio, No. 25 Ladoke Akintola Ave., New Bodija, Ibadan, Oyo State 30667, Nigeria
- Forbes, Barbara J., 92 Clitterhouse Rd., London NW2 1DJ, U.K.
- Giles, Peter M., Hydra Multi-Sonar System, General Dynamics Canada, 3785 Richmond Rd., Ottawa, ON K2H 5B7, Canada
- Good, Jr., Kenneth W., ABP Technologies, Armstrong World Ind., 2500 Columbia Ave., Lancaster, PA 17604
- Gross, Michael L., Kimberly-Clark Corp., 1400 Holcomb Bridge Rd., Roswell, GA 30076
- Guo, Yuqing, Ctr. for Hearing and Deafness, Univ. at Buffalo, 3435 Main St., Buffalo, NY 14214-3007
- Gygi, Brian M., Speech and Hearing Research, East Bay Inst. for Research and Education, 150 Muir Rd., Martinez, CA 94553
- Hadjicostic, Andreas N., Piezotech, 1600 West Main St., Lebanon, IN 46052
- Harner, Stephen A., Sub Zero Sound, 8081 Meadow Hill, Frisco, TX 75034
- Harrison, Matthew J., Bassett Acoustics, Level 12, 132 Arthur St., North Sydney, Sydney NSW 2060, Australia
- He, Anming, Siemens Medical Solutions USA, Inc., Ultrasound Division, 1230 Shorebird Way, Mountain View, CA 94039-7393
- Himmel, Chad N., JEAoustics, 5806 Mesa Dr., Suite 380, Austin, TX 78731-3742
- Hoffmann, Eckhard, FH Aalen, Gartenstrasse 135, Aalen D-73430, Germany
- Hong, Zheng, 1-14-7-912 Takahama, Mihama-ku, Chiba-City, Chiba 261-0003, Japan
- Horne, Peter R., Wrightson Kohnson Maddon & Williams, 4801 Spring Valley Rd., Suite 113, Dallas, TX 75244
- Huaquin, Mario A., Proyectos Acusticos Ltda., Clemente Fabres 874, Santiago RM 6641372, Chile
- Huta, Holly, Starkey Laboratories, 6600 Washington Ave., Eden Prairie, MN 55344

- Ip, Rodney C. W., 30/F Flat D., Blk. 5, Sceneway Garden Lamtin, Hong Kong, China
- Isakson, Marcia J., Applied Research Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029
- Jeong, Hong, POSTECH, EE Dept., Hyoja San 31, Pohang, Kyungbook 790-784, Republic of Korea
- Juan, Betts F., 214 Main St., Box 164, El Segundo, CA 90245
- Kachur, Amanda L., Albert Kahn Associates, Inc., 7430 Second Ave., Detroit, MI 48202
- Kadman, Yoram, Y. Kadman Acoustic and Vibration Engineering, 3 Stern St., Herzlia, Neve Amirim 46412, Israel
- Kim, Inyoung, Biomedical Engineering, Hanyang Univ., SungDong P.O. Box 55, Seoul 133-605, Korea
- Kolarik, Tammy A., 7210B Eden Brook Dr., Apt. T2, Columbia, MD 21046
- Lachs, Lorin, California State Univ., Fresno, Psychology, 5310 North Campus Dr., M/S PH11, Fresno, CA 93740
- Leland, Dennis L., Spectrum Technology Group, Inc., 21917 Foxlair Rd., Laytonsville, MD 20882
- Li, Fenghua, National Lab. of Acoustics, Inst. of Acoustics, CAS, No. 1 Zhongguancun Rd., P.O. Box 2712, Beijing 100080, China
- Line, Kenneth S., Hewlett-Packard, 11311 Chinden Blvd. 5U, MS 536, Boise, ID 83714
- Lussier, Justin R., 123 Bennett St., #2, Brighton, MA 02135
- Mackenzie-Hoy, Terry, Mackenzie-Hoy Consulting Electrical and Acoustics Engineers, 78 Hout St., Cape Town 7405, South Africa
- Marburg, Steffen, Inst. fuer Festkoerpermechanik, Technische Univ. Dresden, Dresden 01062, Germany
- Marchand, Sylvain, 27 Residence de Camponac, Pessac, Gironde 33600, France
- Mariano, Marcos, Caixa Postal 66, Santa Branca S. Paulo, 12380-970 Brazil
- Martens, William L., Graduate Dept. of Information Systems, Univ. of Aizu, Tsuruga, Aizu-Wakamatsu, Fukushima 965-8580, Japan
- Masiello, Brian L., Cavanaugh Tocci Associates, Inc., 327F Boston Post Rd., Sudbury, MA 01776
- Mazza, Riccardo, Strada Valpattonera 131, Torino 10133, Italy
- Medina, Thomas, Naval Undersea Warfare Ctr. Division/Keyport, Acoustics Analysis, C/532, 610 Dowell St., Bldg. 1003T, Keyport, WA 98345
- Meyers, James A., 2712 East 10th St., Bremerton, WA 98310
- Mihammad, Asim I., H. #323 Main Margalla Rd., F-11/3, Islamabad 44000, Pakistan
- Montlick, Terry, 89 Pidge Ave., Pawtucket, RI 02860
- Mueller, Benjamin C., Ostergaard Acoustical Associates, 200 Executive Dr., West Orange, NJ 07052
- Myshinsky, Ernst L., Ship and Industrial Acoustics, Krylov Shipbuilding, Research Inst., Moskovskoye Shosse 44, St. Petersburg 196158, Russia
- Naylor, Graham, Oticon A/S, Kongevejen 243, Snekkersten 3070, Denmark
- Niezrecki, Christopher, Mechanical and Aerospace Engineering, Univ. of Florida, P.O. Box 116250, Gainesville, FL 32611-6250
- Nomura, Hideyuki, Esta-Argo No. 101, Asahi-machi 34 gaiku 7, Kanazawa-shi, Ishikawa 920-0941, Japan
- Nowacek, Douglas P., Mote Marine Lab., Ctr. for Marine Mammal and Sea Turtle Research, 1600 Ken Thompson Parkway, Sarasota, FL 34236
- Ohki, Michio, School of Applied Sciences, National Defense Academy, 1-10-20 Hashirimizu, Yokosuka, Kanagawa 239-8686, Japan
- Ohm, Wonsuk, 307-550 Lang's Rd., Ottawa, ON K1K 4C2, Canada
- Oldham, David J., Architectural and Bldg. Eng., Univ. of Liverpool, Abercrombie Square, Liverpool L69 3BX, U.K.
- Oliver, Patrick J., Research and Development, VAW Systems Ltd., 1300 Inkster Blvd., Winnipeg, MB R2X 1P5, Canada
- Olson, John V., Geophysical Inst., Univ. of Alaska, 903 Koyukuk Dr., Fairbanks, AK 99775
- Osborne, Daniel L., Geophysical Inst., Univ. of Alaska, P.O. Box 75-7320, Fairbanks, AK 99775-7320
- Overweg, Cornelis H., Acentech, Inc., 1429 East Thousand Oaks Blvd., Suite 200, Thousand Oaks, CA 91362
- Page, John H., Physics and Astronomy, Univ. of Manitoba, 301 Allen Bldg., Fort Garry Campus, Winnipeg, MB R3T 2N5, Canada
- Park, Je-Geun, Dept. of Physics, SungKyunKwan Univ., Jangan Gu, Suwon, Kyeonggi Do 440-746, Korea
- Pehl, Gerald J., 608 40th Ave., NE, Columbia Heights, MN 55421
- Pernod, Philippe, Inst. D'Electronique de Microelectronique et de Nanotechnologie, Opto-Acousto-Electronique, Ecole Central de Lille, BP 48, Cite Scientifique, Villeneuve D'Ascq, Cedex 59651, France
- Perren, Gary E., North American Mfg. Co., 4455 East 71st St., Cleveland, OH 44105
- Pollock, Christopher J., Shen Milsom & Wilke, Inc., 417 Fifth Ave., 5th Floor, New York, NY 10016
- Preobrazhensky, Vladimir L., Inst. D'Electronique de Microelectronique et de Nanotechnologie, Opto-Acousto-Electronique, Ecole Central de Lille, BP 48, Cite Scientifique, Villeneuve D'Ascq, Cedex 59651, France
- Pyshin, Alexander V., Ship and Industrial Acoustics, Krylov Shipbuilding Research Inst., Moskovskoye Shosse 44, St. Petersburg 196158, Russia
- Rajapan, Dhilsha, National Inst. of Ocean Technology of Ocean Development, Velachery-Tambaram Main Rd., Pallikaranai, Chennai, Tamilnadu 601302, India
- Ricketts, Todd A., Audiology Dept., Vanderbilt Bill Wilkerson Ctr., 1114 19th Ave., South, Nashville, TN 37212
- Roux, Philippe, Marine Physical Lab., SIO-UCSD, 8820 Shellback Way, NTV Bldg., 4th Floor, La Jolla, CA 92093
- Rozo, Juan Pablo, Acustec de Colombia, Calle 84, #20-39, Bogata, c/Marca, Colombia
- Salvadores, Silvia R., Servicio Naval de Investigacion y Desarrollo, Div. Acustica Submarina, Av. Libertador 327, Vicente Lopez, Buenos Aires 1638, Argentina
- Schultz, William J., CMA, Inc., 205 Earl Rd., Shorewood, IL 60431
- Schwartz, Martin H., Northrop Grumman, 21240 Burbank Blvd., Woodland Hills, CA 91367
- Scott, Kimberly R., 1127 Narcisco St., NE, Albuquerque, NM 87112
- Shaw, Mark T., Engineering, Sonatech, Inc., 879 Ward Dr., Santa Barbara, CA 93111
- Shimada, Shoji, Electrical Engineering, Nagaoka Univ. of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
- Shuri, Hideyuki, Universal Shipbuilding Corp., Technical Research Ctr., 1-3 Kumozukokan-cho, Tsu-City, Mie-Pref 514-0398, Japan
- Skorskii III, Edwin S., Artec Consultants, Inc., 114 West 26th St., Floor 12, New York, NY 10001-6812
- Smith, Gordon P., Dept. of Physics and Astronomy, Western Kentucky Univ., 1 Big Red Way, Bowling Green, KY 42101
- Soker, Joseph N., Halutz 28 St., Jerusalem 96222, Israel
- Soler, Jaume, Acoustic Dimensions, 145 Huguenot St., Suite 406, New Rochelle, NY 10801
- Speaker, William H., 59150 Button Willow Dr., Clark, CO 90428
- Stockmann, Elle S., Research Lab. of Electronics, Speech Communications Group, 77 Massachusetts Ave., Rm. 36-511, Cambridge, MA 02139
- Su, Rosemary, Cavanaugh Tocci Associates, Inc., 327F Boston Post Rd., Sudbury, MA 01776
- Szuberla, Curt A. L., 1102 Norris Ln., Fairbanks, AK 99712-3018
- Taiwo, Michael M., O-Five Studio, No. 25 Ladoke Akintola Ave., New Bodija, Ibadan, Oyo State 30667, Nigeria
- Takano, Sayoko, CREST-JST/ATR, Dept. 1-HIS, 2-2-2 Hiraridai, Seika-cho, Saraku-gun, Kyoto 619-0288, Japan
- Tao, Dongling, 46079 Hanford Rd., Canton, MI 48187
- Trivett, David, Mechanical Engineering, Georgia Inst. of Technology, 771 Ferst Dr., Atlanta, GA 30332
- Vick, Jennell C., Speech Communication Group, MIT, 36-511, 77 Massachusetts Ave., Cambridge, MA 02139
- White, Lee A., 1358 Johnson Rd., Keller, TX 76248-4205
- White, Ronald G., Decoustics, 65 Disco Rd., Toronto, ON M9W 1M2, Canada
- Wilson, Rhonda J., Meridian Audio Limited, Stonehill, Stukeley Meadows, Huntingdon PE29 6EX, U.K.
- Wood, Russell D., OneSource BT, 961 Burke St., Winston Salem, NC 27101
- Wright, Matthew C. M., ISVR, Univ. of Southampton, Southampton SO17 1BJ, U.K.
- Wu, Jian-Da, Dept. of Mechanical and Automation, Da-Yeh Univ., 112 Shan-Jeau Rd., Da-Tsuen, Chang-Hwa 515, Taiwan
- Xu, Mubing, 1502-C King Ave., Columbus, OH 43212
- Yan, Hong Y., Biology Dept., Univ. of Kentucky, 101 Morgan Bldg., Lexington, KY 40506
- Yang, Yuzhen, Deere & Company, Technology Ctr., One John Deere Place, Moline, IL 61265
- Zapala, David A., Audiology, Mayo Clinic-Jacksonville, 4500 San Pablo Rd., Jacksonville, FL 32224

Zeng, Qinqian, United Technology Research Ctr., 411 Silver Ln., East Hartford, CT 06108
 Zhang, Zhaoyan, 3426 Tulane Dr., Apt. 34, Hyattsville, MD 20783
 Zigler, Gary R., Design Engineering, Hufcor, Inc., 2101 Kennedy Rd., Janesville, WI 53545

New Students

Aaronson, Neil, 1135 Michigan Ave., East Lansing, MI 48823
 Al-Khairy, Mohamed A., Program in Linguistics, Univ. of Florida, P.O. Box 115454, Gainesville, FL 32611-5454
 Allen, Peter B., 3502 Pipers Path, San Antonio, TX 78251
 Alvarado-Juarez, Miguel, 210 Monterey Rd., Seaside, CA 93955
 Alves, Steven A., Engineering, Univ. of Western Ontario, 1151 Richmond St., Suite 2, London, ON N6A 5B8, Canada
 Ardila, Mauricio, 624 Hinman Ave., #2, Evanston, IL 60202
 Beers, Christopher J., 11481 Chapin St., Chesterland, OH 44026
 Bell, Andrew, Research School of Biological Sciences, Australian National Univ., P.O. Box 476, Canberra ACT 2601, Australia
 Bernhardt, Elizabeth A., 102 Whitebluff Ln., Apt. LB, Cary, NC 27513
 Berry, Julie C., Universite Laval, Architecture, Quebec, QC G1K 7P4, Canada
 Boukis, Christos G., Electrical Engineering, Imperial College, Exhibition Rd., London SW7 2BT, U.K.
 Bowles, Benjamin G., Boston College, Dept. of Physics, 140 Commonwealth Ave., Chestnut Hill, MA 02467
 Bozinoski, Radoslav, 2448 North Fairview Ln., Rochester Hills, MI 48306
 Bracken, Jeffery A., 267 Gerald St., State College, PA 16801
 Brooks, Christopher J., Happy Landings, Salthill Rd., Fishbourne, Chichester, West Sussex PO19 3QD, U.K.
 Caclin, Anne, IRCAM-CNRS, 1 place Igor Stravinsky, Paris F-75004, France
 Chan, Hsiang-Chih, National Sun Yat-Sen Univ., Inst. of Undersea Technology, 70 Lien-hai Rd., Kaohsiung 804 Taiwan, ROC
 Choi, Sangsook, 1617 "F" St., Apt. 17, Lincoln, NE 68508
 Choy, Bill, VTech Telecommunications, Ltd., 23/F Tai Ping Industrial Center, Block 1, 57 Ting Kok Rd., Tai Po., China
 Christensen, Mads G., Vestergade 9, 3.tv., Norresundby DK-9400, Denmark
 Coudriet, Gregory A., 144 Oakwood Ave., Apt. C1, West Hartford, CT 06119
 Crave, Olivier, 26 Route de Rorh-Mez, Ploemeur, Morbihan 56270, France
 Cruikshank, Matthew E., Columbia College, 600 South Michigan, Chicago, IL 60605
 DeDecker, Paul M., Linguistics Dept., New York Univ., 719 Broadway, New York, NY 10003
 Diakhia, Ousmane, Information and Computer Sciences, Saitama Univ., 255 Shimo-Okubo, Saitama-shi, Saitama-ken 338-8570, Japan
 Duncan, Edward C., Rensselaer Polytechnic Inst., Engineering Science, 145 Sunset Terr., Troy, NY 12180
 Elkhater, Ramzi W., 3205 East Flamingo Rd., Las Vegas, NV 89121
 Everhard, Ian L., 245 West North Ave., Apt. 208, Chicago, IL 60610
 Faust, Bryan K., 8 Old Town Rd., #525, Ayer, MA 01432
 Fedak, Larissa A., 983 Summit Ave., Jersey City, NJ 07307
 Feinberg, David R., Psychology Dept., Univ. of St. Andrews, St. Mary's Quad, St. Andrews, Fife KY16 9JU, Scotland
 Ford, N., 3 Whitwick Rd., Copt. Oak, Markfield, Leicestershire LE67 9QB, U.K.
 Fuchs, Susanne, Ctr. for General Linguistics, ZES, Jaegerstr. 10-111, Berlin 10117, Germany
 Ganapathiraju, Madhavi K., Language Technologies Inst., Carnegie Mellon Univ., NSH 4602, Pittsburgh, PA 15213
 Gokhale, Nachiket H., Aerospace and Mechanical Engineering, Boston Univ., 110 Cummings St., Boston, MA 02215
 Gudigundla, Sai P., 2855 Kendale, Apt. 101, Toledo, OH 43606
 Haller, C. E. Kristian, Mechanical Engineering, Blekinge Inst. of Technology, Valhallav 1, Karlskrona 37151, Sweden
 Hanlon, Ellie H., 342 Dean St., Brooklyn, NY 11217
 Hardy, Pierre, Ecole Centrale de Lyon, Bat. E6, 36 Ave. Guy de Collongue, Ecully 69130, France
 Hastings, Aaron L., Mechanical Engineering, Purdue Univ., 1077 Ray W. Herrick Labs., West Lafayette, IN 47907-1077

Hisagi, Miwako, Speech and Hearing Science, City Univ. of New York, Graduate Ctr., 365 Fifth Ave., Rm. 1707, New York, NY 10016-4309
 Hoyle, Matt, 154 Barnard Ave., Apt. E, Asheville, NC 28804
 Hsieh, Li-Wen, 3-Fl., No. 4, Ln. 74, Shinglin 2nd Rd., Beitou Chiu, Taipei, 112, Taiwan, ROC
 Ingvalson, Erin M., Carnegie Mellon Univ., Psychology, Baker Hall, Pittsburgh, PA 15213
 Johnston, Miles C., 2327 Floyd Ave., #2, Richmond, VA 23220
 Jones, Guenevere J., School of Aquatic and Fisheries Science, Univ. of Washington, P.O. Box 355020, Seattle, WA 98195-5020
 Jorge, Alvarez D., Casilla 640, Santa Cruz de La Sierra, Bolivia
 Joyce, Andrea L., Dept. of Entomology, Texas A&M Univ., Minnie Belle Heep Bldg. 412, College Station, TX 77843
 Kang, Kookjin, Dept. of Sensor Engineering, Kyungpook National Univ., 1370 Sankyukdong, Bukgu, Daegu 702-701, Korea
 Kawakami, Akira, Akiba Heights A-1, 98-2 Yoneda, Yoneda-cho, Takasago, Hyogo 676-0805, Japan
 Kennedy, Daniel J., 40A St. Benedicts St., Norwich NR2 4AQ, England
 Khioe, Fung Wah, Dept. of Chinese, Translation and Linguistics, City Univ. of Hong Kong, 83 Tat Chee Ave., Kowloon, Hong Kong, China
 Khodayari-Rostamabad, Ahmad, No. 73, Saqafi Ave., Seraj St., Farjam St., Tehran 16867-36163, Iran
 Konepally, Niranjana R., 2058 Grand Blvd., Apt. 101, Montreal, QC H4B 2W8, Canada
 Kusel, Elizabeth T., Mathematical Sciences, Rensselaer Polytechnic Inst., 110 8th St., Amos Eaton Rm. 301, Troy, NY 12180
 Kuster, Martin, Renswoudelaan 132, Den Haag ZH 2546XK, The Netherlands
 Lampropoulos, Iraklis E., 2180 14th St., Apt. #2, Troy, NY 12180
 Larsen, Jeffrey B., 2218 8th St. Cir., Charleston, IL 61920
 Lee, Hyun-Kook, Inst. of Sound Recording, Univ. of Surrey, Guildford, Surrey GU2 7XH, England
 Lefkowitz, Kimberly, A., 16 Stonehedge Rd., West Nyack, NY 10994
 Li, Deyu, Mechanical Engineering, Univ. of Pittsburgh, Benedum Hall 560, 3700 O'Hara St., Pittsburgh, PA 15261
 Li, Kang, Speech and Hearing Sciences, Univ. of Arizona, 1131 East 2nd St., Tucson, AZ 85721
 Li, Qingmei, Ocean Engineering, Florida Atlantic Univ., 101 North Beach Rd., Dania Beach, FL 33004
 Liaw, Lily, 5190 Panama Ave., Richmond, CA 94804
 Low, Stephen W. K., SIM Open Univ. Center, Faculty of Science, Mathematics and Technology, 461 Clementi Rd., Singapore 599491, Singapore
 Maguluri, Gopi N., AME, Boston Univ., 110 Cummings St., Boston, MA 02215
 Marshall, Vincent T., Biological Sciences, Univ. of Missouri, 213 Tucker Hall, Columbia, MO 65211-7400
 Martinez, Meysa A., 509 Alafaya Woods Blvd., Apt. C, Oviedo, FL 32765
 McCabe, Marie E., 3424 Brookline Ave., Apt. 7, Cincinnati, OH 45220
 Mehta, Pavak A., Mechanical Engineering, The Ohio State Univ., 206 West 18th Ave., Columbus, OH 43210
 Messum, Piers R., 112 Warner Rd., London SE5 9HQ, U.K.
 Methuku, Reddy S., Dept. of Audiology, All India Inst. of Speech and Hearing, Mysore-6, Karnataka 57006, India
 Modarresi, Golnaz, 4116 Burney Dr., Austin, TX 78731
 Mooney, Aran T., Zoology, Marine Mammal Research, Univ. of Hawaii, P.O. Box 1106, Kailua, HI 96734
 Morris, Gary L., 106 Mervine St., Monterey, CA 93940
 Moseley, Stephen C., 677B Waupelani Dr., State College, PA 16801
 Naidu, Manish, 635 Cotanche St., #505, Greenville, NC 27858
 Nichols, Matt J., Columbia College, 676 North LaSalle, Chicago, IL 60610
 Nieva, Luis, 237 Cummins Highway, Apt. A31, Roslindale, MA 02131
 Nuar, Nicholas C., P.O. Box 27366, Lansing, MI 48909
 O'Gorman, David E., Eaton-Peabody Lab. of Auditory Physiology, 243 Charles St., Boston, MA 02114
 Oh, Suntaek, Ocean Acoustics Lab., Dept. of Earth & Marine Sci., Hanyang Univ., 1271 SA-I Dong, Ansan-Si, Kyungki-do 425-791, Republic of Korea
 Owen, Neil R., Applied Physics Lab., Ctr. for Industrial and Medical Ultrasound, 1013 NE 40th St., Seattle, WA 98105
 Pacheco, Andres A., El Ciruelillo, 3981, Vitacura, Santiago, Chile
 Paek, Insu, Mechanical Engineering, Herrick Labs., Purdue Univ., 140 South Intramural Dr., West Lafayette, IN 47907-2031

Park, Jong Beom, Mechanical Engineering, Purdue Univ., Rm. 243, Herrick Lab., West Lafayette, IN 47907
 Pugh, Adam G., Vanderbilt Station B, Box 28444, Nashville, TN 37235-2844
 Reller, Christoph, Rosalind Close 28, Colchester, Essex CO4 3JH, U.K.
 Ripley, Jennifer, L., 221 Roberts St., Fairmount, WV 26554
 Rollins, Sarah, 660 North 200 East, #14, Provo, UT 84606
 Rozell, Christopher J., ECE, Rice University, MS 366, P.O. Box 1892, Houston, TX 77251-1892
 Saad, Omar, 30 Cambridge Park Dr., #2136, Cambridge, MA 02140
 Sainclair, Suzanne S., Communications Sciences and Disorders, Penn State Univ., 110 Moore Bldg., University Park, PA 16802
 Sarpun, Ismail, Adnan Menderes Blv. Diker, Apt. Mo. 136/4, Afyon 03040, Turkey
 Seal, Chris R., Wake Forest, P.O. Box 28352, Winston-Salem, NC 27109
 Seckman, Aaron M., Studiomedica Recording, 1030 David St., Evanston, IL 60201
 Seo, Jongbum, Biomedical Engineering Dept., Univ. of Michigan, Biomedical Engineering Dept., 2350 Hayward, 3304 G. G. Brown, Ann Arbor, MI 48109-2125
 Shirai, Setsuko, Dept. of Linguistics, Univ. of Washington, Box 354340, Seattle, WA 98195-4340
 Shub, Daniel E., Harvard-MIT Div. Of Health Sci. and Tech., Hearing Research Ctr., 44 Cummington St., Boston, MA 02215
 Sieler, Ryan G., 525 8th Ave., SE, Apt. 2, Minneapolis, MN 55414
 Smith, Heather M., Physics Dept., Brigham Young Univ., N283, ESC, Provo, UT 84602
 Sparks, Geoffrey S., 1101 Saint Paul St., Unit 811, Baltimore, MD 21202
 Stumpf, Kelly R., 205 Hartford Rd., Apt. D-5, New Britain, CT 06053
 Sui, Lei, Aerospace and Mechanical Engineering, Boston Univ., 110 Cummington St., Boston, MA 02215
 Theodore, Rachel M., 201 North Willow, Kent, OH 44240
 Thompson, Lauren M., 5162 Brook Way, #3, Columbia, MD 21044
 Thompson, Robert L., PSC 79, Box 161, APO, AE 09714
 Tornberg, Jarkko, Signal to Noise, Inc., Kilpisenkatu 8, Jyväskylä 40740, Finland
 Valente, Daniel P., Graduate Program in Acoustics, Pennsylvania State Univ., P.O. Box 30, State College, PA 16804
 Vasan, Srinu S., B. E. Information Technology, Sona College of Technology, Thiagarajar Rd., Salem 5, Salem, Tamil Nadu 636-005, India
 Vier, Matthew S., Calyton State Univ., Spivey Hall, 5900 North Lee St., Morrow, GA 30260
 Watkins, Emily S., 927 Meadowbrook Dr., Syracuse, NY 13224-1952
 Welby, Pauline, Linguistics Dept., Ohio State Univ., 1712 Neil Ave., #222, Columbus, OH 43210
 Wickstrom, Sueann, 968 NW 1st Ave., Chosholm, MN 55719
 Wieberg, Kimberly, M., 8207 Vista Colorado, Las Vegas, NV 89123
 Wochner, Mark S., 601 Marjorie Mae St., State College, PA 16803
 Wong, Pusan, 4559 Bedford Ave., Brooklyn, NY 11235
 Wong, Wai Yi P., 101 Curl Dr., Rm. 607, Columbus, OH 43210
 Wullens, Frederic, Dept. of Applied Acoustics, Chalmers Univ. of Technology, Sven Hultins Gata 8A, Goteborg 41296, Sweden
 Xiaoyan, Shen, Third Inst. of Oceanography State Oceanic Administration, 178 Daxue Rd., Xiamen, Fujian 361005, China
 Yu, Linxiao, Ctr. for Nondestructive Evaluation, 1915 Scholl Rd., Ames, IA 50011-3042
 Yuan, Hanfeng, Electrical Eng. and Computer Sci., Massachusetts Inst. of Tech., 50 Vassar St., Rm. 36-757, Cambridge, MA 02139
 Zhu, Wenhao, 1383 Bethamy Ln., Ottawa, ON K1J 8P5, Canada

New Electronic Associates

Ban, Thomas O., 15134 Beachview Ave., White Rock, BC V4B 1P6, Canada
 Beuselinck, Michael T., 6931 Reynolds St., Pittsburgh, PA 15208
 Bishop, Dorothy V. M., Experimental Psychology, Univ. of Oxford, South Parks Rd., Oxford OX1 3UD, U.K.
 Cohen, Penny D., 397 Ogden Ave., Teaneck, NJ 07666
 Collins, Simon W., 76 Paxton St., Malvern East VIC 3145, Australia
 Coulston, Rachel F., 2135 NW Flanders St., Apt. 202, Portland, OR 97210
 Dhami, Chandrajett S., Kalyan, Mumbai, Maharashtra 421306, India

Dunbar, John A., Dept. of Geology, Baylor Univ., 1311 South Fifth St., Waco, TX 76798
 Hobelsberger, Max, Dorfstr. 16, Wuerenlingen 5303, Sweden
 Houghton, Benjamin A., Shen Milsom & Wilke, Inc., 3300 North Fairfax Dr., Ste. 302, Arlington, VA 22201
 Hunt, Dave L., Dave Hunt Audio, 30 Stronsa Rd., London W12 9LB, U.K.
 Irza, John, Syngus Technology, Inc., 23 Richardson Ave., Arlington, MA 02476
 Johansen, Espen S., Weatherford, 16600 Park Rd., Houston, TX 77084
 Jones, Edward, Electronic Engineering, National Univ. of Ireland, Galway, University Rd., Galway, Ireland
 Kellett, Paul, Human Factors in Sound and Vibration, QinetiQ, Iveley Rd., Farnborough, Hants GU14 0LX, U.K.
 Krasnic, Alexandre S., 57 Brookfield Rd., London W4 1DF, U.K.
 Latimer, Paul J., 1045 Bannister Cir., Waldorf, MD 20602
 Martellotta, Francesco, Politecnico di Bari, Dipt. Di Fisica Technica, via Orabona 4, Bari 70125, Italy
 Nagar, Ron, Glucon, P.O. Box 3098 K. Arye, Petach Tikva 49130, Israel
 Pardue, Sally J., Mechanical Engineering, Tennessee Technological Univ., 115 West 10th St., Box 5014, Cookeville, TX 38505
 Peters, Dennis J., Science Applications Intl. Corp., Gulf Coast Office, 1140 Eglon Parkway, Shalimar, FL 32579
 Pompei, F. Joseph, 51 Water St., Watertown, MA 02472
 Sanchez, Maria M., Estebanez Calderson 7, 8C, Madrid 28020, Spain
 Schaub, Edward, Physics Dept., Baylor Univ., P.O. Box 97316, Waco, TX 76798-7316
 Schwartz, David A., Duke Univ., Neurobiology, 3209 DUMC, Durham, NC 27710
 Strong, Henry, 8710 Belmont Rd., Potomac, MD 20854
 Torres, Eddie G., RBF Consulting, 14725 Alton Parkway, Irvine, CA 92618

New Corresponding Electronic Associates

Adeyemi, Rahman O., Awobi Intl. Nigeria Limited, BB16F Alaba International Market, Lagos State, 23401, Nigeria
 Alonso, Victor, P.O. Box 10728, San Juan PR 00922-0728, Puerto Rico
 Camargo, Fernanda S., Rue Vitor Costa 440, Sao Paulo 04150-060, Brazil
 Diazgranados, Maria Claudia, Foundation Omacha, KR. 7-27-40 P. 4, Bogota DC, Colombia
 Hou, Zhaorong, Intel China, Ltd., Intel China Research Ctr., 601, North Tower, Kerry Center, 1 Guanghai Rd., Chaoyang District, Beijing 100020, P. R. China
 Kaibin, Qui, Geomechanics
 MacInnes, Craig S., Electrical Engineering, Pontificia Univ. Catolica de Peru, Av. Universitario cdra 18, Lima, Peru
 Manthravadi, Umashankar, L-35 G Sheikh Serai II, New Delhi 110017, India
 Moreno, Eduardo, Pezuela 86, Cojimar, La Habana 19140, Cuba
 Nepomuceno, Jose A., Acustica & Sonica, Rue Fradique Coutinho, 955 Sala 01, San Paulo 05416-011, Brazil
 Padmanabhan, Chandramouli, 108 Kumarappa St. Nungambakkkan, Chennai, Tamilnadu 600034, India
 Schlumberger Technologies, Ltd., 8 Floor, Chuanxin Bldg., Tsinghua Science Park, Beijing 100084, China

Member to Honorary-Fellow

M. S. Longuet-Higgins

Members Elected Fellows

H. P. Gottlieb, J. S. Jaffe, R. M. Keolian, O. Leroy, R. Lim, R. L. McKay, Z.-H. Michalopoulou, J. H. Rindel, R. S. Schlauch, B. Schulte-Fortkamp, S. D. Sommerfeldt, M. A. Svirsky, F. E. Toole, D. P. Walsh, B. A. Wright

Associates to Members

J. F. Antaki, P. H. Bedenbaugh, V. B. Biesel, M. Fatemi, W. O. Hughes, R. A. Iah, R. G. Ingham, J. A. Simmen, J. M. Spender, R. S. L. Tengham, S. Wang, B. Yang

Students to Associates

N. Abdelli, C. Bazua-Duran, S. V. Bharadwaj, H. A. Bradley, T. L. Brooks, M. H. Burk, C. L. Compton, D. B. Copeland, C. L. Davies, A. B. Dittberner, D. A. Ellis, S. H. Ferguson, K. E. Fisher, M. M. George, A. Giacometti, B. W. Harrison, K. J. Hays-Stang, P. P. Henry, B. W. Y. Hornsby, K. H. Kim, D. P. Knight, P. J. Lesniewski, S. K. Levey, J. S. Martin, R. D. Mason, G. Modarresi, D. J. Parker, A. G. Petculescu, P. Ratilal, J. L. Raymond, D. B. Reeder, A. E. Russo, K. R. Shandera, T. G. Simmons, D. L. Sokolov, M. E. Swearingen, N. Thubthong, R. L. Trousil, S. J. Volaric, M. A. Walton, P. J. Wolfe, B. Wrede, K. Yoneyama, Z. Zhang, S. Zhou

Members to Electronic Associates

L. Fishman, B. K. Kimberley, L. C. Kryzac, F. M. J. Rocaboy

Associates to Electronic Associates

S. Deguchi, K. J. Delaney, P. J. Gendron, R. A. Hedges, R. Kahn, A. Kinoshita, C. J. Long, A. Pelletier, E. M. Ring, B. M. Sadler, T. Tateno, P. Traykovski, A. Weber

Students to Electronic Associates

G. J. Berg, W.-S. Chen, G. Essl, R. M. Harwell, T. McCabe, A. A. Scharine, S. M. Spitzer, G. Tzanetakis, G. K. Vallabha

Members to Corresponding Electronic Associates

E. C. Biassoni, L. Krige, C.-W. Lim

Associates to Corresponding Electronic Associates

A. T. Barros, I. Psencik, H. C. Yehia

Students to Corresponding Electronic Associates

A. A. El-Khateeb, F. Fiates

Electronic Associate to Corresponding Electronic Associate

M. S. Hnilo

Reinstated

K. A. Owusu-Anash, L. Wang—*Members*
R. Patel—*Student*
D. W. Gow—*Electronic Associate*

Resigned

M. G. Studdert-Kennedy—*Fellow*
R. Batherman, G. R. Ebbeson, D. W. Fitting, I. Fraser, L. Kay, V. S. Pascucci, H. Stenneken, S. Suresh, E. B. Viveiros, S. T. Zavtrak—*Members*
D. Bodnar, J. A. Bauhs, G. J. Gannon, L.-M. Chen, G. V. Dalrymple, F. R. Fricke, K. Holliger, M. Latzel, A. S. Lessem, M. Lydolf, A. Morris, J. V. Pierpoint, P. G. Schreiner III, R. Smitsm, O. Stearman, R. Vinokur, A. Wong, J. P. Zamzes—*Associates*
J. M. Mencik, A. Purwanto, B. Tramm, Y. Xiang—*Students*

Deceased

R. C. Bilger, E. F. Elkins, L. W. Sepmeyer—*Fellows*
W. L. Baker, P. W. Barnett, F. P. Beguin, H. W. Kolbe, A. S. Nikiforov, C. B. Pattarini—*Members*
P. Brodeur, D. L. Kirk—*Associates*

Fellows	902
Members	2591
Associates	2842
Students	1014
Total	7349

ACOUSTICAL NEWS—INTERNATIONAL

Walter G. Mayer

Physics Department, Georgetown University, Washington, DC 20057

French Acoustical Society honors Gilles Daigle

Gilles Daigle of the CNR Canada, a Fellow of the *Society*, was awarded the 2002 Médaille Etrangère by the French Acoustical Society (S.F.A.). This medal, formerly known as the Silver Medal, is presented once a year to an outstanding acoustician outside France. Gill Daigle is the sixth North American recipient of the medal. The others are L. Beranek 1966; W. G. Mayer 1988; G. Weinreich 1991; O. B. Wilson 1992; H. Überall 1996; and H. Levine 1998.

International Meetings Calendar

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

April 2003

- 6–10 **IEEE International Conference on Acoustics, Speech, and Signal Processing**, Hong Kong, Hong Kong. (Web: www.en.polyu.edu.hk/%7Ecassp03)
- 7–9 **WESPAC8**, Melbourne, Australia. (Web: www.wespac8.com)

May 2003

- 19–21 **5th European Conference on Noise Control (Euro-noise 2003)**, Naples, Italy. (Fax: +39 81 239 0364; Web: www.euronoise2003.it)
- 22–23 **2nd International Styrian Noise, Vibration & Harshness Congress**, Graz, Austria. (Fax: +43 316 873 4002; Web: www.acgraz.com)

June 2003

- 8–13 **XVIII International Evoked Response Audiometry Study Group Symposium**, Puerto de la Cruz, Tenerife, Spain. (Web: www.ierasg-2003.org)
- 16–18 **Acoustics 2003—Modeling & Experimental Measurements**, Cadiz, Spain. (Fax: +44 238 029 2853; Web: www.wessex.ac.uk/conference/2003/acoustics/index.html)
- 29–3 **8th Conference on Noise as a Public Health Problem**, Amsterdam-Rotterdam, The Netherlands. (Fax: +31 24 360 1159; Web: www.icben2003.nl)
- 30–3 **Ultrasonics International (UI'03)**, Granada, Spain. (Fax: +44 1295 253 334; Web: www.ui03.com)

July 2003

- 7–11 **10th International Congress on Sound and Vibration**, Stockholm, Sweden. (Fax: +46 88 661 9125; Web: www.congex.com/icsv10)
- 14–16 **8th International Conference on Recent Advances in Structural Dynamics**, Southampton, UK. (Web: www.isvr.soton.ac.uk/sd2003)

August 2003

- 6–9 **Stockholm Music Acoustics Conference 2003 (SMAC03)**, Stockholm, Sweden. (Web: www.speech.kth.se/music/smac03)
- 25–27 ***Inter-Noise 2003**, Jeju Island, Korea. (Fax: +82 42 869 8220; Web: www.internoise2003)
- 25–29 **XIII Session of the Russian Acoustical Society**, Moscow, Russia. (Fax: +7 095 126 0100; Web: www.akin.ru)

September 2003

- 1–4 **Eurospeech 2003**, Geneva, Switzerland. (Web: www.symporg.ch/eurospeech2003)
- 7–10 **World Congress on Ultrasonics**, Paris, France. (Fax: +33 1 46 33 56 73; Web: www.sfa.asso.fr/wcu2003)
- 16–19 **Autumn Meeting of the Acoustical Society of Japan**, Nagoya, Japan. (Fax: +81 3 5256 1022; Web: wwwsoc.nii.ac.jp/asj/index-e.html)
- 18–19 **Surface Acoustics 2003**, Salford University, Manchester, UK. (Web: www.ioa.org.uk/salford2003)
- 23–25 **2nd International Symposium on Fan Noise**, Senlis, France. (Fax: +33 4 72 44 49 99; Web: www.fannoise2003.org)

October 2003

- 15–17 **34th Spanish Congress on Acoustics**, Bilbao, Spain. (Fax: +34 91 411 7651; Web: www.ia.csic.es/sea/index.html)
- 15–17 **Acoustics Week in Canada**, Edmonton, AB, Canada. (Fax: +1 780 414 6376; Web: caa-aca.ca/edmonton-2003.html)

December 2003

- 10–12 **3rd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications**, Firenze, Italy. (Fax: +39 55 479 6767; Web: www.maveba.org)

March 2004

- 17–19 **Spring Meeting of the Acoustical Society of Japan**, Atsugi, Japan. (Fax: +81 3 5256 1022; Web: wwwsoc.nii.ac.jp/asj/index-e.html)
- 22–25 ***Joint Congress of the French and German Acoustical Societies (SFA-DEGA)**, Strasbourg, France. (Fax: +49 441 798 3698; Web: www.sfa.asso.fr/cfa-daga2004)
- 31–3 **International Symposium on Musical Acoustics (ISMA2004)**, Nara, Japan. (Fax: +81 77 495 2647; Web: www2.crl.go.jp/jt/a132/isma2004)

April 2004

- 5–9 **18th International Congress on Acoustics (ICA2004)**, Kyoto, Japan. (Fax: +81 66 879 8025; Web: www.ica2004.or.jp)
- 11–13 **International Symposium on Room Acoustics (ICA2004 Satellite Meeting)**, Hyogo, Japan. (Fax: +81 78 803 6043; Web: rad04.iis.u-tokyo.ac.jp)

June 2004

- 8–10 **Joint Baltic–Nordic Acoustical Meeting**, Mariehamn, Åland, Finland. (Fax: +358 09 460 224; e-mail: asf@acoustics.hut.fi)

July 2004

- 5–8 **7th European Conference on Underwater Acoustics (ECUA 2004)**, Delft, The Netherlands. (Fax: +31 70 322 9901; Web: www.ecua2004.tno.nl)
- 11–16 **12th International Symposium on Acoustic Remote Sensing (ISARS)**, Cambridge, UK. (Fax: +44 161 295 3815; Web: www.isars.org.uk)

August 2004

- 23–27 **2004 IEEE International Ultrasonics, Ferroelectrics,**

22–25	<p>and Frequency Control 50th Anniversary Conference, Montreal, Canada. (Fax: +1 978 927 4099; Web: www.ieee-uffc.org/index2-asp)</p> <p>*Inter-noise 2004, Prague, Czech Republic. (Inter-Noise Secretariat, Technická 2, 16627 Praha 6, Czech Republic; Web: www.internoise2004.cz)</p>	<p>August 2005 7–10</p> <p>28–2</p>	<p>*Inter-Noise, Rio de Janeiro, Brazil. (Contacts to be announced later)</p> <p>Forum Acusticum Budapest 2005, Budapest, Hungary. (Contacts to be announced later)</p>
<p>September 2004 13–17</p>	<p>4th IberoAmerican Congress on Acoustics, 4th Iberian Congress on Acoustics, 35th Spanish Congress on Acoustics, Guimarães, Portugal. (Fax: +351 21 844 3028; e-mail: dsilva@lnec.pt)</p>	<p>Preliminary Announcement 2008 June/July</p>	<p>*Joint Meeting of European Acoustical Association (EAA), Acoustical Society of America (ASA), and Acoustical Society of France (SFA), Paris, France. (Details to be announced later)</p>

REVIEWS OF ACOUSTICAL PATENTS

Lloyd Rice

11222 Flatiron Drive, Lafayette, Colorado 80026

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

Reviewers for this issue:

GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*
 ALIREZA DIBAZAR, *Department of BioMed Engineering, University of Southern California, Los Angeles, California 90089*
 MARK KAHRS, *Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261*
 HASSAN NAMARVAR, *Department of BioMed Engineering, University of Southern California, Los Angeles, California 90089*
 DAVID PREVES, *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*
 KEVIN P. SHEPHERD, *Mail Stop 463, NASA Langley Research Center, Hampton, Virginia 23681*
 WILLIAM THOMPSON, JR., *Pennsylvania State University, University Park, Pennsylvania 16802*
 ERIC E. UNGAR, *Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
 ROBERT C. WAAG, *Department of Electrical and Computer Engineering, Univ. of Rochester, Rochester, New York 14627*

6,466,514

43.30.Sf METHOD FOR ELIMINATION OF PASSIVE NOISE INTERFERENCE IN SONAR

Darrin W. Kabel, assignor to Garmin Corporation
 15 October 2002 (Class 367/98); filed 12 June 2000

A fish-finder type of sonar system employs some signal processing to mitigate the masking effects of ambient noise. First, a reading of the ambient noise is made to establish a detection threshold. Then the projector is excited and the subsequent echo signals are recorded. If the intensity of the reflected signal is less than the established detection threshold, this signal is eliminated from further processing. The detection threshold is increased in value over time to compensate for any increases in gain in the receiver and it is periodically updated to compensate for changes in the ambient noise.—WT

6,449,566

43.30.Yj ACOUSTIC SCATTERING MEASUREMENT AND PROCESSING FOR DETERMINING VARIANCES IN MULTIPLE FEATURES

John Oeschger, assignor to the United States of America as represented by the Secretary of the Navy
 10 September 2002 (Class 702/54); filed 6 November 2000

The aim of this patent is to provide a method for determining variances of such features as temperature and salinity in an area within a body of water, without placing sensors in the actual area of interest. Acoustic sources are placed on one side of a selected scattering direction and an acoustic receiver is placed at a mirror-imaged location (relative to the scattering direction) for each source, so that a Bragg scattering wave vector associated with each source/receiver pair is parallel to the selected scattering direction. The sources are operated separately to direct a broadband ultrasonic pulse at the region of interest. The resulting waves are detected by the paired receivers, converted to spectral waveforms, and used to determine a band of Bragg wave numbers. The complex acoustic scattering relationship is solved as a function of the band of Bragg wave numbers for each source/receiver pair and used to evaluate variations in multiple features of the medium.—EEU

6,441,703

43.35.Cg MULTIPLE FREQUENCY ACOUSTIC REFLECTOR ARRAY AND MONOLITHIC COVER FOR RESONATORS AND METHOD

Carl M. Panasik, assignor to Texas Instruments Incorporated
 27 August 2002 (Class 333/189); filed 18 January 2000

Efficient crystal or thin-film resonators, such as those used in televisions, radios, and cell phones, require terminations with impedances that differ greatly from those of the resonators. A solid layer of a given thickness provides a high-impedance termination at the frequency at which the thickness encompasses a quarter-wavelength, and thus is useful for only a limited frequency range. The present patent pertains to a termination consisting of a series of layers, conceived to provide suitable impedances for multiple resonators supported on it.—EEU

6,437,478

43.35.Pt SURFACE ACOUSTIC WAVE RESONATOR FILTER

Yasushi Yamamoto To, assignor to NEC Corporation
 20 August 2002 (Class 310/313 B); filed in Japan 24 August 1992

Surface acoustic wave devices that are small and relatively inexpensive are widely used as band-pass filters in such communication devices as portable telephones. The device described in this patent consists of interdigital electrodes on a piezoelectric substrate and is constructed so that it couples to a surface wave that propagates in the longitudinal direction and also to one that propagates in the transverse direction. Wide pass-band characteristics are obtained by coupling of these resonant modes or by coupling together of the first and second longitudinal modes.—EEU

6,437,479

43.35.Pt SURFACE ACOUSTIC WAVE DEVICE

Michio Miura *et al.*, assignors to Fujitsu Limited; Fujitsu Media Devices Limited
20 August 2002 (Class 310/313 R); filed in Japan 29 June 2001

This device consists of an interdigital transducer attached to a piezoelectric substrate. In order to increase the resistance of this device to problems associated with different thermal expansions of these components, a trough in the substrate is filled with a material that has a low coefficient of thermal expansion.—EEU

6,448,513

43.35.Zc ELECTRONIC WEIGHING APPARATUS UTILIZING SURFACE ACOUSTIC WAVES

Vyacheslav D. Kats and Arnold S. Gordon, assignors to Circuits & Systems Incorporated
10 September 2002 (Class 177/210 R); filed 2 February 2001

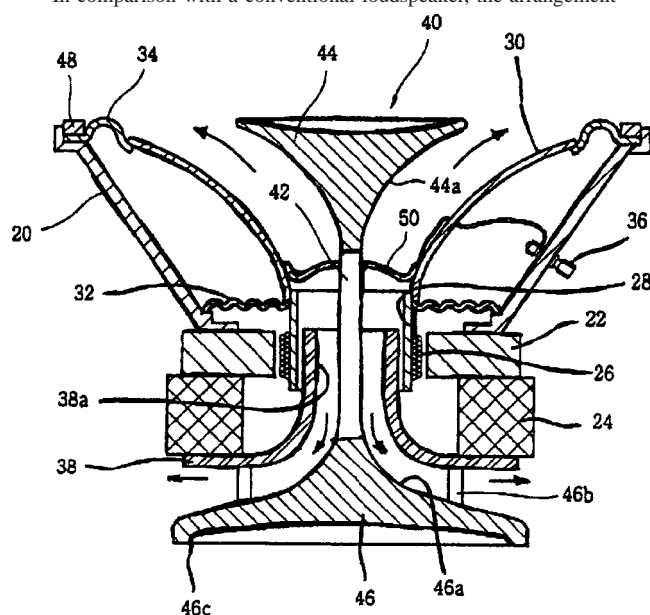
Precision electronic weighing is achieved by using surface acoustic wave devices essentially as strain gages to measure the deflections of elastic elements. The attendant surface acoustic wave oscillator is designed to oscillate in the frequency range in which it has the best phase linearity.—EEU

6,438,246

43.38.Ja SPEAKER APPARATUS

Jae-Nam Kim, assignor to Samsung Electronics Company, Limited
20 August 2002 (Class 381/337); filed in the Republic of Korea 3 November 1997

In comparison with a conventional loudspeaker, the arrangement



shown is said to provide more uniform sound radiation and more efficient heat dissipation.—GLA

6,435,301

43.38.Ja APPARATUS FOR THE REDISTRIBUTION OF ACOUSTIC ENERGY

Emanuel LaCarrubba, Sausalito, California
20 August 2002 (Class 181/155); filed 5 May 2000

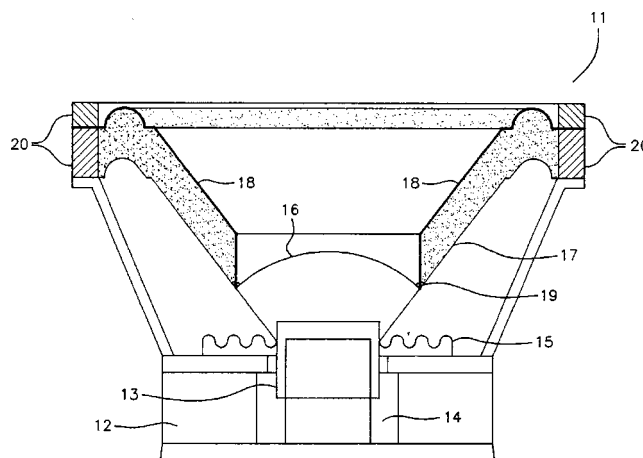
This patent is a continuation of United States Patent No. 6,068,080 [reviewed in J. Acoust. Soc. Am. **109**(1), 22 (2001)]. It describes an acoustic reflector consisting of elliptical surfaces, intended to provide a broad sound radiation pattern with controlled directivity in two dimensions and with minimal frequency response anomalies.—EEU

6,445,803

43.38.Ja SPEAKER

Chuan How Boon and Jason Noe Kia-Chong Boon, both of Auckland, New Zealand
3 September 2002 (Class 381/184); filed in New Zealand 16 December 1999

Auxiliary cone **18** is spaced above conventional cone **17** and driven through collar **19**. It is also driven by air trapped between the two cones.



The arrangement is said to improve sound quality at low levels and remove problems associated with high-frequency noises. Whether the noises themselves are removed is left for us to guess.—GLA

6,445,806

43.38.Ja TUNED ELASTIC LOUDSPEAKER ENCLOSURE

Michael L. Jacobson, Leesburg, Virginia
3 September 2002 (Class 381/345); filed 8 February 2000

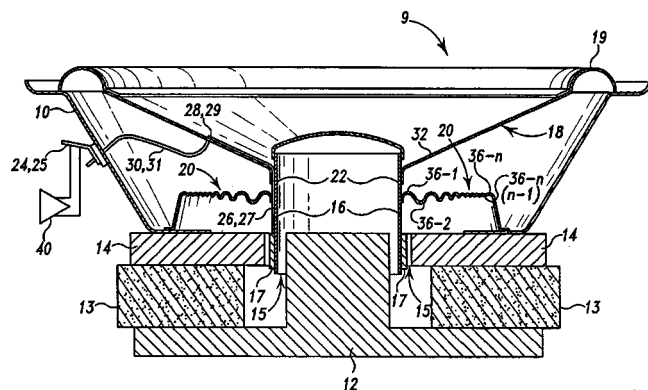
If a loudspeaker is mounted on an enclosure having flexible walls, such as a plastic jug, then wall vibrations will generate sound in addition to that emanating from the front of the speaker cone. This idea crops up from time to time and a number of variants have been patented. In this case, (a) the enclosure is made from an integral, elastic, self-supporting material with a single opening for the speaker, and (b) the speaker is mounted on a vibration isolator rather than being attached directly to the jug.—GLA

6,449,375

43.38.Ja LOUDSPEAKER SPIDER WITH REGRESSIVE ROLLS

Steven W. Hutt, assignor to Harmon International Industries, Incorporated
10 September 2002 (Class 381/403); filed 22 September 1999

Centering spider 20 is formed with deep corrugations at its inner edge and shallow corrugations near the perimeter. The patent teaches that with conventional equal corrugations there is a greater amount of material in the outer rolls and, as a result, "...the stiffness of the rolls varies regressively outwardly." True enough if only hoop stresses are considered but the varia-



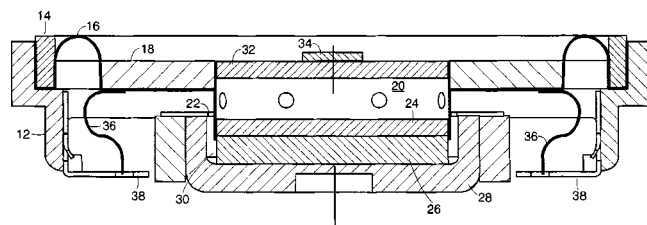
tion in radial stiffness is just the opposite. I tried laying a short straight edge along the corrugations of a conventional spider and found no observable dishing at maximum excursions. However, my simple experiment cannot predict the more complicated relationship between suspension linearity and acoustic performance. The device may in fact reduce low-frequency distortion, as the patent asserts.—GLA

6,449,376

43.38.Ja PLANAR-TYPE LOUDSPEAKER WITH AT LEAST TWO DIAPHRAGMS

Carlos Beltran *et al.*, assignors to Boston Acoustics, Incorporated
10 September 2002 (Class 381/423); filed 20 September 1999

A conventional voice coil assembly drives coaxial planar diaphragms 18 and 32. In addition, a small damper 34 may be attached to the inner diaphragm. The idea is to restrict the number of natural modes in each



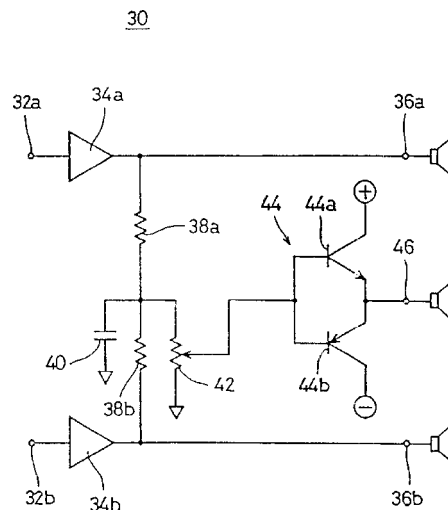
diaphragm such that a combination of mechanical coupling and acoustic interference provides relatively flat response over the desired range of frequencies.—GLA

6,438,237

43.38.Lc 3D WOOFER DRIVE CIRCUIT

Takeshi Nakamura, assignor to Murata Manufacturing Company, Limited
20 August 2002 (Class 381/27); filed in Japan 1 October 1998

The circuit shown is a simple, inexpensive way to drive a subwoofer. The low-pass section is not restricted to a first-order filter. In the patent claims, elements 38a, 38b, 40, and 42 are defined collectively as "an im-



pedance element" which may include "a low-pass filter." That is exactly how your reviewer connected a subwoofer to his stereo system 25 years ago.—GLA

6,447,359

43.38.Md MEMORIAL NOVELTY DOLL DEVICE HAVING INTEGRAL SOUND PRODUCING MEANS AND KIT AND METHOD FOR CUSTOMIZING THE SAME

Carlos D. B. Crump, Little Rock, Arkansas
10 September 2002 (Class 446/100); filed 3 August 2001

A macabre concept: somehow (not specified in the patent) customize a doll to match a deceased friend or relative's visage. Add a sound chip to store the voice. The author contends that it will help "keep memories alive."—MK

6,449,887

43.38.Md WATER GLOBE WITH TOUCH SENSITIVE SOUND ACTIVATION

Jin K. Song, assignor to Jin K. Song
17 September 2002 (Class 40/406); filed 25 April 2000

One of the wonders of "snow globes" is their delightfully simple operation. This patent would eliminate the fun by adding a stirrer so that it snows as long as the batteries last, and of course, provide audio output.—MK

6,452,525

43.38.Md ONE-BIT SECOND ORDER SIGMA-DELTA MODULATOR INCLUDING A POLYNOMIAL VECTOR QUANTIZER

Truong-Thao Nguyen, assignor to Research Foundation of the City University of New York
17 September 2002 (Class 341/143); filed 5 April 2001

The author proposes the use of a nonlinear operator (specifically, a squarer) inside the feedback loop of a sigma-delta modulator so that the mean square of the quantization error follows the same decay as higher order modulators.—MK

6,454,625

43.38.Md INTERACTIVE TALKING DOLLS

Peter Sui Lun Fong and Chi Fai Mak, assignors to Peter Sui Lun Fong
24 September 2002 (Class 446/297); filed 13 June 2001

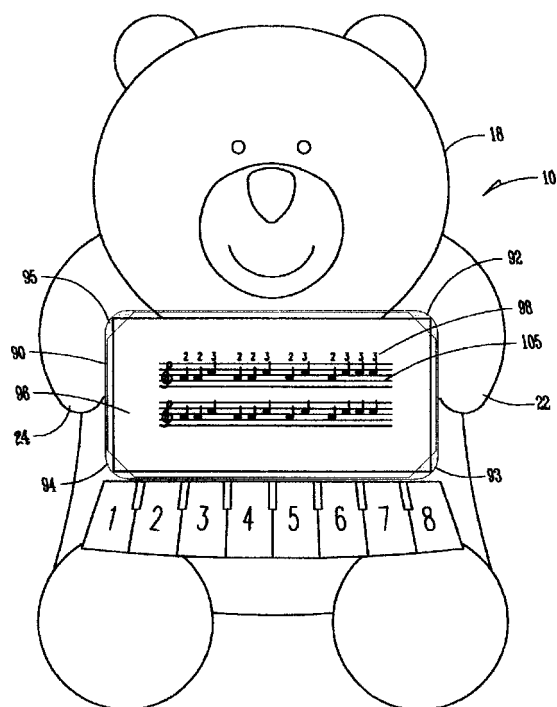
A pair of dolls can communicate status via infrared transducers and so they can speak and hold conversations with each other. Imagine "Chatty Cathy" with a brother. The patent does include a real schematic and seems entirely plausible.—MK

6,454,627

43.38.Md MUSICAL ENTERTAINMENT DOLL

Danny Wai Keung Mak, assignor to Well Creation Limited
24 September 2002 (Class 446/369); filed 30 April 2001

Take a plushy bear, add a numbered, piano-style keyboard, and then



present a score with numbered notes and you have one frustrated 4 year old.—MK

6,457,681

43.38.Md CONTROL, SOUND, AND OPERATING SYSTEM FOR MODEL TRAINS

Michael Paul Wolf *et al.*, assignors to Mike's Train House, Incorporated
1 October 2002 (Class 246/187 A); filed 7 December 2000

The increasing complexity of electronics has benefitted the model train community as well as audio/visual users. First, it was pulse width modulation of the track voltage to control engine motors. Next, the addition of microprocessors allowed addressable control of different engines and other features. This patent first proposes the use of spread spectrum modulation to compensate for the noisy transmission environment. Second, it proposes transmission of digitally encoded sounds and speech to speakers located on the trains. There are other details presented for more "realistic" operation.—MK

6,448,719

43.38.Ne APPARATUS AND METHOD FOR PULSATING LIGHTS IN RESPONSE TO AN AUDIO SIGNAL

Jeffrey G. Denny, San Diego, and Steven D. Karnes, Mission Viejo, both of California
10 September 2002 (Class 315/200 A); filed 8 June 2001

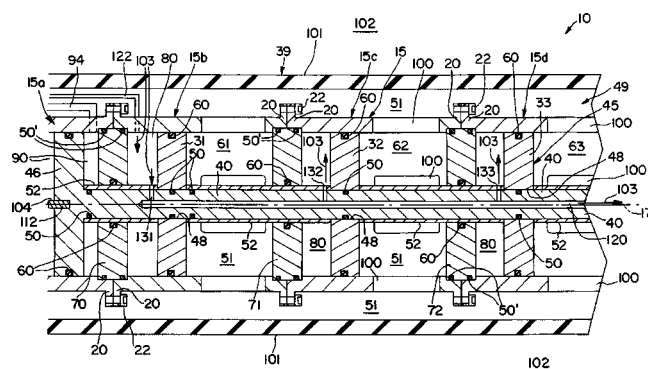
The problem, as stated by the authors, is that existing acoustic control of holiday lighting displays is "monotonous, aesthetically dull, lacking luster, and visual stimulation is absent." The solution is straightforward: use audio filters and triacs. Now, how exactly is this different from existing "color organ" technology?—MK

6,464,035

43.38.Pf STREAMLINED, READILY TOWABLE MARINE SEISMIC ENERGY SOURCE FOR CREATING INTENSE SWEEP-FREQUENCY AND PULSE-CODED SIGNALS IN A BODY OF WATER

Stephen Chelminski, assignor to Bolt Technology Corporation
15 October 2002 (Class 181/120); filed 9 November 2000

The figure is a cross-sectional view of a portion of the middle of a cylindrical shaped, towed, underwater vehicle about 18 in. in diameter and almost 10 ft in length. The cylinder is terminated at both ends by hydrodynamically shaped end caps. Actuator piston 46 is excited into vibration by a remotely controlled hydraulic circuit. A set of vibration pistons 31, 32, 33,... (eight are discussed but more or less are possible) are coupled to piston rod 40, which is attached to piston 46 and which extends along the axis 17 of the vehicle. These output pistons are in turn excited into vibration and radiate into water cavities 61, 62, 63,..., bounded in the axial direction by the pistons and by bulkheads 71, 72,... This sound then communicates with the water-filled annular cavity 51 via a series of holes 100 in cylindrical wall 15 (there are four holes around the circumference at each axial station). Finally, sound is radiated into the surrounding water medium through the elastomeric tube 39 whose surface area comprises a large radiator. Air chambers



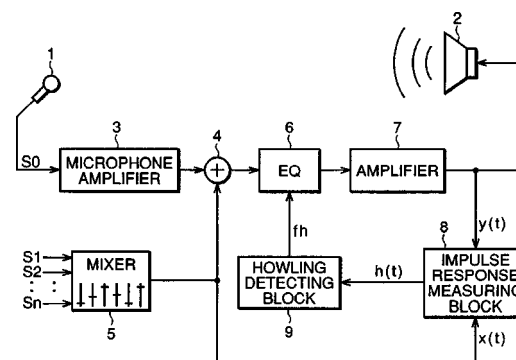
80 are created behind each output piston because of the bulkheads 70, 71, 72,... and these chambers, as well as the hollow head and tail sections of the vehicle, are filled with low-pressure air supplied by an air line from the tow-ship. A set of piston positioning sleeves 52 on the drive rod 40 maintain the output pistons 31, 32, 33,... in their proper axial positions. A position-sensing transducer associated with actuator piston 46 provides a signal so that the displacements of additional units being deployed simultaneously can be synchronized.—WT

6,438,240

43.38.Si CIRCUIT TO IMPROVE TRANSDUCER SEPARATION IN HANDSFREE TELEPHONE

Rob Mcleod, assignor to Mitel Corporation
20 August 2002 (Class 381/71.7); filed 18 February 1997

To minimize direct feedback from the loudspeaker to the microphone of a hands-free telephone, a second microphone is mounted inside the enclosure and its output is subtracted from that of the primary microphone.—GLA



envelope, and identifies the frequency of impending howl. The information is used to update equalization and/or gain. To derive impulse response from the program signal itself, a nonhowling microphone or direct sound source such as an electronic keyboard must be included in the setup.—GLA

6,445,804

43.38.Tj ULTRA-DIRECTIONAL SPEAKER SYSTEM AND SPEAKER SYSTEM DRIVE METHOD

Kouji Hirayanagi, assignor to NEC Corporation
3 September 2002 (Class 381/303); filed in Japan 25 November 1997

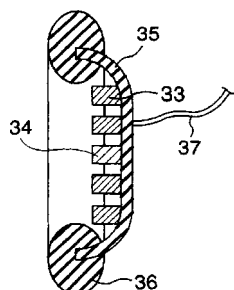
Ultrasonic burglar alarm technology can be used to detect the presence of a listener in a given area and switch the sound system from background music to commercial messages. A beam of modulated ultrasonic energy—a parametric loudspeaker—can generate audible sound in empty air. The patent explains how these two known technologies can be combined to make a self-tracking parametric sound source from which the targeted listener cannot escape, no matter where he moves.—GLA

6,449,370

43.38.Si DIGITAL ELECTRO-ACOUSTIC TRANSDUCER

Yoshinobu Yasuno and Yasuhiro Riko, assignors to Matsushita Electric Industrial Company, Limited; Yasuhiro Riko
10 September 2002 (Class 381/71.6); filed in Japan 16 February 1998

This digital noise-cancelling earphone consists of a small condenser microphone 34 surrounded by a number of equally small sound generating elements 33, all in lightweight ear cup 35. The feedback loop, drive cir-



cuitry, and electroacoustical transduction are all digital. The patent is interesting and informative.—GLA

6,442,277

43.38.Vk METHOD AND APPARATUS FOR LOUDSPEAKER PRESENTATION FOR POSITIONAL 3D SOUND

Charles D. Lueck and Alec C. Robinson, assignors to Texas Instruments Incorporated
27 August 2002 (Class 381/17); filed 19 November 1999

Like several other recent patents, the goal here is to create convincing movement of phantom sound sources from a pair of loudspeakers without

hogging computer memory and processing time. The patent describes an interesting method that allows about half of the required signal processing to be performed offline and stored as a relatively small file. Filtering by head-related transfer functions needs to be performed for only one source position.—GLA

6,444,892

43.38.Vk SOUND SYSTEM AND METHOD FOR CREATING A SOUND EVENT BASED ON A MODELED SOUND FIELD

Randall B. Metcalf, Cantonment, Florida

3 September 2002 (Class 84/723); filed 25 May 2001

This is a continuation of United States Patent 6,239,348 [reviewed in J. Acoust. Soc. Am. **111**(3), 1150 (2002)]. The device is intended to capture and then reproduce the 3D sound field produced by a given sound source—something like an all-electronic, virtual player piano. Some interesting elaborations have been added to the first patent and the number of claims has been increased from 24 to 66. However, only four references are listed, all cited by the examiner. To this reviewer it seems that a fair amount of relevant prior art has been ignored.—GLA

6,445,798

43.38.Vk METHOD OF GENERATING THREE-DIMENSIONAL SOUND

Richard Spikener, Brooklyn, New York

3 September 2002 (Class 381/63); filed 21 January 1998

Each channel of a conventional stereo recording is delayed and recombined with the original. The process is then repeated using different delay values, "...imparting to the sound a three-dimensionality which cannot be achieved by conventional reverberation and echo methods."—GLA

6,439,359

43.40.Tm DAMPER FOR VEHICLES

Rentaro Kato *et al.*, assignors to Tokai Rubber Industries, Limited
27 August 2002 (Class 188/379); filed in Japan 2 September 1998

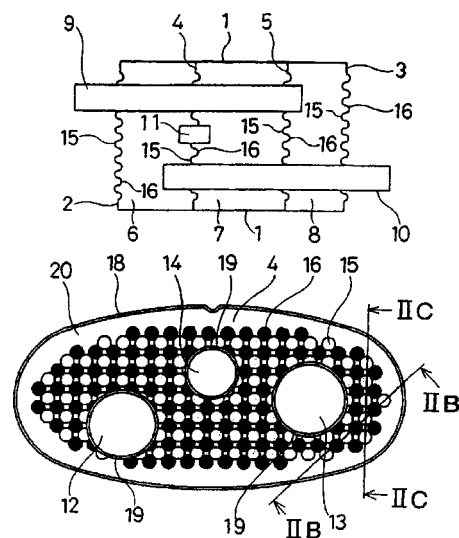
Cylindrical bodies are located in tubular elements of vehicle suspension systems and configured so that these bodies can slide along their lengths and impact against the elements' end walls. Sliding friction and end impacts remove energy from the vibrations of concern and thus provide damping. In other embodiments, tubes that incorporate sliding elements are affixed to vibrating components, such as oil pans and engine heads.—EEU

6,427,802

43.50.Gf MUFFLER

Masayuki Yokoi *et al.*, assignors to Sango Company, Limited
6 August 2002 (Class 181/282); filed in Japan 26 May 1998

Rather than being flat surfaces, the end plates and inner plates of this automobile muffler have an array of alternating convex and concave



dimples. This is said to result in diffuse acoustic reflections and a reduction in high-frequency exhaust noise.—KPS

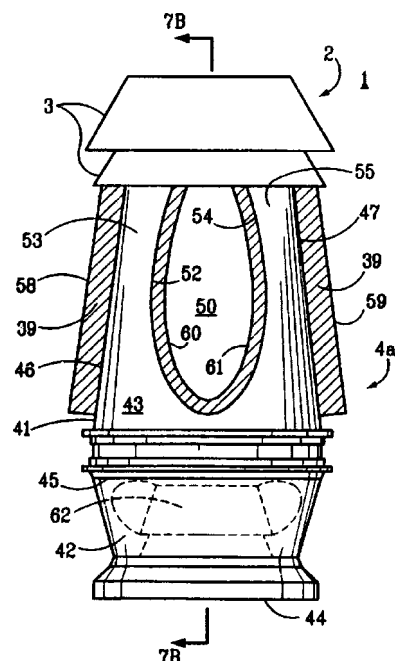
6,431,974

43.50.Gf ACOUSTIC WIND BAND

Paul Antony Tetley and Charles A. Gans, assignors to Met Pro Corporation

13 August 2002 (Class 454/33); filed 29 March 2000

A muffler for use on the exhaust of a building HVAC system consists of sound-absorbing material covered with perforated surfaces 46, 47, 52,



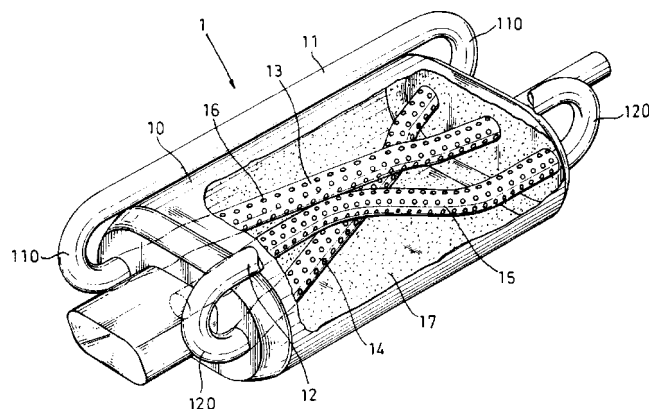
and 54. Furthermore, discharge noise is further reduced by entrainment of ambient air by means of frustum-shaped devices 3.—KPS

6,435,302

43.50.Gf MOTOR VEHICLE MUFFLER

Shun-Lai Chen and Changhwa Hsien, Taiwan, Province of China
20 August 2002 (Class 181/252); filed 21 December 2000

This motor vehicle muffler has intersecting, perforated pipes inside a casing filled with sound absorbing material. Exhaust gases flow in through



tube 13 and proceed through tubes 11, 14, 12, and 15, where they exit the muffler. Much is made of the fact that the internal tubes cross at a central location inside the casing.—KPS

6,435,303

43.50.Gf SOUND ABSORBING STRUCTURE

Glenn E. Warnaka, assignor to Future Technologies LLC
20 August 2002 (Class 181/286); filed 15 January 2000

This compact, lightweight, flat panel structure is a multi-ply sandwich construction. Its interior walls create a series of cavities that act as quarter-wave resonators and/or Helmholtz resonators to absorb sound. The panel is typical of those used in aerospace construction. The sound-absorbing qualities of the flat panel are dependent on its geometry and the panel does not use open cell foam or fibrous materials.—CJR

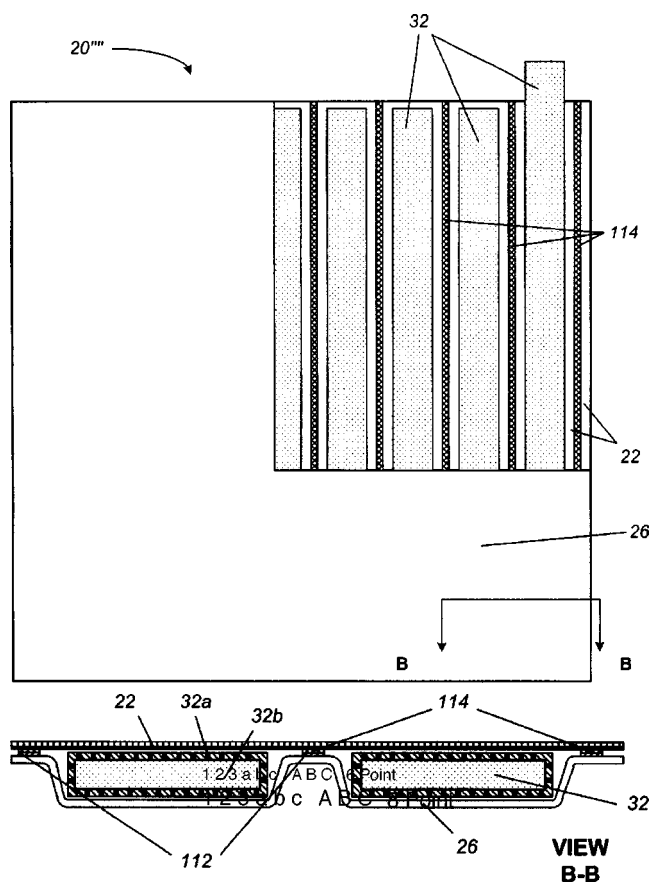
comb shaped core, manufactured from corrugated metallic ribbons. Sound waves directed into the cowl through a perforated face are channeled through the sound receiving chamber.—CJR

6,446,751

43.50.Gf APPARATUS AND METHOD FOR REDUCING NOISE LEVELS

Krishan K. Ahuja *et al.*, assignors to Georgia Tech Research Corporation
10 September 2002 (Class 181/295); filed 5 September 2000

This flexible curtain can both block sound and absorb sound. The curtain has pockets into which either sound insulating or sound absorbing



elements can be inserted. The inserts can be readily removed to permit periodic laundering of the curtain fabric.—CJR

6,439,340

43.50.Gf ACOUSTICALLY TREATED STRUCTURALLY REINFORCED SOUND ABSORBING PANEL

Shawn S. Shirvan, assignor to Astech Manufacturing, Incorporated
27 August 2002 (Class 181/213); filed 17 November 2000

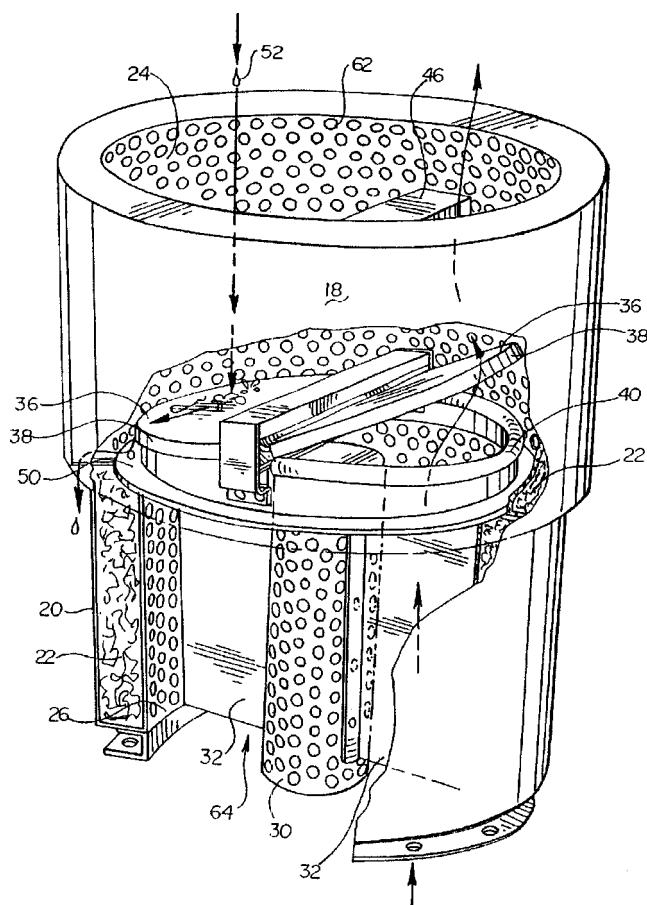
This reinforced sound-absorbing panel is for particular application at the inlet or cowl of a jet engine to reduce engine noise leaked into the environment. The panel includes a sound receiving chamber with a honey-

6,457,550

43.50.Gf WEATHERPROOF SOUND ATTENUATING DEVICE

Charles L. Barry *et al.*, assignors to Twin City Fan Companies, Limited
1 October 2002 (Class 181/224); filed 6 March 2001

This unit efficiently combines several functions: an exhaust silencer (for a duct outlet), a damper (to close the duct when air is not exhausted), and a windband (to prevent wind from pulling the damper open). In this



configuration, the windband is attached to the sound-insulating structure or silencer, and is arranged to prevent water from entering the silencer. The windband also provides additional sound attenuation.—CJR

6,459,578

43.50.Gf CHASSIS HAVING REDUCED ACOUSTIC NOISE AND ELECTROMAGNETIC EMISSIONS AND METHOD OF COOLING COMPONENTS WITHIN A CHASSIS

Guy R. Wagner, assignor to Agilent Technologies, Incorporated
1 October 2002 (Class 361/694); filed 24 April 2001

This chassis for housing electronic components includes a shell featuring an intake duct and an exhaust duct. An intake port directs cooling air into the chassis and an exhaust port directs cooling air out of the chassis. These two ducts include active and/or passive noise-attenuating features for attenuating noise generated within the chassis, thereby promoting quieter operation. Examples of noise attenuation methods include strategic deployment of sound absorption materials and contouring of airflow guides. The ducts are also dimensioned optimally to attenuate electromagnetic radiation generated within the chassis by preventing it from escaping.—DRR

6,466,752

43.50.Gf METHOD OF IMPROVING SOUND QUALITY AND IMAGE FORMATION APPARATUS

Koichi Tsunoda, assignor to Ricoh Company, Limited
15 October 2002 (Class 399/91); filed in Japan 16 October 2000

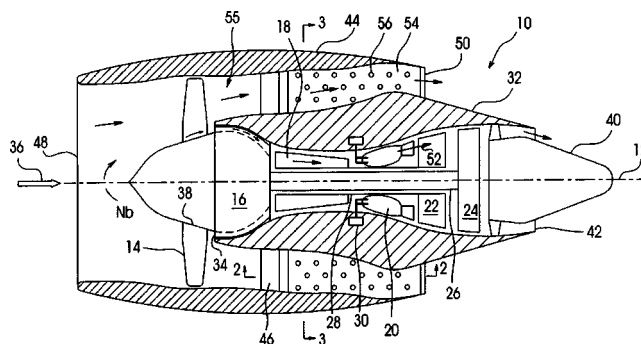
As anyone who has unjammed a copier knows, today's copiers have increasingly complex paper paths. As the paper makes its way toward the output tray, it traverses many rollers and belts and this produces noise. The patent begins by describing a numerical evaluation function and then shows how design of noise-isolating plates can decrease this function.—MK

6,439,840

43.50.Nm BYPASS DUCT FAN NOISE REDUCTION ASSEMBLY

Man-Chun Tse, assignor to Pratt & Whitney Canada Corporation
27 August 2002 (Class 415/119); filed 30 November 2000

A method to attenuate fan noise that propagates downstream through the aft fan duct of a turbofan aircraft engine consists of the placement of one or more perforated baffle plates 54. The purpose of these plates is to create multiple axial flow paths, thus destroying the symmetry and disrupting the



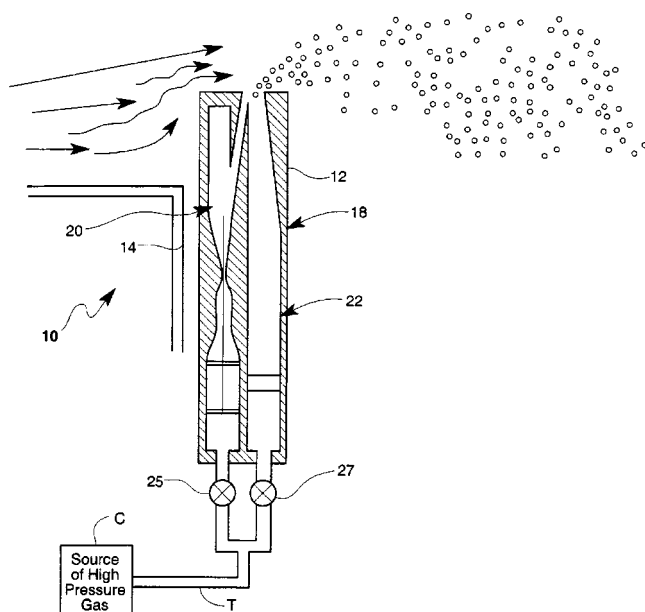
spinning acoustic modes. These perforated baffle plates have 20%–50% open area and the hole diameters are between 0.5 and 2 times the plate thickness.—KPS

6,446,904

43.50.Nm AIRCRAFT WEAPONS BAY HIGH FREQUENCY ACOUSTIC SUPPRESSION APPARATUS

Michael J. Stanek, assignor to the United States of America as represented by the Secretary of the Air Force
10 September 2002 (Class 244/1 N); filed 5 October 2001

A means to suppress flow-induced acoustic resonances in an aircraft weapons bay consists of deployable spoiler 12, which serves to generate high-frequency pulses of pressurized gas. The spoiler has two primary components, nozzle 22, supplied with pressurized air through valve 27, and reso-



nator 20, supplied with pressurized air through valve 25. The resonator serves to perturb the flow from the nozzle, thus creating high-frequency pulses which suppress the acoustic resonances within the weapons bay.—KPS

6,443,256

43.55.Ev DUAL LAYER ACOUSTICAL CEILING TILE HAVING AN IMPROVED SOUND ABSORPTION VALUE

Mirza A. Baig, assignor to USG Interiors, Incorporated
3 September 2002 (Class 181/286); filed 27 December 2000

This ceiling tile has two layers of acoustical materials formed together as one. The base mat layer has either no or very little mineral wool. The surface layer has a high mineral wool content, which provides improved sound absorption values, with or without perforating or fissuring the tile. The result is a panel with NRC value of at least 0.50. The material is made by a water felting process to form both the base mat layer and the fiber-rich surface layer together, in a quick manufacturing process.—CJR

6,457,554

43.55.Ev ACOUSTIC ABSORBER FOR ABSORBING NOISES IN BUILDINGS

Samw Hong Jen Wang, Taichung, Taiwan, Province of China
1 October 2002 (Class 181/282); filed 15 March 2001

This absorber would be part of a wall for reducing sound in a building, but it does not use fibrous materials that would get wet. The housing includes a casing that has openings for receiving and dissipating the noise.—CJR

6,457,284

43.55.Vj STRUCTURE FOR INSTALLING A VISCOUS VIBRATION DAMPING WALL AND METHOD OF INSTALLING THE SAME

Kazuhiko Isoda, assignor to Shimizu Corporation
1 October 2002 (Class 52/167.1); filed in Japan 2 September 1998

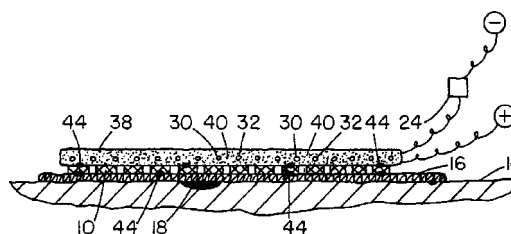
This patent describes a simplified vibration-damping wall unit and its installation that purports to be of lower cost and quicker installation than other viscous-damping systems. By using friction-type high-strength bolt connections, various intermediate plates can be eliminated and the thickness of other plates can be reduced.—CJR

6,445,304

43.58.Wc MEDICAL ALARM SYSTEM

John J. Bandean, Jr. and John J. Bandean III, both of Bristol, Tennessee
3 September 2002 (Class 340/604); filed 11 August 2000

A bleeding or oozing wound site 18 can be covered with a conductive



bandage that will generate an alarm by energizing output transducer 24.—MK

6,453,599

43.58.Wc FISHING LURE WITH SOUND ATTRACTION

Richard S. Mathews and Robert G. Link, assignors to Pop-A-Long LLC
24 September 2002 (Class 43/42.31); filed in Canada 26 May 2000

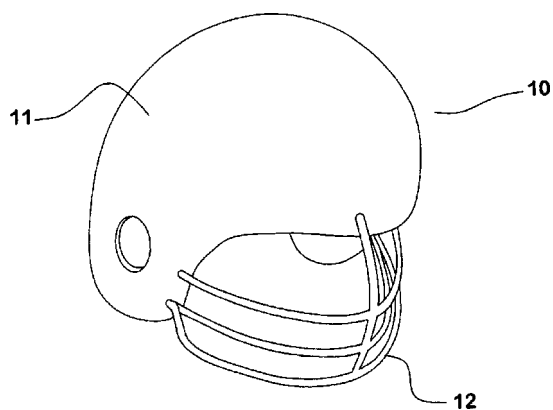
In the long running contest between man and fish, this device uses a striking plate so that clicking or popping sounds are created. Additionally, "the sliding movement of the aft section with respect to the fore section as the lure is pulled through the water produces a 'wounded prey' effect."—MK

6,464,558

43.58.Wc SPORTS RELATED COMBINED NOISEMAKER AND VISUAL DISPLAY DEVICE

Ronald W. Ayres, Tyler, Texas
15 October 2002 (Class 446/419); filed 13 October 2000

Quite simply, this is a shaker/maracas for use in sporting events. A



hollow cavity inside a helmet provides space for vibrating beads or pellets.—MK

6,453,281

43.60.Bf PORTABLE AUDIO DATABASE DEVICE WITH ICON-BASED GRAPHICAL USER-INTERFACE

Timothy L. Walters *et al.*, assignors to VXI Corporation
17 September 2002 (Class 704/200); filed 30 July 1996

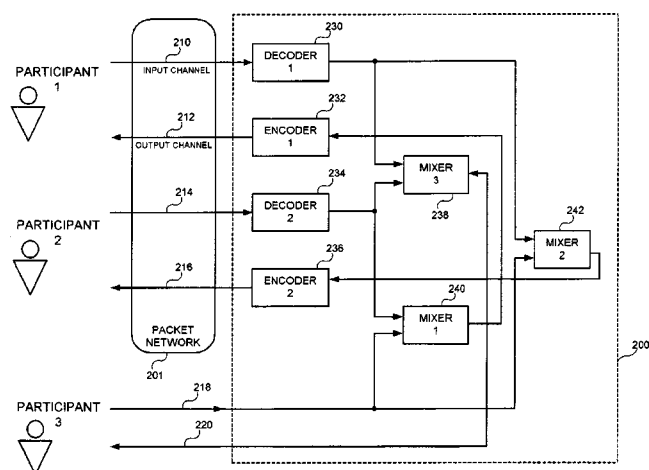
This patent describes a method for storing and retrieving information to and from an external audio database device. The device includes a graphics display in a hand-held housing. A speaker and microphone relay audio information to and from a user. The hand-held audio database device includes a memory which stores graphic icons and supports a hierarchical memory structure comprising categories denoted by the graphic icons. A user-actuated navigation control permits the user to navigate the categories in the hierarchical memory structure and to select a desired category. A processor controls the display during use and acts to store the audio information in the desired category of the memory.—DRR

6,463,414

43.60.Qv CONFERENCE BRIDGE PROCESSING OF SPEECH IN A PACKET NETWORK ENVIRONMENT

Huan-Yu Su *et al.*, assignors to Conexant Systems, Incorporated
8 October 2002 (Class 704/270.1); filed 12 April 2000

The device relates, generally, to the transmission of speech using voice-over-packet networks and, more particularly, to techniques for improving voice-over-IP (VoIP) conference bridges and transcoders. The device provides a conference bridge or transcoder configured to intelligently handle multiple speech channels in the context of a packet network, wherein the various speech channels may use a variety of speech encoding standards.



In general, the conference bridge establishes framing and alignment of multiple incoming speech channels associated with multiple participants, extracts parameters from the speech samples, mixes the parameters, and reencodes the resulting speech samples for transmission back to the participants. Priority assignment and speech enhancement (e.g., noise reduction, reshaping, etc.) are also performed.—DRR

6,454,718

43.64.Nf INTRA AURAL INTEGRATED VITAL SIGNS MONITOR

Vaughan L. Clift, Houston, Texas
24 September 2002 (Class 600/483); filed in Australia 10 November 1997

This intra-aural probe for monitoring multiple physiological parameters simultaneously includes a pressure sensor, a temperature sensor, and a light sensor. The probe is dimensioned to fit snugly within the outer aural cavity. The probe measures blood pressure, pulse rate, and volume of blood by detecting contraction and expansion in the vascular bed of the lining tissue of the external auditory canal in the course of a cardiac cycle. The probe measures respiratory rate and volume by detecting pressure changes across or movement of the tympanic membrane induced by corresponding pressure changes in the naso-pharynx and trachea during respiration. The probe measures temperature in the auditory canal. It also measures electroencephalographic or electrocardiographic voltages by detecting small potential differences between the external auditory canal region and a suitable reference position, measures electroencephalographic currents by detecting magnetic flux, and measures changes in oxygen saturation of the hemoglobin traveling in the capillary bed lining the external auditory canal.—DRR

6,460,489

43.64.Ri DEVICE FOR HANDLING LIVESTOCK USING VIBRATION AND NOISE AS A STIMULATION ON EXTERNAL PORTIONS OF THE BODY

Timothy O'Byrne, Consort, Alberta, Canada
8 October 2002 (Class 119/908); filed 17 April 2001

This device provides an alternative to the electric livestock prods used to administer electric shocks to an animal. The preferred embodiment of the device is battery powered and adapted to be easily hand held by the user. The distal end of the prod, to be applied to the animal, contains an electric motor that generates a vibration and/or sound that provides a stimulus causing the animal to move on. A button activates the motor to initiate the prodding.—DRR

6,463,408

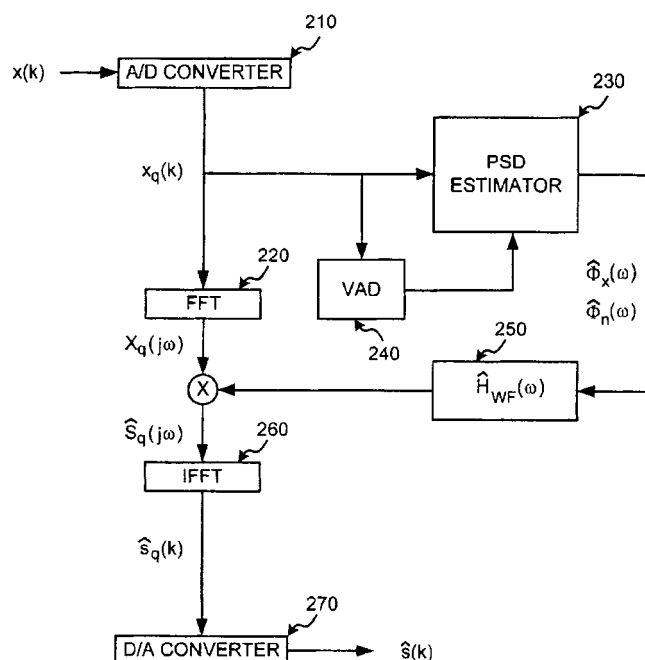
43.72.Ew SYSTEMS AND METHODS FOR IMPROVING POWER SPECTRAL ESTIMATION OF SPEECH SIGNALS

Leonid Krasny and Soontorn Oraintara, assignors to Ericsson, Incorporated

8 October 2002 (Class 704/217); filed 22 November 2000

This patent describes a method to improve the power spectrum estimation of speech signals. Power spectral density (PSD) estimator 230 estimates both mixed and noise signals and uses the control signal from voice

200 →



activity detector 240 to separate the noise-only PSDs. Wiener filter 250 includes the logic to perform spectral subtraction. Finally, inverse FFT 260 converts the noise-reduced speech signals into a bit stream for transmission.—HHN

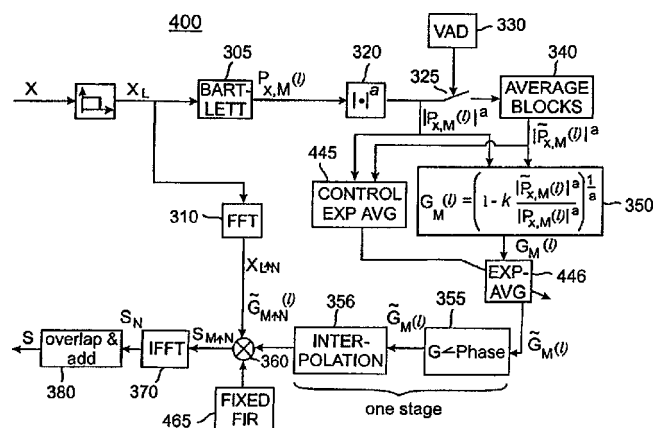
6,459,914

43.72.Ew SIGNAL NOISE REDUCTION BY SPECTRAL SUBTRACTION USING SPECTRUM DEPENDENT EXPONENTIAL GAIN FUNCTION AVERAGING

Harald Gustafsson *et al.*, assignors to Telefonaktiebolaget LM Ericsson (publ)

1 October 2002 (Class 455/570); filed 27 May 1998

This patent relates to speech enhancement by means of a spectral subtraction technique using linear convolution, causal filtering, and controlled exponential averaging 305. Successive blocks of a spectral subtrac-



tion gain function are averaged based on the difference between an estimated power spectral density of the mixed signals and an averaged estimate of the noise. Interpolation processor 356 provides a smoothed, noise-reduced speech signal.—HHN

6,453,041

43.72.Ne VOICE ACTIVITY DETECTION SYSTEM AND METHOD

Erol Eryilmaz, assignor to Agere Systems Guardian Corporation
17 September 2002 (Class 379/392.01); filed 15 October 1998

This device can be used in a half-duplex speakerphone which operates in transmit, silence, and receive modes. Depending on detection of voice activity, the speakerphone switches from silence to either transmit or receive mode. For switching between different modes of operation, the patent presents a new voice activation detector which utilizes the integral of absolute values of the signal derivatives. The voice activity is detected in both transmission and receive paths by comparing the current voice energy with the background noise energy. During transmission (or receive) the direction of speakerphone operation can be changed by comparing the transmit energy with received voice energy (or receive with transmit). After detecting the direction of voice activity, the speakerphone begins transitioning to the desired mode.—AAD

6,453,285

43.72.Ne SPEECH ACTIVITY DETECTOR FOR USE IN NOISE REDUCTION SYSTEM, AND METHODS THEREFOR

David V. Anderson *et al.*, assignors to Polycom, Incorporated
17 September 2002 (Class 704/210); filed 10 August 1999

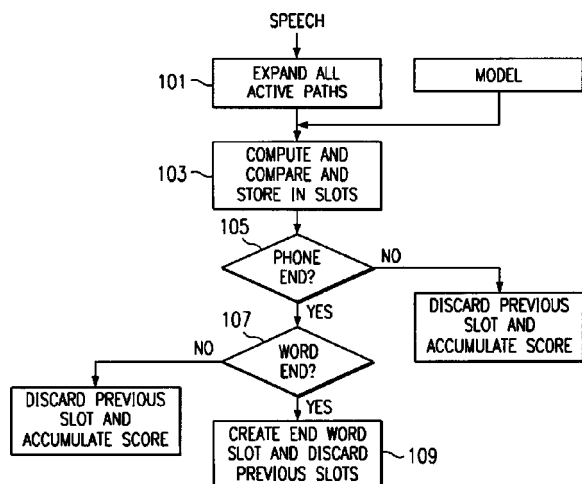
This patent introduces a method to detect voice activity in noisy speech signals. The method first extracts some statistics from the spectrum of the noisy signal, to use as features. To detect voice activation, a finite state machine makes a decision based on previous and current states of the signal. According to the patent, the noise and signal power spectra are estimated using an adaptive Wiener filter. Short time features are then extracted from the filtered signals which are the inputs to the state machine. The patented method can be used in noise reduction applications as well.—AAD

6,456,970

43.72.Ne MINIMIZATION OF SEARCH NETWORK IN SPEECH RECOGNITION

Yu-Hung Kao, assignor to Texas Instruments Incorporated
24 September 2002 (Class 704/256); filed 15 July 1999

Each word in this speech recognition system is represented by a Markov model with states and transitions among states. Acoustical characteristics are associated with each state and transitions represent characteristics while transitioning through a word at normal rates of speaking. A search network consisting of input acoustic vectors is constructed by knowing what the present state is and to which state(s) the next transition(s) will go. The



object is to find the best path through the search space. The search network is expanded for each speech frame by expanding phonetic, phone, lexical, sentence, and word hidden Markov models. Slots, or search network building blocks, define what state one is at in the search network. At least one slot is created in the search path for each acoustic vector. The algorithm is optimized for minimal peak slot usage by pruning unnecessary slots based on unneeded information. Optimization is said to reduce peak slot usage by about 20%.—DAP

6,466,908

43.72.Ne SYSTEM AND METHOD FOR TRAINING A CLASS-SPECIFIC HIDDEN MARKOV MODEL USING A MODIFIED BAUM-WELCH ALGORITHM

Paul M. Baggenstoss, assignor to the United States of America as represented by the Secretary of the Navy
15 October 2002 (Class 704/256); filed 14 January 2000

A training method is proposed to reduce the number of data samples needed to train hidden Markov models (HMMs) used for modeling physical characteristics in speech recognition systems. Estimating parameters of class-specific HMMs from training data is accomplished by iterating a modified version of the Baum-Welch algorithm using likelihood ratios that compare each of the states to a common state. Each iteration of the modified Baum-Welch algorithm includes class-specific forward and backward procedures to calculate forward and backward probabilities, respectively, and

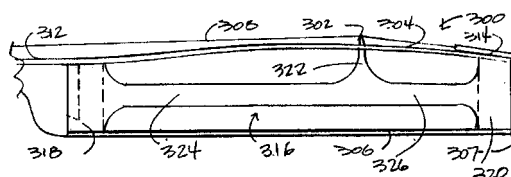
HMM reestimation formulas for updating the parameters of the class-specific HMM based on these probabilities.—DAP

6,459,024

43.75.Gh STRUCTURAL TORSION BRACE FOR AN ACOUSTIC MUSICAL INSTRUMENT

James R. Baker, assignor to James R. Baker
1 October 2002 (Class 84/291); filed 19 September 1997

This device is intended to eliminate or reduce electronic feedback in amplified acoustic string instruments due to top plate vibration. Accordingly,



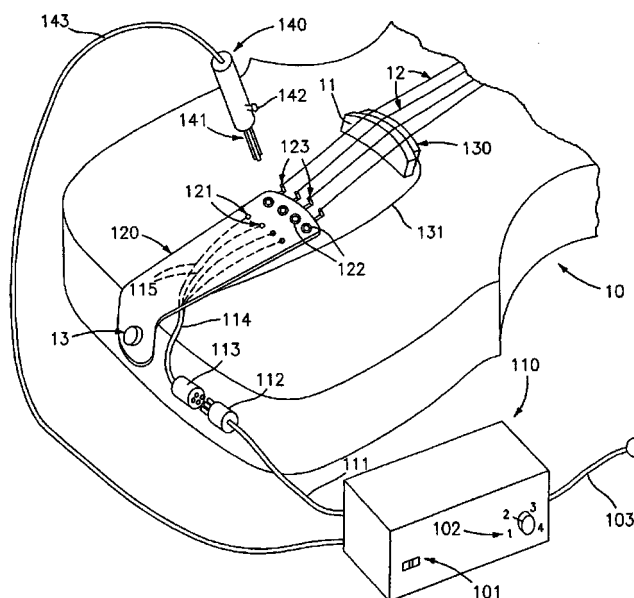
a brace is introduced that attaches sparingly to the top plate as shown in the figure. However, given the influence of plate vibration on instrumental timbre, the full impact of this design has not been fully assessed.—MK

6,465,723

43.75.Gh AUTOMATIC STRING INSTRUMENT TUNER KIT

Lynn M. Milano, Sound Beach, New York
15 October 2002 (Class 84/454); filed 7 March 2001

Continuing in the automatic tuning vein, the author proposes using



tuning wrench 141 that turns socket head 122. The pitch of the string can be sensed by pickup 121 that is fed back to controller 110.—MK

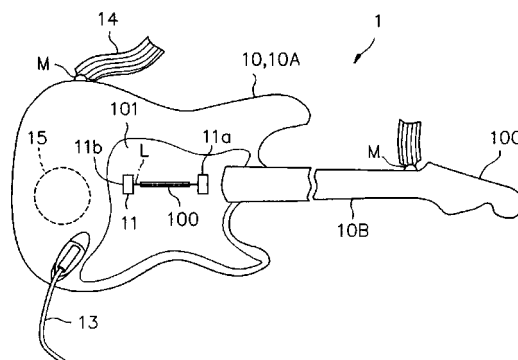
6,452,079

43.75.Mn KEYBOARD MUSICAL INSTRUMENT EQUIPPED WITH HAMMER STOPPER PROMPTLY DRIVEN FOR ROTATION BY MEANS OF RIGID LINK WORK

Nobuo Sugiyama and Shigeaki Sato, assignors to Yamaha Corporation

17 September 2002 (Class 84/423 R); filed in Japan 28 April 2000

The device described is a modification of the piano action so that the player can silence the acoustic tones and use electronic tones instead and "not disturb the neighborhood."—MK



infrared light beam **100** between emitter **11a** and sensor **11b**. However, in spite of the claims to the contrary, it is no doubt "less natural to play" than a real stringed instrument.—MK

6,452,081

43.75.St STRINGED INSTRUMENT FINGER POSITIONING GUIDE AND METHOD FOR TEACHING STUDENTS TO READ MUSIC

Steven F. Ravagni *et al.*, all of Issaquah, Washington

17 September 2002 (Class 84/477 R); filed 15 May 2001

The authors propose using a thin polymer with light emitters to guide the fingers of guitar novices to the proper place on the keyboard—nothing more, nothing less.—MK

6,459,029

43.75.St METHOD AND APPARATUS FOR DISPLAYING MULTIPLE IMAGES OF MUSICAL SCORE DATA ON SCREEN OF DISPLAY

Haruki Uehara, assignor to Yamaha Corporation

1 October 2002 (Class 84/477 R); filed in Japan 21 June 2000

Imagine an LCD display mounted above the keyboard on a piano. Now as the score pointer advances from top to bottom, the background color will change for each staff. When the score pointer reaches the end of the "page," the color selection will start from the top again.—MK

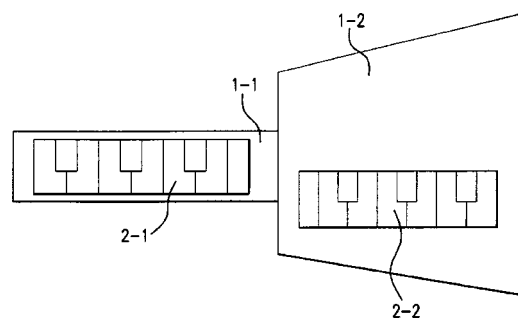
6,444,891

43.75.Tv ELECTRONIC GUITAR WITH ITS KEYS ARRANGED IN COMPLEX ARRAY

Po Wo Koo, Hong Kong, Hong Kong Special Administration of the People's Republic of China

3 September 2002 (Class 84/684); filed in China 9 November 2000

This patent is based on the premise that karaoke singers can play a



piano keyboard—therefore, the central idea is to put black and white keys on a guitar form. After all, you can't rock without a guitar in hand.—MK

6,459,028

43.75.Wx PERFORMANCE DATA MODIFYING METHOD, PERFORMANCE DATA MODIFYING APPARATUS, AND STORAGE MEDIUM

Akira Yamauchi, assignor to Yamaha Corporation

1 October 2002 (Class 84/477 R); filed in Japan 21 March 2000

The issue here is naturalness in a synthetic performance. If the score contains ornaments such as turns, trills, and mordants, then the underlying MIDI score must be modified accordingly.—MK

6,444,887

43.75.Tv STRING-PLUCKING TYPE ELECTRONIC MUSICAL INSTRUMENT WITH PHOTO SENSOR FOR GENERATING SIGNAL FOR SOUND

Tadashi Hiraoka and Isamu Kubota, assignors to Namco Limited
3 September 2002 (Class 84/600); filed in Japan 16 September 1999

In a gaming guitar for use in man-machine contests, a robust stringless user interface must be offered to the user. Here, the player interrupts an

6,459,030

43.75.Wx METHOD FOR PREVENTING POLYPHONY SHORTAGE IN AN ELECTRONIC ORGAN

Marco Di Paolo *et al.*, assignors to Roland Europe S.p.A.

1 October 2002 (Class 84/620); filed in Italy 27 April 2001

Implementing a multivoice/polyphonic synthesizer on a DSP microprocessor will eventually result in oscillator starvation as the processor can execute only a fixed number of oscillators for a given sampling rate. The

patent is geared towards organs and so the recovery strategy is specific about eliminating stops according to a priority: first comes Flute Celeste II, last comes Blockflöte 2'.—MK

6,464,585

43.75.Wx SOUND GENERATING DEVICE AND VIDEO GAME DEVICE USING THE SAME

Shigeru Miyamoto *et al.*, assignors to Nintendo Company, Limited
15 October 2002 (Class 463/35); filed in Japan 20 November 1997

The mind of Nintendo is partially exposed in this gaming patent. It describes how to use a typical gaming joystick controller to construct a sequence of musical notes. It includes flowcharts and ROM layout for those who wish to understand the internals.—MK

6,450,128

43.80.Nd BIRD TRAINING METHOD AND APPARATUS THEREFOR

Mark A. Boyce, Gilbert, Arizona
17 September 2002 (Class 119/713); filed 15 May 2001

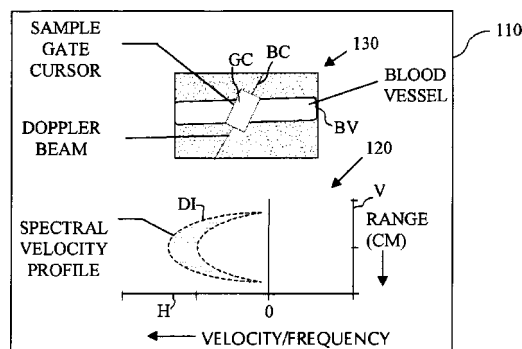
The device is designed to train birds (generally of the parrot species) to refrain from uttering preselected vocalizations. It functions by locating a bird in an illuminated control area, capturing a sound uttered by the bird to produce an audio signal, determining if the sound is the preselected vocalization, and, upon the occurrence of the preselected vocalization, triggering a timing unit to automatically darken the control area for a specified time interval.—DRR

6,450,959

43.80.Qf ULTRASOUND B-MODE AND DOPPLER FLOW IMAGING

Larry Y. L. Mo *et al.*, assignors to GE Medical Systems Global Technology Company
17 September 2002 (Class 600/441); filed 23 March 2000

This ultrasound scanning system contains multiple ultrasound range gates for generating Doppler signal samples that represent different increments of depth within a subject undergoing study. A logic unit generates



Doppler frequency signals and also B-mode data. A display provides a B-mode image upon which a Doppler image may be superimposed. The Doppler image is arranged to illustrate depth increments within the target area versus Doppler velocity or frequency.—DRR

6,453,201

43.80.Qf IMPLANTABLE MEDICAL DEVICE WITH VOICE RESPONDING AND RECORDING CAPACITY

Douglas R. Daum *et al.*, assignors to Cardiac Pacemakers, Incorporated
17 September 2002 (Class 607/62); filed 28 December 1999

An implantable medical device such as a cardiac pacemaker or a cardioverter/defibrillator is equipped with the capability of receiving and acting on communications in the form of speech spoken by the patient. The device includes an acoustics transducer, which, along with associated filtering circuitry, enables the voice communication to be used to affect the operation of the device or recorded for later playback.—DRR

6,454,714

43.80.Qf ULTRASONIC HARMONIC FLASH SUPPRESSION

Gary Ng and James R. Jago, assignors to Koninklijke Philips Electronics N.V.
24 September 2002 (Class 600/443); filed 20 October 2000

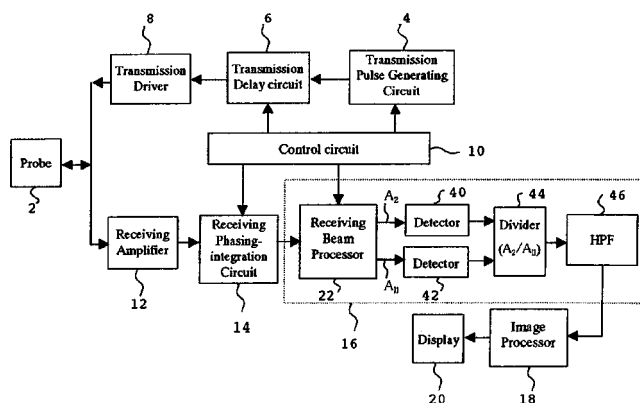
A method is described for reducing the flash artifact of ultrasonic harmonic images. Harmonic images are separated by pulse inversion separation which uses multiple transmit pulses that may be subject to motion artifacts. The motion artifacts are detected and subtracted from the harmonic signals to produce harmonic images with reduced flash artifacts. The motion artifacts may also be reduced by notch filtering. In another embodiment, the amount of motion in the image is detected and the flash artifact is reduced in accordance with the detected motion. The amount of artifact signal that is removed is variable in accordance with anticipated image motion or clinical application.—DRR

6,458,084

43.80.Qf ULTRASONIC DIAGNOSIS APPARATUS

Jing-Wen Tsao *et al.*, assignors to Aloka Company, Limited
1 October 2002 (Class 600/443); filed in Japan 17 February 2000

In this ultrasonic apparatus, one transmission signal includes two fundamental frequencies, f_0 and $2f_0$. A second transmission signal then has a polarity reverse that of the first transmission signal. A receiving beam processor 22 generates a signal in which the echoes are subtracted. The fundamental components are canceled out by addition and a secondary harmonic component A_2 generated by the nonlinear interactions remains in the sum



signal. In the subtraction signal, the fundamental component remains while the secondary harmonic component is canceled out by subtraction. The fundamental A_{II} of the frequency $2f_0$ is extracted from the subtraction signal and a ratio (A_2/A_{II}) is calculated by a divider 44. High-pass filter 46 then indicates, by differentiation, the nonlinear parameter at each depth into the body.—DRR

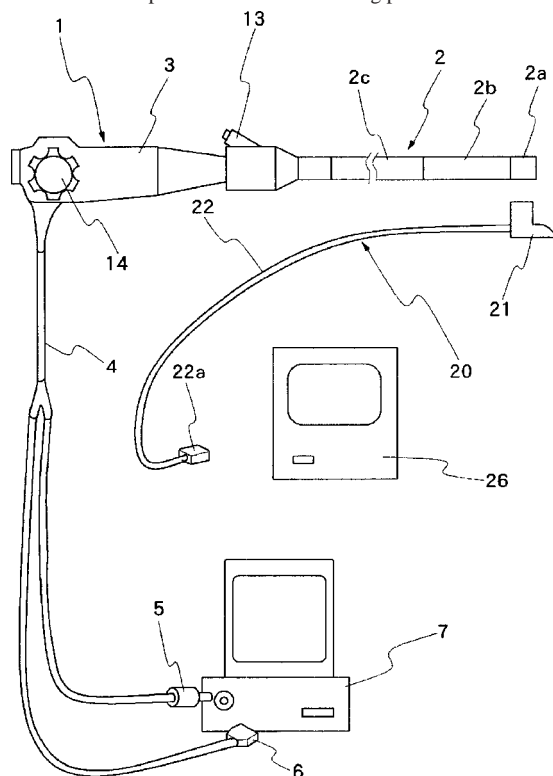
6,461,304

43.80.Qf ULTRASOUND INSPECTION APPARATUS DETACHABLY CONNECTED TO ENDOSCOPE

Toshizumi Tanaka *et al.*, assignors to Fuji Photo Optical Company, Limited

8 October 2002 (Class 600/462); filed in Japan 30 March 1999

The aim of this apparatus is to provide ultrasonic scanning to supplement the mechanical scanning of an endoscope. The ultrasonic inspection apparatus consists of an ultrasound transducer in the form of a rectangular array of transducer chips. The ultrasonic scanning portion is detachable from



the endoscope so as to protrude ahead by a predetermined amount. In order to position the ultrasonic scanning portion in a fixed manner, an endoscope-placement segment is connected to the base end of the ultrasonic scanner.—DRR

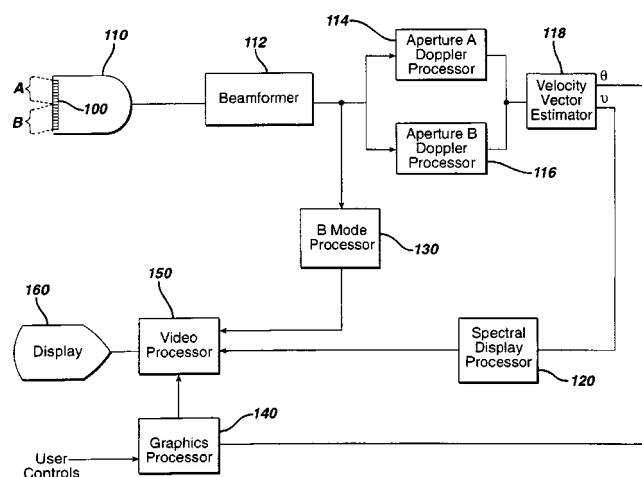
6,464,637

43.80.Qf AUTOMATIC FLOW ANGLE CORRECTION BY ULTRASONIC VECTOR

Aline Laure Criton and Helen Frances Routh, assignors to Koninklijke Philips Electronics N.V.

15 October 2002 (Class 600/441); filed 21 November 2000

An ultrasonic system produces an image that includes a blood vessel. A blood flow direction indicator is displayed over the blood vessel to indicate the direction of the flow within the vessel. The direction of the blood



flow is used to correct the Doppler estimate for angle of insonation. A vector processor automatically sets the orientation of the blood flow indicator. The automatically calculated angle is used to display the correct flow velocity without user intervention.—DRR

6,450,975

43.80.Sh ULTRASONIC TRANSMISSION GUIDE WIRE

Lawrence Brennan *et al.*, assignors to Advanced Cardiovascular Systems, Incorporated

17 September 2002 (Class 600/585); filed 30 December 1999

The device constitutes a guide wire configured to conduct ultrasound energy for use in the ultrasonic treatment of lesions located in hollow anatomical structures. This angioplasty transmission guide wire has regions of reduced cross section diameter. This is said to improve flexibility and to compensate for degradation of longitudinal displacement due to acoustic losses along the length of the wire. The guide wire has a core of ultrasonic transmission material and an outer jacket of shrink tubing surrounding a usable portion at the distal end of the elongated shaft. Threads at the proximal end accept a connecting ultrasonic transducer. The distal end consists of a smooth ball tip for atraumatic application of ultrasound energy to biological tissue.—DRR

6,451,013

43.80.Sh METHODS OF TONSIL REDUCTION USING HIGH INTENSITY FOCUSED ULTRASOUND TO FORM AN ABLATED TISSUE AREA CONTAINING A PLURALITY OF LESIONS

F. Barry Bays and James B. Hissong, assignors to Medtronic Xomed, Incorporated

17 September 2002 (Class 606/27); filed 31 July 2000

This method of tonsil reduction by thermal ablation uses high-intensity focused ultrasound energy. An ultrasound emitter is introduced into a patient's oral cavity and positioned next to an external surface of one of the patient's tonsils. Ultrasound energy from the transducer is emitted into the tissue of one tonsil and focused to ablate the tonsil tissue. A series of ablated lesions are created and the ablated tissue is either surgically removed or is allowed to remain in the patient's body. The lesions end at a predetermined depth so that muscular tissue is not damaged.—DRR

6,454,713

43.80.Sh ULTRASOUND THERAPEUTIC APPARATUS

Yoshiharu Ishibashi *et al.*, assignors to Kabushiki Kaisha Toshiba
24 September 2002 (Class 600/439); filed in Japan 31 March 1995

This therapeutic device for treating diseased tissue consists of an ultrasonic wave generator to generate therapeutic ultrasonic waves and an *in vivo* imaging probe for obtaining a tomographic image in the focusing region. The imaging probe receives echoes of the pulses emitted by the transducer. The driving conditions for the wave generating device are adjusted on the basis of the received echo signal that contains information about the actual intensity of the applied therapeutic waves within the body.—DRR

6,454,730

43.80.Sh THERMAL FILM ULTRASONIC DOSE INDICATOR

Dennis Hechel *et al.*, assignors to Misonix Incorporated
24 September 2002 (Class 601/2); filed 2 April 1999

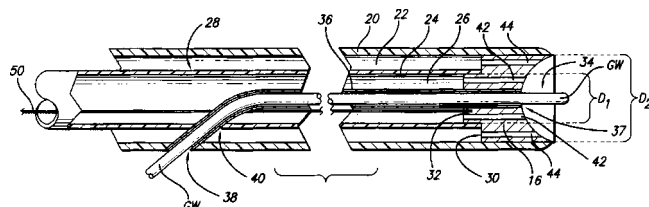
In the application of applying ultrasound treatment to a portion of the human body, the body portion will experience a temperature rise when an ultrasound dosage limit has been reached. The method described here involves deploying on the body portion an indicator adapted to provide a visual change at the determined temperature and applying ultrasound to the body until the indicator provides the specified visual change.—DRR

6,454,737

43.80.Sh ULTRASONIC ANGIOPLASTY-ATHERECTOMY CATHETER AND METHOD OF USE

Henry Nita and Timothy Mills, assignors to Advanced Cardiovascular Systems, Incorporated
24 September 2002 (Class 604/22); filed 25 June 1999

The device is designed to remove obstructive matter from an anatomical structure or passageway by ultrasonic means. It consists of an elongated pliable catheter with an attached distal tip and an ultrasound transmitter that extends longitudinally to convey ultrasound to the distal tip. The tip has a



concave indentation in the distal surface and at least one passageway extending inside the catheter to communicate with a catheter lumen. This facilitates suctioning of severed matter that comes in contact with the concave indentation. The device may also incorporate means for infusing fluids (e.g., irrigation fluid or medications) separately or concurrently with the aspiration of the severed obstruction matter.—DRR

6,454,757

43.80.Sh ULTRASONIC METHOD FOR ABLATING AND REMOVING OBSTRUCTIVE MATTER FROM ANATOMICAL PASSAGEWAYS AND BLOOD VESSELS

Henry Nita and Timothy C. Mills, assignors to Advanced Cardiovascular Systems, Incorporated
24 September 2002 (Class 604/517); filed 2 September 1999

This is a device intended for treating obstructions of tubular anatomical passageways of the urinary tract, including the urethra and fallopian tubes. It involves inserting an ultrasonic catheter into the passageway and passing ultrasound to the distal end of the catheter. Liquid is infused into the distal end and expelled through lateral fluid outflow apertures at a rate that will cause fluidic dilation of the anatomical passageway.—DRR

6,458,143

43.80.Sh ULTRASONIC TREATMENT INSTRUMENT FOR MEDICAL OPERATION

Toshiya Sugai, assignor to Olympus Optical Company, Limited
1 October 2002 (Class 606/169); filed in Japan 4 October 1999

This is an ultrasonic device to be used in medical operations. The device includes a handpiece having a cover that serves as a grip and an ultrasonic transducer is housed within the cover. A vibration-transmitting member connected to the transducer extends from the handpiece in order to convey the ultrasonic energy. A treatment portion is provided at the distal end of the transmitter for the purpose of treating living body tissues. The transmitting member can be bent in a number of places so that the treatment portion can be positioned at or near the central axis of the handpiece.—DRR

6,461,586

43.80.Sh METHOD OF MAGNETIC RESONANCE FOCUSED SURGICAL AND THERAPEUTIC ULTRASOUND

Evan C. Unger, assignor to Imarx Therapeutics, Incorporated
8 October 2002 (Class 424/9.32); filed 10 July 2000

This involves a method of combining magnetic resonance imaging and focused surgical/therapeutic ultrasound. A contrast medium for magnetic resonance imaging is administered, introducing gas-filled vesicles into a patient undergoing surgery. The contrast medium is used to scan the patient with magnetic resonance imaging to identify the surgical work region. Ultrasound is then applied to carry out the surgery. The ultrasound application may be simultaneous with magnetic resonance imaging and the scanning and surgical steps may be performed repeatedly until the desired effect is achieved. The gas-filled vesicles may contain a therapeutic agent to be released in a localized region upon application of the ultrasound.—DRR

6,464,660

43.80.Sh BALLOON CATHETERS HAVING ULTRASONICALLY DRIVEN INTERFACE SURFACES AND METHODS FOR THEIR USE

Axel Briskin and Vartan Ghazarossian, assignors to Pharmasonics, Incorporated
15 October 2002 (Class 604/22); filed 14 March 2001

A catheter consists of a catheter body having an oscillating driver, an interface surface mechanically coupled to the driver, and an inflatable balloon deployed near the interface surface. The balloon may serve as an angioplasty balloon, in which case the interface surface will deliver ultrasonic or other vibratory energy into a blood vessel as part of an angioplasty or

related procedure. Alternatively, the catheter may include a pair of axially separated isolation balloons, in which case the interface surface can deliver ultrasonic or other vibratory energy into a treatment region between these two balloons. The energy can thus act to mix or enhance penetration of a treatment confined between the two balloons in performing a vascular treatment procedure.—DRR

6,450,172

43.80.Vj BROADCAST AUDIBLE SOUND COMMUNICATION FROM AN IMPLANTABLE MEDICAL DEVICE

Jerome T. Hartlaub *et al.*, assignors to Medtronic, Incorporated
17 September 2002 (Class 128/899); filed 25 January 2000

This patent involves a method for broadcasting patient warnings or messages from an implantable medical device (IMD). The IMD transmits radio signals that can be received and reproduced in the form of voice statements or other audible sounds. The sounds are stored in an analog memory and triggered by a condition arising during programming, during interrogation, or by a warning trigger event. An external radio receiver receives the broadcast signal and the audible sounds are demodulated and reproduced as understandable messages comprising IMD information during programming and interrogation phases.—DRR

6,450,961

43.80.Vj ULTRASOUND IMAGING USING FLASH ECHO IMAGING TECHNIQUE

Eiichi Shiki and Yoshitaka Mine, assignors to Kabushiki Kaisha Toshiba
17 September 2002 (Class 600/458); filed in Japan 3 June 1999

Two images are produced by this system. The first image uses echoes from a transmission that can destroy an ultrasound contrast agent. The second image comes from a transmission that has less capability to destroy the contrast agent. This permits imaging in an interval of a flash echo phenomenon in an object that contains a contrast agent and in an interval when the phenomenon is absent.—RCW

6,450,962

43.80.Vj ULTRASONIC DIAGNOSTIC METHODS AND APPARATUS FOR GENERATING IMAGES FROM MULTIPLE 2D SLICES

Helmut Brandl *et al.*, assignors to Kretztechnik AG
17 September 2002 (Class 600/458); filed 18 September 2001

Ultrasonic echo information is received from a volume, scan converted, volume rendered, and projected onto a 2D cross-section.—RCW

6,454,715

43.80.Vj METHODS AND APPARATUS FOR BLOOD SPECKLE DETECTION IN AN INTRAVASCULAR ULTRASOUND IMAGING SYSTEM

Tat-Jin Teo, assignor to Scimed Life Systems, Incorporated
24 September 2002 (Class 600/443); filed 20 April 2001

The strong frequency dependence of scattering from blood is used to distinguish between scattering from blood and scattering from tissue that lacks a strong frequency dependence of scattering.—RCW

6,458,082

43.80.Vj SYSTEM AND METHOD FOR THE DISPLAY OF ULTRASOUND DATA

John I. Jackson *et al.*, assignors to Acuson Corporation
1 October 2002 (Class 600/441); filed 29 September 1999

A 2-D ultrasound echo data set is acquired, transformed into a 3-D data set based on a component of echoes in the set, and then transformed back into an altered 2-D data set to display image information in an easier-to-interpret format.—RCW

6,458,083

43.80.Vj ULTRASONIC HARMONIC IMAGING WITH ADAPTIVE IMAGE FORMATION

James R. Jago *et al.*, assignors to Koninklijke Philips Electronics N.V.
1 October 2002 (Class 600/443); filed 3 March 2000

Fundamental frequency, harmonic contrast, and tissue harmonic echo information are received. The fundamental and harmonic content of the echo information are analyzed and the relative content of an output signal is adjusted to take advantage of different characteristics in the echo information. Images formed using the signal are an adaptive blend of fundamental, tissue harmonic, and harmonic contrast echo information.—RCW

6,461,298

43.80.Vj THREE-DIMENSIONAL IMAGING SYSTEM

Aaron Fenster *et al.*, assignors to Life Imaging Systems
8 October 2002 (Class 600/437); filed 27 August 1999

An ultrasound array transducer is swept along a linear scanning path over a volume of interest and successive cross-sectional images of the volume are digitized. The digitized images are used to produce a three-dimensional image quickly. The three-dimensional image can be rotated about an arbitrary axis, a surface of the image can be translated to provide different cross-sectional views, and a selected surface can also be rotated about an arbitrary axis.—RCW

6,461,299

43.80.Vj MEDICAL DIAGNOSTIC ULTRASOUND SYSTEM AND METHOD FOR HARMONIC IMAGING WITH AN ELECTROSTATIC TRANSDUCER

John A. Hossack, assignor to Acuson Corporation
8 October 2002 (Class 600/437); filed 22 December 1999

Waveforms applied to an electrostatic transducer are predistorted to account for the nonlinear transfer characteristic of the transducer. The force on the transducer array is measured. The measured force is used to develop a bias voltage applied to the transducer.—RCW

6,461,303

43.80.Vj METHOD OF DETECTING ULTRASOUND CONTRAST AGENT IN SOFT TISSUE, AND QUANTITATING BLOOD PERFUSION THROUGH REGIONS OF TISSUE

Bjorn Angelsen, Trondheim, Norway
8 October 2002 (Class 600/458); filed 19 January 2001

Transmit pulse parameters are chosen to produce an incident pressure pulse for imaging a contrast agent at a particular depth with minimal varia-

tion over the range of the image. Nonlinearly distorted backscattered ultrasound signals are received from both the tissue and the contrast agent. The receive signals are passed through a beamformer that provides a steerable sensitivity.—RCW

6,464,638

43.80.Vj ULTRASOUND IMAGING SYSTEM AND METHOD FOR SPATIAL COMPOUNDING

Darwin P. Adams and Karl E. Thiele, assignors to Koninklijke Philips Electronics N.V.

15 October 2002 (Class 600/443); filed 5 October 2000

Ultrasound images are compounded in the elevation dimension. A controller in this system obtains image information from imaging planes prior to compounding and the information is used to compound the images. The compounded images consist of ultrasonic scanned beams that originate from spatially separate vantage points.—RCW

6,464,640

43.80.Vj METHODS AND APPARATUS FOR ULTRASOUND IMAGING WITH AUTOMATIC COLOR IMAGE POSITIONING

Ismayil M. Guracar and John I. Jackson, assignors to Acuson Corporation

15 October 2002 (Class 600/453); filed 6 January 2000

A Doppler signal at a spatial location of interest is summed to obtain a value along a surface. A corresponding signal is also found at subsequent times and summed with the initial signal to obtain a time-dependent surface integral value, which is displayed. Other quantities based on ultrasound

parameters such as Doppler energy and b-mode intensity can be calculated and displayed as quantities or waveforms as a function of time. Comparisons of data are provided.—RCW

6,464,641

43.80.Vj METHOD AND APPARATUS FOR AUTOMATIC VESSEL TRACKING IN ULTRASOUND IMAGING

Lihong Pan *et al.*, assignors to GE Medical Systems Global Technology Company LLC

15 October 2002 (Class 600/453); filed 25 February 2000

A range gate is maintained inside a moving vessel during the imaging process. The gate can be maintained at the center of the vessel, a distance from a vessel boundary, or at a ratio of distances between two boundaries. Successive image frames are processed and used to revise information about the vessel boundary positions and the vessel orientation in the vicinity of the range gate.—RCW

6,464,644

43.80.Vj METHOD OF ULTRASONIC IMAGING AND ULTRASONIC DIAGNOSTIC APPARATUS

Hiroshi Hashimoto, assignor to GE Medical Systems Global Technology Company, LLC

15 October 2002 (Class 600/458); filed in Japan 8 March 2000

Ultrasonic waves sufficiently strong to disperse a contrast agent are transmitted in a number of beams and received echoes are used to form an image. Ultrasonic waves of a reduced strength that will not disperse the contrast agent are then transmitted and received echoes are used to produce another image. These operations are repeated to display ultrasonic images from which time changes in blood flow are recognized at a glance.—RCW

Acoustic and mechanical response of reservoir rocks under variable saturation and effective pressure

C. L. Ravazzoli^{a)}

Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, Paseo del Bosque, S/N, (1900) La Plata, Argentina

J. E. Santos

Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, Paseo del Bosque, S/N, (1900) La Plata, Argentina, Department of Mathematics, Purdue University, 150 North University Street, West Lafayette, Indiana 47907-2068, and C.O.N.I.C.E.T., Argentina

J. M. Carcione

Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (OGS), Borgo Grotta Gigante 42c, 34010 Sgonico, Trieste, Italy

(Received 24 May 2001; revised 15 September 2002; accepted 17 December 2002)

We investigate the acoustic and mechanical properties of a reservoir sandstone saturated by two immiscible hydrocarbon fluids, under different saturations and pressure conditions. The modeling of static and dynamic deformation processes in porous rocks saturated by immiscible fluids depends on many parameters such as, for instance, porosity, permeability, pore fluid, fluid saturation, fluid pressures, capillary pressure, and effective stress. We use a formulation based on an extension of Biot's theory, which allows us to compute the coefficients of the stress-strain relations and the equations of motion in terms of the properties of the single phases at the *in situ* conditions. The dry-rock moduli are obtained from laboratory measurements for variable confining pressures. We obtain the bulk compressibilities, the effective pressure, and the ultrasonic phase velocities and quality factors for different saturations and pore-fluid pressures ranging from normal to abnormally high values. The objective is to relate the seismic and ultrasonic velocity and attenuation to the microstructural properties and pressure conditions of the reservoir. The problem has an application in the field of seismic exploration for predicting pore-fluid pressures and saturation regimes.

© 2003 Acoustical Society of America. [DOI: 10.1121/1.1554696]

PACS numbers: 43.20.Jr, 43.20.Hq [ANN]

I. INTRODUCTION

Rock-acoustic models relating the petrophysical characteristics of rocks—and their pore fluids at *in situ* conditions—with the acoustic properties, are essential tools in reservoir geophysics. Important applications are time-lapse seismics, geomechanics of petroleum-rich rocks, and reservoir characterization, evaluation, and monitoring.

The fundamental concepts about the stress-strain relations and the dynamics of porous deformable rocks fully saturated by single-phase fluids were established in the pioneer works of Biot.^{1–3} However, when the pore volume is occupied by more than one fluid phase, a different treatment is required, depending on the behavior of the fluids and their distribution within the pore space. Theoretical formulations for the study of the deformation and wave propagation in porous rocks with partial, miscible, or segregated fluid saturation have been presented by different authors. In this regard, Dutta and Odé⁴ analyzed the attenuation and dispersion of seismic waves in a brine-saturated rock containing spherical gas pockets using a model based on Biot's equations and White's model.⁵ Another application of Biot's formulation

was presented by Mochizuki⁶ to analyze experimental data in partially saturated rocks. The two fluids were modeled as a single phase one by means of a volume average for the density, an apparent viscosity, and an effective fluid compressibility. Some years later, using variational principles, Berryman *et al.*⁷ derived a theory for wave propagation in porous rocks saturated by segregated fluids (liquid and gas) and also for the case of mixing of liquid and gas. A different approach, based on a scattering theory, was employed by Toksöz *et al.*⁸ for pores of different shapes to analyze seismic velocities for variable pressure and saturation conditions. In this model two phase fluids are treated as inclusions (miscible or immiscible) and is also valid when each fluid occupies different parts of the pore space.

It must be remarked that none of these models incorporate the capillary forces existing when two immiscible fluids share the same pore volume. To take this effect into account, Santos *et al.*^{9,10} developed a theory based on a Lagrangian formulation and the principle of virtual complementary work. Consequently, unlike the previously mentioned models, this theory takes into account the residual saturations of both fluid phases and the fact that pressure variations induced by wave propagation are different in the two fluid phases. Some applications of this model were recently pub-

^{a)}Electronic mail: claudia@fcaglp.fcaglp.unlp.edu.ar; fax: 0054 221 4236591.

lished in Refs. 11, 12 for the estimation of gas hydrate concentration and free gas saturation from well logs and seismic data. Also, an interesting analysis and discussion about this subject was presented by Corapcioglu.¹³

In recent years, the oil industry has shown an increasing interest in the use of high-frequency seismic data for studying the changes in pore pressure and poroelastic moduli during production. Of utmost importance is the determination of fluid saturation and its distribution, and the pressure regime of the hydrocarbon fluids before drilling, since the prediction and detection of rocks with abnormally high fluid pressures (*overpressure*) is important to minimize risks and avoid dangerous blowouts during drilling operations. The occurrence of overpressure phenomena is observed in sedimentary basins worldwide and is associated with different mechanisms.¹⁴ The most important are gas to oil conversion (cracking)¹⁵ and disequilibrium compaction.¹⁶

In connection with these subjects, our main objective in this work is to present the application of a generalized version of the model described in Refs. 9, 10 to analyze the mechanical and acoustical properties of a consolidated reservoir sandstone saturated by immiscible fluids under variable confining and fluid pressures.

Ultrasonic laboratory experiments on dry samples¹⁷ for different confining pressures provide the dry-rock moduli versus effective pressure. These moduli and the effective stress coefficients are the most important parameters, since they characterize the acoustic properties of the rock. Our results indicate that the model is suitable for studying the transition zones within hydrocarbon reservoirs, providing a unified relationship between the properties of the rock frame and the pore fluids and their static and dynamic responses.

The paper is organized as follows. In Sec. II, we give the generalities of the model and a brief review of definitions and concepts regarding compressibilities and effective pressure. In the next section, we describe the equations of motion, the elastic coefficients of the model, and the propagation of elastic waves in a porous isotropic rock saturated by immiscible fluids, and give expressions for the phase velocities and quality factors. In Sec. IV, we apply the previous formulation to analyze the effects of variable saturation and fluid pressures on bulk compressibilities, effective pressure, phase velocities, and quality factors. We consider fluid pressures ranging from the normal hydrostatic state to abnormally high values, close to the fracturing limit. We use a realistic capillary pressure curve, and frequency-dependent dissipation, satisfying appropriate thermodynamic conditions.

II. BULK COMPRESSIBILITIES AND EFFECTIVE PRESSURE OF A THREE PHASE MEDIUM

Let us consider a volume V_b of a porous medium fully saturated by two immiscible fluids (such as oil, brine, or free gas) under a confining pressure P_c . In these situation, the distribution of fluids in the pore space depends on the wettability of the rock, i.e., its relative preference to be covered by a certain phase. The fluid in contact with the pore surface is the so-called wetting phase. In the following, the subscripts “w” and “n” refer to the wetting and nonwetting

fluids, respectively. Then, we denote S_w and S_n the averaged wetting and nonwetting fluid saturations, respectively, which are assumed to satisfy the condition of complete pore volume saturation: $S_n + S_w = 1$. The model assumes that both fluids can flow in the pore space, and, consequently, we only consider saturation states within the range^{18–21}

$$S_{rn} < S_n < 1 - S_{rw}, \quad S_{rw} < S_w < 1 - S_{rn}, \quad (2.1)$$

where S_{rn} and S_{rw} denote the residual saturations of both phases. The saturation S_{rn} is the amount of nonwetting fluid that remains in the pore space when the capillary pressure tends to zero, and S_{rw} —also called irreducible saturation—is the remaining wetting fluid when the capillary pressure reaches its maximum value.

Let ΔP_n , ΔP_w , and ΔS_n denote, respectively, infinitesimal changes in the pressures of the nonwetting and wetting fluids and the nonwetting fluid saturation, with respect to corresponding reference values \bar{P}_n , \bar{P}_w , and \bar{S}_n associated with the initial equilibrium state. Recall that the fluid pressures are related through the capillary relation^{18,20,21}

$$P_{ca}(S_n) = P_{ca}(\bar{S}_n + \Delta S_n) = \bar{P}_n - \bar{P}_w + \Delta P_n - \Delta P_w \\ = P_{ca}(\bar{S}_n) + \Delta P_{ca}. \quad (2.2)$$

Based on experimental data and ignoring hysteresis, we will assume that P_{ca} is a positive and strictly increasing function of the nonwetting fluid saturation.

Next, using some formulas given by Santos *et al.*,¹⁰ we will define and compute the bulk compressibilities for this type of media using a generalization of the arguments given by Zimmerman²² for single-phase fluids. First, from the strain–stress relations in that formulation we obtain an expression of the specific change in bulk volume $e_b = \Delta V_b / \bar{V}_b$ from a reference undeformed state \bar{V}_b due to changes in P_c , P_w , and P_n . Thus, from (39), (41a) and the expressions for $(3D^* + 1/2N)$, F_1 , F_2 in Ref. 10 and applying the capillary relation (2.2), it can be shown that²³

$$e_b = -C_m \Delta P_c - \delta(\bar{S}_n + \beta) \Delta P_n - \delta(\bar{S}_w - \beta) \Delta P_w, \quad (2.3)$$

where $\delta = C_s - C_m$, C_s and C_m being the compressibilities of the solid grains and dry matrix, $\beta = P_{ca}(\bar{S}_n) / P'_{ca}(\bar{S}_n)$ and $P'_{ca}(\bar{S}_n) = dP_{ca} / dS_n$.

Since the variations of the bulk volume V_b are due to the changes of three different pressures that may vary independently, we can define the following bulk compressibilities:

$$C_{bc} = -\frac{1}{\bar{V}_b} \left(\frac{\Delta V_b}{\Delta P_c} \right)_{P_w, P_{ca}} = \frac{1}{K_m} = C_m, \quad (2.4)$$

$$C_{bw} = \frac{1}{\bar{V}_b} \left(\frac{\Delta V_b}{\Delta P_w} \right)_{P_c, P_{ca}} = C_m - C_s, \quad (2.5)$$

$$C_{bca} = \frac{1}{\bar{V}_b} \left(\frac{\Delta V_b}{\Delta P_{ca}} \right)_{P_c, P_w} = (\bar{S}_n + \beta)(C_m - C_s), \quad (2.6)$$

where $K_m = C_m^{-1}$ is the bulk modulus of the dry matrix. The first two compressibilities are completely analogous to those

of the single-phase fluid case. In particular, (2.5) shows that P_w plays the same role as the “pore pressure” P_p used in the single-phase case.²²

Next, combining (2.4)–(2.6) and the capillary relation (2.2), Eq. (2.3) can be rewritten in differential form as follows:

$$e_b = -C_{bc}(P_c, P_w, P_{ca})dP_c + C_{bw}(P_c, P_w, P_{ca})dP_w + C_{bca}(P_c, P_w, P_{ca})dP_{ca}. \quad (2.7)$$

Assuming that e_b is an exact differential of the variables P_c , P_w , and P_{ca} and that C_s is constant, from (2.4), (2.5), and (2.7), we get

$$C_{bc} = C_{bc}(P_c - P_w, P_{ca}), \quad (2.8)$$

and from (2.6) and (2.7),

$$C_{bca} = C_{bca}(P_c - P_w, P_{ca}) = (\bar{S}_n + \beta)(C_{bc}(P_c - P_w, P_{ca}) - C_s), \quad (2.9)$$

Next, following Zimmerman,²² to obtain an effective pressure law for bulk volume deformation we compute the total strain E_b , integrating (2.7) in the (P_c, P_w, P_{ca}) state variables, along the following paths:

$$\text{Path 1: } (0, 0, P_{ca}(S_n^*)) \rightarrow (P_c, 0, P_{ca}(S_n^*)),$$

$$\text{Path 2: } (P_c, 0, P_{ca}(S_n^*)) \rightarrow (P_c, \bar{P}_w, P_{ca}(S_n^*)),$$

$$\text{Path 3: } (P_c, \bar{P}_w, P_{ca}(S_n^*)) \rightarrow (P_c, \bar{P}_w, P_{ca}(\bar{S}_n)),$$

obtaining

$$E_b = \int_0^{P_c} e_b dp_1 + \int_0^{\bar{P}_w} e_b dp_2 + \int_{P_{ca}(S_n^*)}^{P_{ca}(\bar{S}_n)} e_b dp_3 \approx -\bar{C}_{bc}(P_c - \bar{P}_w, P_{ca}(S_n^*))P_{ef}^b, \quad (2.10)$$

resulting in a *bulk effective pressure law* P_{ef}^b given by

$$P_{ef}^b = P_c - \bar{n}_{b1}\bar{P}_w - \bar{n}_{b2}(P_{ca}(\bar{S}_n) - P_{ca}(S_n^*)), \quad (2.11)$$

where

$$\bar{n}_{b1} = 1 - C_s/\bar{C}_{bc}(P_c - \bar{P}_w, P_{ca}(S_n^*)), \quad (2.12)$$

$$\bar{n}_{b2} = \bar{C}_{bca}(P_c - \bar{P}_w, P_{ca}(S_n^*), P_{ca}(\bar{S}_n))/\bar{C}_{bc}(P_c - \bar{P}_w, P_{ca}(S_n^*)). \quad (2.13)$$

It must be observed that the integration in the capillary pressure variable (Path 3) does not begin at $P_{ca}=0$ since such a value corresponds to the irreducible saturation S_{rn} . To avoid loss of generality we start the process at an arbitrary saturation S_n^* within the range (2.1).

The coefficients \bar{C}_{bca} and \bar{C}_{bc} , are the mean (or “secant”) values of the compressibilities C_{bca} and C_{bc} in the intervals $[0, P_c]$ and $[P_{ca}(S_n^*), P_{ca}(\bar{S}_n)]$, respectively,

$$\bar{C}_{bc}(P_c, P_{ca}(\bar{S}_n)) = \frac{1}{P_c} \int_0^{P_c} C_{bc}(p, P_{ca}(\bar{S}_n)) dp, \quad (2.14)$$

$$\begin{aligned} \bar{C}_{bca}(P_c - \bar{P}_w, P_{ca}(S_n^*), P_{ca}(\bar{S}_n)) &= \frac{1}{P_{ca}(\bar{S}_n) - P_{ca}(S_n^*)} \int_{P_{ca}(S_n^*)}^{P_{ca}(\bar{S}_n)} C_{bca}(P_c - \bar{P}_w, p_3) dp_3 \\ &= \frac{1}{P_{ca}(\bar{S}_n) - P_{ca}(S_n^*)} \int_{S_n^*}^{\bar{S}_n} [s + \beta(s)] \\ &\quad \times [C_{bc}(P_c - \bar{P}_w, P_{ca}(s)) - C_s] P'_{ca}(s) ds, \end{aligned} \quad (2.15)$$

where we have used (2.6) and the fact that $p_3 = p_3(s)$ is the capillary pressure function and the variable s takes values in the saturation range $[S_n^*, \bar{S}_n]$.

It must be remarked that the effective pressure in this case not only depends on the confining and wetting fluid pressures but also on the changes in capillary pressure.

III. EQUATIONS OF MOTION AND ELASTIC COEFFICIENTS

In this section we present the equations of motion for this kind of media, which will be used in Sec. IV to compute seismic wave velocities and quality factors. Our objective is to analyze the combined effects of saturation and confining and pore pressures on the properties of the different wave modes.

Let \mathbf{u}^s , $\tilde{\mathbf{u}}^n$, and $\tilde{\mathbf{u}}^w$ denote the time Fourier transforms of the averaged absolute displacements for the solid, nonwetting, and wetting fluid phases, respectively. If ϕ denotes the effective porosity, then we introduce the relative fluid displacements \mathbf{u}^l and \mathbf{u}^w , given by $\mathbf{u}^l = \phi(\tilde{\mathbf{u}}^l - \mathbf{u}^s)$, $l = n, w$.

For a spatially homogeneous medium, the frequency-domain equations of motion take the following form:

$$\begin{aligned} (i) \quad & -\omega^2(\rho\mathbf{u}^s + \rho_n\bar{S}_n\mathbf{u}^n + \rho_w\bar{S}_w\mathbf{u}^w) \\ & = (K_c + \frac{4}{3}N)\nabla(\nabla\cdot\mathbf{u}^s) - N\nabla\times\nabla\times\mathbf{u}^s + B_1(\nabla\cdot\mathbf{u}^n) \\ & \quad + B_2\nabla(\nabla\cdot\mathbf{u}^w); \\ (ii) \quad & -\omega^2(\rho_n\bar{S}_n\mathbf{u}^s + g_n\mathbf{u}^n + g_{nw}\mathbf{u}^w) + i\omega d_n\mathbf{u}^n - i\omega d_{nw}\mathbf{u}^w \\ & = B_1\nabla(\nabla\cdot\mathbf{u}^s) + M_1\nabla(\nabla\cdot\mathbf{u}^n) + M_3\nabla(\nabla\cdot\mathbf{u}^w); \\ (iii) \quad & -\omega^2(\rho_w\bar{S}_w\mathbf{u}^s + g_{nw}\mathbf{u}^n + g_w\mathbf{u}^w) + i\omega d_w\mathbf{u}^w - i\omega d_{nw}\mathbf{u}^n \\ & = B_2\nabla(\nabla\cdot\mathbf{u}^s) + M_3\nabla(\nabla\cdot\mathbf{u}^n) + M_2\nabla(\nabla\cdot\mathbf{u}^w), \end{aligned} \quad (3.1)$$

where $\omega = 2\pi f$ denotes the angular frequency and f is the frequency. The coefficients ρ_n and ρ_w are the mass densities of the fluids and ρ is the mass density of the bulk material, which is related to the grain density ρ_s and the porosity as

$$\rho = (1 - \phi)\rho_s + \phi(\bar{S}_n\rho_n + \bar{S}_w\rho_w). \quad (3.2)$$

The quantities g_n , g_w , g_{nw} , are mass coupling coefficients; as in the classical Biot's theory for single-phase fluid saturation, they represent the inertial effects associated with dynamic interactions between the three different phases. The coefficients d_n , d_w , and d_{nw} represent viscous drag effects that can be expressed in terms of the viscosities and relative permeabilities of each fluid phase.

For frequencies higher than characteristic frequencies to be defined later, these coefficients need to be modified. For simple pore geometries like plane slits and cylindrical ducts Auriault *et al.*²⁴ gave explicit expressions for frequency-dependent relative permeability tensors using homogenization techniques. In Sec. IV we will define the mass and viscous coupling coefficients using an extension of the ideas described in Refs. 2 and 25.

The elastic moduli K_c , B_1 , B_2 , M_1 , M_2 , M_3 are given by the following formulas that can be obtained with the arguments given in Ref. 10 and taking into account the reference pressures \bar{P}_w , \bar{P}_n of both fluid phases:

$$K_c = \frac{K_s(K_m + \Xi)}{(K_s + \Xi)}, \quad \Xi = \frac{K_f(K_m - K_s)}{\phi(K_f - K_s)},$$

$$K_f = \alpha \left(\frac{\gamma \bar{S}_n}{K_n} + \frac{\bar{S}_w}{K_w} \right)^{-1}, \quad \alpha = 1 + (\bar{S}_n + \beta)(\gamma - 1), \quad (3.3)$$

$$\gamma = \left(1 + \frac{P'_{ca}(\bar{S}_n)\bar{S}_n\bar{S}_w}{K_w} \right) \left(1 + \frac{P'_{ca}(\bar{S}_n)\bar{S}_n\bar{S}_w}{K_n} \right)^{-1}.$$

The remaining coefficients can be obtained by using the following relations:

$$B_1 = \chi K_c [(\bar{S}_n + \beta)\gamma - \beta + (\gamma - 1)\zeta],$$

$$B_2 = \chi K_c [(\bar{S}_w + (1 - \gamma)\zeta)], \quad (3.4)$$

$$M_1 = -M_3 - \frac{B_1}{\delta K_m}, \quad M_2 = \frac{r}{q} B_2 + \frac{\zeta}{q},$$

$$M_3 = -B_2 \left[\frac{r}{q} + \frac{1}{K_m \delta} \right] - \frac{\zeta}{q}, \quad (3.5)$$

with

$$\chi = \left[\delta + \phi \left(\frac{1}{K_m} - \frac{1}{K_c} \right) \right] \left\{ \alpha \left[\delta + \phi \left(\frac{1}{K_m} - \frac{1}{K_f} \right) \right] \right\}^{-1},$$

$$r = \frac{\bar{S}_n + \beta}{K_s} + \frac{1}{K_c - K_m} \left[q B_2 + (\bar{S}_n + \beta) \left(1 - \frac{K_c}{K_s} \right) \right],$$

$$q = \phi \left(\frac{1}{K_n} + \frac{1}{P'_{ca} \bar{S}_n \bar{S}_w} \right),$$

where $\zeta = \bar{P}_w / P'_{ca}(\bar{S}_n)$. According to Gassman's theory,²⁶ it will be assumed that the coefficient N is identical to the shear modulus of the dry rock N_m .

It is known that wave dispersion and attenuation in real saturated porous media is higher than that predicted by Biot-type theories, in which dissipation is mainly due to viscodynamic effects. The main discrepancies are related to the complexity of the pore shapes, heterogeneities in the spatial distribution of the fluids, and other physical properties, the presence of clays and the anelasticity of the rock matrix. These factors give rise to a variety of attenuation mechanisms that can be included in the formulation, introducing viscoelasticity. The resultant poroviscoelastic moduli are complex and frequency-dependent functions and their computation will be explained in detail in Sec. IV C.

TABLE I. Properties of Bandera sandstone.

Porosity, ϕ	20%
Absolute permeability, K	$3.5 \times 10^{-15} \text{ m}^2$
Grain density, ρ_s	2700 kg/m^3
Grain bulk modulus, K_s	44 GPa

The plane wave analysis performed in Ref. 9 shows that in these types of media, three different compressional waves (Type I, Type II, and Type III) and one shear wave (or S wave) propagate.

The Type-I wave is the analog of the classical fast P wave propagating in elastic or viscoelastic isotropic solids and is associated with the motion of both solid and fluids moving nearly in phase. Type II and Type III are slow waves strongly attenuated in the low-frequency range. The first one is analogous to the classical Biot's slow wave, with the two fluids moving nearly in phase and in opposite phase with the solid. For the Type-III waves, both fluids move in nearly opposite phase with each other.

Denoting by k_j , $j = \text{I, II, III, S}$, the complex wave numbers, the phase velocities C_j and quality factors Q_j are obtained by using the formulas

$$C_j = \frac{\omega}{\text{Re}(k_j)}, \quad Q_j = -\frac{1}{2} \frac{\text{Re}(k_j)}{\text{Im}(k_j)}, \quad j = \text{I, II, III, S}. \quad (3.6)$$

IV. APPLICATION TO THE STUDY OF REAL SANDSTONES

In this section, we apply the theoretical formulation described in previous sections to predict and investigate the behavior of the compressibilities, wave velocities, and quality factors of hydrocarbon reservoir rocks at *in situ* conditions. We use the ultrasonic laboratory data published by King,¹⁷ who measured compressional and shear wave velocities of a set of sandstones (dry and fully saturated) under variable confining pressure. Among the available samples, we chose *Bandera* sandstone, a well-cemented and consolidated feldspathic graywacke.²⁷ The petrophysical and geotechnical characteristics of the sample (porosity, permeability, grain density, and bulk modulus) are taken from King¹⁷ and Mann and Fatt²⁷ (see Table I). The mineral bulk modulus K_s is taken from Zimmerman.²² The sandstone is assumed to be isotropic.

To introduce the variation of the matrix properties with effective pressure, we obtained the dependence of the (elastic) shear and bulk dry-rock moduli $N_m(P_c)$ and $K_m(P_c)$ versus confining pressure. For that purpose, we assume that for each P_c , the measured velocities can be described in terms of the simplified elastic isotropic approximation. Then, we performed exponential regressions of the form

$$f^{-1}(P_c) = y_0 + A_1 \exp(-(P_c - P_0)/t_1) + A_2 \exp(-(P_c - P_0)/t_2), \quad (4.1)$$

where f stands for N_m or K_m . The coefficients are given in Table II.

Since fluid pressure for a dry sample can be assumed to be zero, the effective pressure equals the confining pressure.

TABLE II. Coefficients of Eq. (4.1) for Bandera sandstone. For pressures given in MPa the moduli result in GPa.

	P_0	γ_0	A_1	t_1	A_2	t_2
$(K_m)^{-1}$	2.318 591 921 065	7.9435×10^{-2}	8.687×10^{-2}	3.509	8.553×10^{-2}	19.32
$(N_m)^{-1}$	1.524 982 994 273	7.9698×10^{-2}	3.342×10^{-2}	8.22	1.956×10^{-2}	41.26

Then, to obtain the elastic properties of the skeleton for a given combination of confining and fluid pressures, we simply replace P_c by the effective pressure P_{ef}^b [given by (2.11)] in Eq. (4.1). The whole set of poroelastic moduli are then computed using (3.3)–(3.5).

The mass and viscous coupling coefficients in the equations of motion (3.1) are computed by using the expressions given in Ref. 9 modified by frequency correction factors:

$$g_l(\omega) = \frac{F_s \rho_l \bar{S}_l}{\phi} + \mu_l(\bar{S}_l)^2 D_l \frac{F_l(\theta_l)}{\omega}, \quad l = n, w; \quad (4.2)$$

$$g_{nw}(\omega) = \epsilon \frac{F_s}{\phi} (\rho_n \rho_w, \bar{S}_n, \bar{S}_w)^{1/2} + \frac{(\mu_n \mu_w)^{1/2} \bar{S}_n \bar{S}_w k_{rnw}}{D} \frac{F_l(\theta_{nw})}{\omega}; \quad (4.3)$$

$$d_l(\omega) = \mu_l(\bar{S}_l)^2 D_l F_R(\theta_l), \quad l = n, w; \quad (4.4)$$

$$d_{nw}(\omega) = \frac{(\mu_n \mu_w)^{1/2} \bar{S}_n \bar{S}_w k_{rnw}}{D} F_R(\theta_{nw}); \quad (4.5)$$

where $D = K(k_{rn}k_{rw} - k_{rnw}^2)$, $D_n = k_{rn}/D$, $D_w = k_{rn}/D$. The constants μ_n , μ_w are the fluid viscosities and K , $k_{rn}(S_n)$, $k_{rw}(S_n)$ and $k_{rnw}(S_n)$ are the absolute and coupling permeability functions, respectively. The relative permeability functions and the capillary pressure $P_{ca}(S_n)$ used here are computed using the relations²⁸

$$k_{rn}(S_n) = \left(1 - \frac{1 - S_n}{1 - S_{rn}}\right)^2, \quad k_{rw}(S_n) = \left(\frac{1 - S_n - S_{rw}}{1 - S_{rw}}\right)^2, \quad (4.6)$$

$$P_{ca}(S_n) = A \left(\frac{1}{(S_n + S_{rw} - 1)^2} - \frac{S_{rn}^2}{S_n^2(1 - S_{rn} - S_{rw})^2} \right). \quad (4.7)$$

Also, we take $k_{rnw}(S_n) = \sqrt{\epsilon k_{rn}(S_n) k_{rw}(S_n)}$. These relations are based on laboratory experiments performed on different porous rocks during imbibition and drainage processes (neglecting hysteresis effects). However, it must be pointed out that those functions depend on the nature of the porous medium and the wettability of the pore fluids involved. In what follows the capillary pressure amplitude coefficient A equals 3 KPa. The residual saturation values are $S_{rn} = S_{rw} = 5\%$ and, consequently, the saturation S_n in all the calculations varies in the range

$$0.06 \leq \bar{S}_n \leq 0.94. \quad (4.8)$$

The factor ϵ in the definition of k_{rnw} is taken to be 0.01. The parameter F_s is the formation structure factor and is related to the tortuosity of the pores. It can be estimated using the following relation:²⁹ $F_s = 1/2(1 + 1/\phi)$.

The complex-valued frequency-dependent function $F(\theta_j) = F_R(\theta_j) + iF_I(\theta_j)$, $j = n, w, nw$ in (4.2)–(4.5) is the “universal” frequency correction function defined by Biot² for single-phase fluids:

$$F(\theta) = \frac{1}{4} \frac{\theta T(\theta)}{1 - \frac{2}{i\theta} T(\theta)}, \quad T(\theta) = \frac{\text{ber}'(\theta) + i\text{bei}'(\theta)}{\text{ber}(\theta) + i\text{bei}(\theta)},$$

with $\text{ber}(\theta)$ and $\text{bei}(\theta)$ being the Kelvin functions of the first kind and zeroth order. The arguments θ_j for $F(\theta_j)$, $j = n, w, nw$ in (4.2)–(4.5) can be estimated as in Refs. 2 and 25,

$$\theta_j = a_p^j \sqrt{\frac{\omega \rho_j}{\mu_j}}, \quad a_p^j = 2 \sqrt{\frac{K k_{rj} A_0}{\phi}}, \quad j = n, w, nw, \quad (4.9)$$

where A_0 denotes the Kozeny–Carman constant.^{18,30} This correction function $F(\theta)$ is needed for frequencies higher than the minimum of the characteristic frequencies, defined as

$$\omega_c^l(\bar{S}_n) = \frac{\bar{S}_l \mu_l \phi D_l}{F_s \rho_l}, \quad l = n, w, \quad (4.10)$$

$$\omega_c^{nw}(\bar{S}_n) = \frac{\phi(\mu_n \mu_w)^{1/2} (\bar{S}_n \bar{S}_w)^{1/2} k_{rnw}}{\epsilon D F_s (\rho_n \rho_w)^{1/2}}.$$

These expressions have been obtained, taking into account that we have two immiscible fluids and three relative permeability functions, following the ideas given in Ref. 7 for the case of segregated fluids.

In the following numerical tests, we study the mechanical response of the sandstone saturated by brine (the wetting phase) and oil (nonwetting phase) at a fixed confining pressure (i.e., at a fixed depth), for variable fluid pressures and saturation states. The model represents a transition zone within a hydrocarbon reservoir. We assume that the rock is located at a depth of about 3000 m at a confining pressure of about 70 MPa. The physical properties of brine and oil—assumed to be independent of pressure—are presented in Table III.

A. Effective pressure law

It is a well-established fact that knowledge of the effective pressure law of a porous saturated rock is essential for the study of its mechanical behavior. In many cases, it has

TABLE III. Physical properties of brine and oil.

Fluid	Density	Viscosity	Bulk modulus
Brine	1030 kg/m ³	0.001 N s/m ²	2.25 GPa
Oil	700 kg/m ³	0.01 N s/m ²	0.57 GPa

been observed that when both confining and formation fluid pressures change, only the difference between the two, i.e., the *differential pressure*, has a significant effect on elastic properties and velocities. In particular, when this difference is small, the differential pressure gives a reasonable estimate of the effective pressure.³¹ Since in our model we have two different fluid pressures that can vary independently, such a comparison cannot be directly done. Instead, we introduce two differential pressures, associated with each fluid phase, namely

$$P_{wd} = P_c - \bar{P}_w, \quad P_{nd} = P_c - \bar{P}_n. \quad (4.11)$$

As stated in the Introduction, we are interested in the study of acoustic wave propagation in overpressured formations. Such a physical condition of the rock is commonly quantified in terms of an “excess pore pressure,” defined as the difference between the fluid pressure and the hydrostatic pressure P_h exerted by a column of water extending from the surface up to the depth under consideration. Thus, we define the excess pressure of each fluid phase as

$$\Delta P_{wh} = \bar{P}_w - P_h, \quad \Delta P_{nh} = \bar{P}_n - P_h. \quad (4.12)$$

In the following numerical tests, the wetting fluid pressure \bar{P}_w (analogous to “pore pressure”) is varied from the water hydrostatic pressure P_h for the considered depth (about 29 MPa) to the confining pressure value, assuming that the wetting fluid pressure cannot exceed the confining pressure.

To obtain the coefficients \bar{n}_{b1} and \bar{n}_{b2} given by Eqs. (2.12) and (2.13), we first estimate the secant compressibilities \bar{C}_{bc} and \bar{C}_{bca} , disregarding the possible dependence of the C_{bc} coefficient on P_{ca} . The initial saturation state in these computations is $S_n^* = 0.053$. Next, using Eq. (2.11), we obtain the effective pressure law for the deformation of the bulk volume $P_{ef}^b(P_c, \bar{P}_w, P_{ca}(\bar{S}_n))$ for each fluid pressure and saturation state. The effective pressure law is then analyzed as a function of the wetting fluid pressure and saturation.

Figure 1(a) shows the resulting P_{ef}^b versus wetting fluid pressure for $\bar{S}_n = 0.4$ (at constant $P_{ca} \approx 10$ KPa). We compare the results of our model with other possible estimates, such as the wetting-differential pressure P_{wd} and a general law proposed by Gangi and Carlson³¹ of the form

$$P_{ef}^b = P_c - n_G \bar{P}_w, \quad \text{with } n_G = n_0 - n_1 P_{wd}. \quad (4.13)$$

In these equations, the “pore pressure” is replaced by its equivalent \bar{P}_w , and we have considered $n_0 = 1$ and $n_1 = 0.014 \text{ MPa}^{-1}$, with pressures given in MPa.

For a fixed saturation, we observe significant differences, particularly in the low pore pressure range. As expected, the increase in wetting fluid pressure causes a reduction in the effective pressure acting on the solid matrix, which produces a “softening” effect in the formation. Thus, the behavior of the effective pressure law is consistent with the previously defined laws and explains the well-known effect of the decreasing of seismic velocities observed in overpressured formations.

Figure 1(b) shows the effect of changes in saturation—and therefore capillary pressure—on the effective pressure

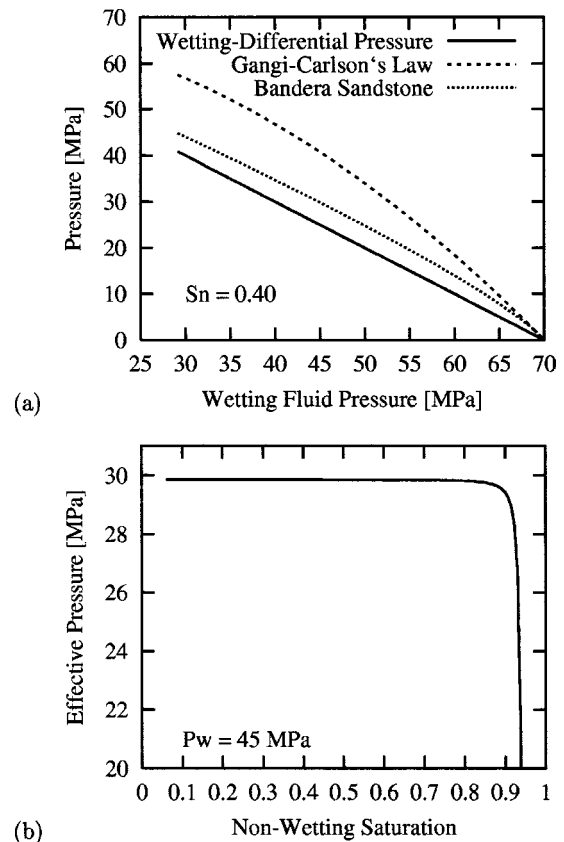


FIG. 1. (a) A comparison of effective pressure, wetting-differential pressure (P_{wd}), and Gangi–Carlson’s law for $\bar{S}_n = 0.40$. (b) Effective pressure versus saturation \bar{S}_n (increasing P_{ca}) for $\bar{P}_w = 45$ MPa and $P_c = 70$ MPa.

law, keeping the wetting fluid pressure \bar{P}_w constant at 45 MPa. Although the effective pressure is almost independent of \bar{S}_n , a marked decrease can be appreciated (of about 33%) when \bar{S}_n approaches its upper limit, i.e., near the residual brine saturation. This reduction also depends on confining and wetting fluid pressures and is associated with the rapid growth of the capillary pressure and the \bar{n}_{b2} coefficient. The behavior of the effective pressure coefficients versus pressure and saturation is shown in Figs. 2(a) and (b), respectively. In the first plot, we observe that when \bar{P}_w approaches P_c (i.e., near the fracture limit) $n_{b1} \rightarrow 1$ and $n_{b2} \rightarrow 0$, and $P_{ef}^b \rightarrow P_{wd}$. This is due to the asymptotic limit of the compressibility \bar{C}_{bc} that tends to infinity for $\bar{P}_w \rightarrow P_c$.

It is worthwhile to remark that, when considering elastic wave propagation in this kind of media, the different wave modes induce infinitesimal oscillatory changes in the stresses and pressures of the solid and fluid phases with respect to their reference equilibrium values. Under this assumption the stress–strain process takes place within the linear range. However, those increments are not taken into account in the evaluation of the effective pressure law, which is computed in terms of P_c , \bar{P}_w , and $P_{ca}(\bar{S}_n)$.

B. Bulk compressibilities versus saturation and pressure

In this section, we apply the equations given in Sec. II to compute the bulk compressibilities of the sandstone. A simi-

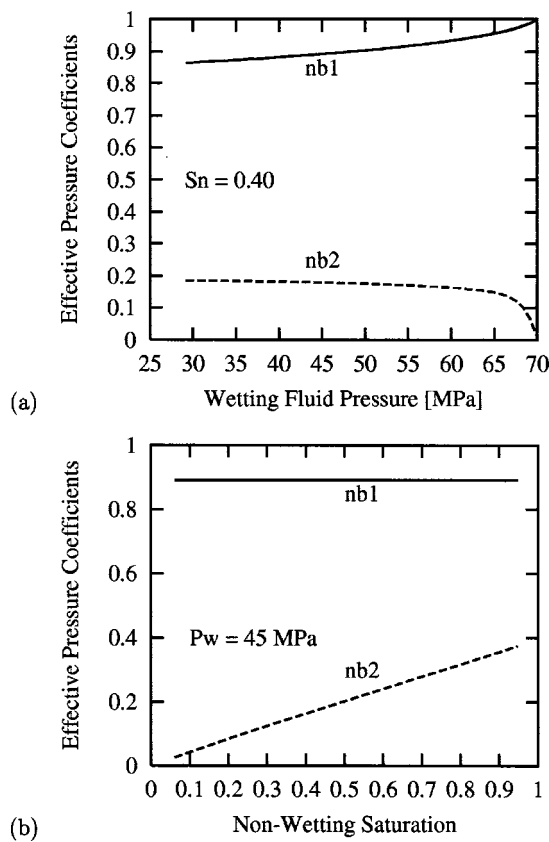


FIG. 2. Effective pressure coefficients n_{b1} and n_{b2} (a) versus pore pressure \bar{P}_w (for $\bar{S}_n = 0.40$), (b) versus saturation \bar{S}_n (increasing P_{ca}) for $\bar{P}_w = 45$ MPa.

lar analysis was presented in Ref. 32. To estimate the elastic response for different combinations of the state variables, using the law (4.1) previously obtained for $K_m(P_c)$, we evaluate the coefficient $C_m = K_m^{-1}$ at each effective pressure. In this way, we obtain the variation of the compressibilities (2.4)–(2.6) with effective pressure, assuming constant grain compressibility C_s and an elastic reversible behavior. As pointed out in Ref. 33, this is valid for well-cemented and consolidated rocks, but in loose unconsolidated sands the compaction is both anelastic and irreversible.

Figure 3(a) shows the compressibilities C_{bc} , C_{bw} , and C_{bca} versus effective pressure, normalized by C_s . Note in (2.4)–(2.5) that, unlike the C_{bca} coefficient, the other two compressibilities do not depend explicitly on P_{ca} . The nonlinear decreasing behavior observed in all the cases can be attributed to the closure of microcracks, low aspect ratio pores, and loose grain contacts, which increase the stiffness of the rock. In Fig. 3(b), we observe the variation of C_{bca} versus nonwetting saturation for a given wetting fluid pressure. The magnitude of this coefficient suggests that capillary pressure changes may have a non-negligible influence on bulk volume changes, particularly in the range of intermediate to high nonwetting saturations and for abnormally high pore pressures (low effective pressures).

C. Poroviscoelastic moduli

In the previous sections we considered bulk effective pressures and compressibilities, using purely elastic approxi-

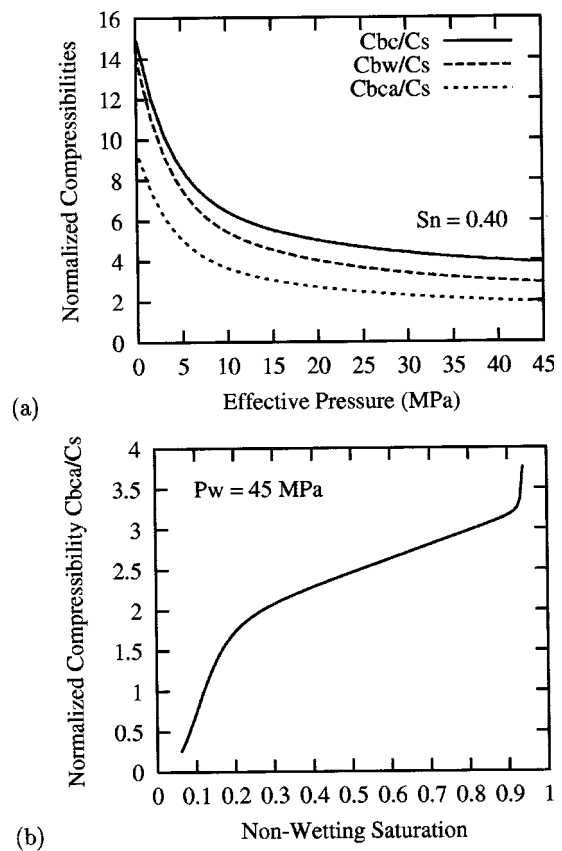


FIG. 3. (a) Normalized compressibilities C_{bc} ($=C_m$), C_{bw} and C_{bca} of Bandera sandstone versus bulk effective pressure; (b) C_{bca} versus nonwetting fluid saturation.

mations of the elastic moduli. This approach is frequently used in many static or quasistatic problems in rock mechanics, which can be regarded as reversible processes. However, when considering dynamic problems such as wave propagation, the irreversible character of the deformation process plays a very important role.

The dissipation due to a viscous solid–fluid interaction was already introduced in the equations of motion with the terms d_n , d_w , and d_{nw} . In addition, as stated in Sec. III, viscoelastic dissipation must be incorporated in the formulation. The theoretical basis for the extension of the poroelastic formulation to include viscoelasticity were given by Biot^{34,3} and later applied by different authors.^{35–38} Using principles of irreversible thermodynamics, Biot established a general correspondence rule that allows us to extend our formulation to the poroviscoelastic case, replacing the real elastic coefficients by complex frequency-dependent moduli in the frequency domain.

Many mechanisms and models are proposed in the literature to explain the irreversible behavior observed in real rocks.³⁹ However, for practical purposes, we use the phenomenological model defined by Liu *et al.*⁴⁰ to describe attenuation by making the shear and undrained modulus complex and frequency dependent, while all the other coefficients remain real. This is a linear and causal model, whose associated complex modulus behaves properly in all the frequency range, and satisfies the restrictions imposed by the laws of thermodynamics.⁴¹

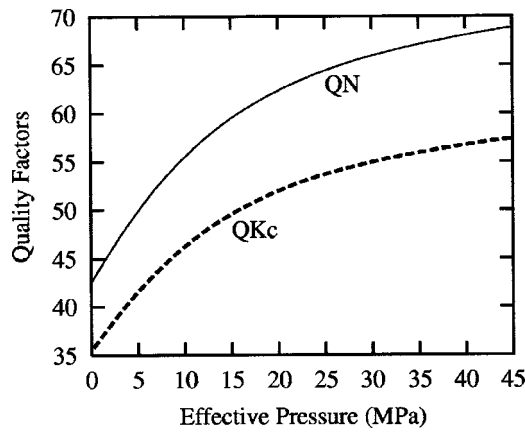


FIG. 4. Assumed values for \hat{Q}_{Kc} and \hat{Q}_N coefficients versus effective pressure.

Using this model, the complex moduli K_c , N are computed as

$$K_c(\omega, P_{ef}^b, \bar{S}_n) = \frac{K_c^r(P_{ef}^b, \bar{S}_n)}{R_{K_c}(\omega, P_{ef}^b) - iT_{K_c}(\omega, P_{ef}^b)},$$

$$N(\omega, P_{ef}^b) = \frac{N^r(P_{ef}^b)}{R_N(\omega, P_{ef}^b) - iT_N(\omega, P_{ef}^b)}. \quad (4.14)$$

The real coefficients $K_c^r(P_{ef}^b)$ and $N^r(P_{ef}^b)$ denote the relaxed closed bulk and shear moduli, respectively. They are chosen so that the high-frequency limits of (4.14) match the values of $N(P_{ef}^b)$ and $K_c(P_{ef}^b, \bar{S}_n)$ obtained from the experimental law (4.1) and (3.3), respectively. The frequency-dependent functions R_l and T_l , $l = K_c, N$, associated with a continuous spectrum of relaxation times, characterize the viscoelastic behavior and are given by^{40,42}

$$R_l(\omega, P_{ef}^b) = 1 - \frac{1}{\pi \hat{Q}_l(P_{ef}^b)} \ln \frac{1 + \omega^2 T_{1,l}^2}{1 + \omega^2 T_{2,l}^2}, \quad (4.15)$$

$$T_l(\omega, P_{ef}^b) = \frac{2}{\pi \hat{Q}_l(P_{ef}^b)} \tan^{-1} \frac{\omega(T_{1,l} - T_{2,l})}{1 + \omega^2 T_{1,l} T_{2,l}}, \quad l = K_c, N. \quad (4.16)$$

The parameters \hat{Q}_l , $T_{1,l}$, and $T_{2,l}$ are taken such that the resulting quality factor $Q_l = T_l/R_l$ is approximately equal to \hat{Q}_l in the range of frequencies where the equations are solved. For this computation we also introduced the dependence of the quality factors associated to K_c and N (i.e., \hat{Q}_{Kc} and \hat{Q}_N), on effective pressure,⁴³ disregarding any possible dependence of such coefficients on saturation. The corresponding curves are shown in Fig. 4.

D. Wave velocities

We apply the formulation described in Sec. III to analyze the phase velocities [given by (3.6)] versus effective pressure and saturation, in the ultrasonic frequency range ($f = 1$ MHz) commonly used in laboratory experiments.

Figure 5(a) shows the well-known behavior of the type-I P- and S-wave velocities, which have a marked increase with

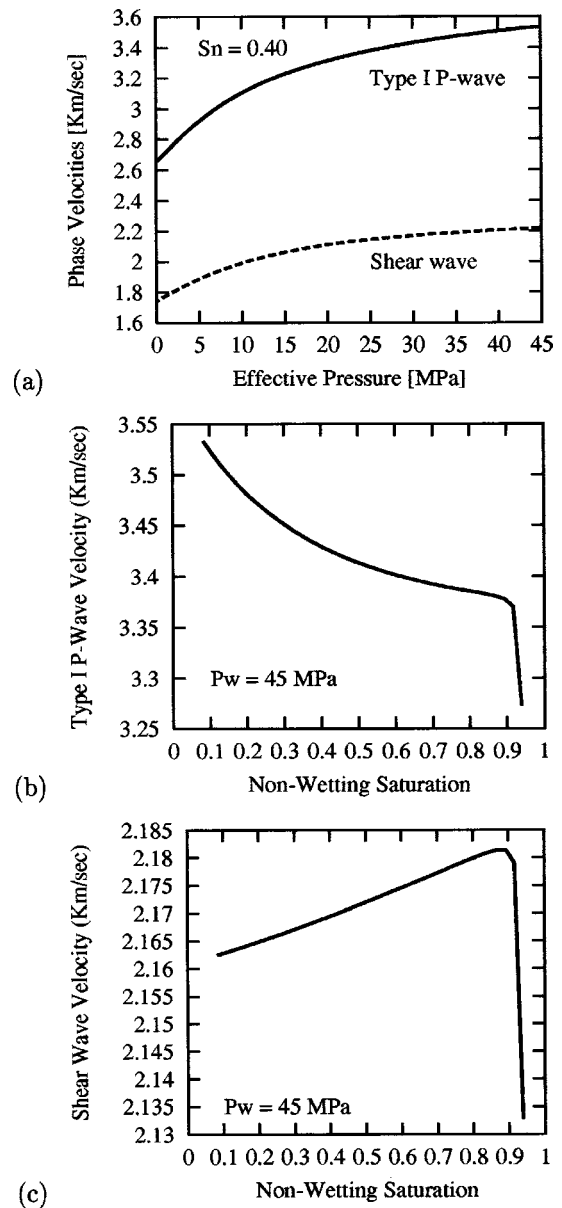


FIG. 5. Type I P- and S-wave phase velocities versus effective pressure (a), and versus nonwetting fluid saturation [(b) and (c)].

increasing effective pressure.^{8,44,47} Figure 5(b) shows the type-I P-wave velocity versus saturation for $\bar{P}_w = 45$ MPa (that is, for an excess wetting fluid pressure $\Delta P_{wh} \approx 16$ MPa). As expected, there is a maximum at minimum oil saturation, and a decreasing trend at intermediate states. The absolute minimum corresponds to the effect observed in the effective pressure law near the residual brine saturation [see Fig. 1(b)].

The shear wave velocity, shown in Fig. 5(c) increases almost linearly with increasing oil saturation, showing the same abrupt decrease associated with the anomaly of the effective pressure near $\bar{S}_n = 1 - S_{rw}$. Except at this point, the changes are only associated with density effects, since the bulk density [given by (3.2)] decreases with increasing \bar{S}_n and by hypothesis the matrix shear modulus does not depend on saturation.

In Fig. 6(a), we plot the type-II and type-III P-wave

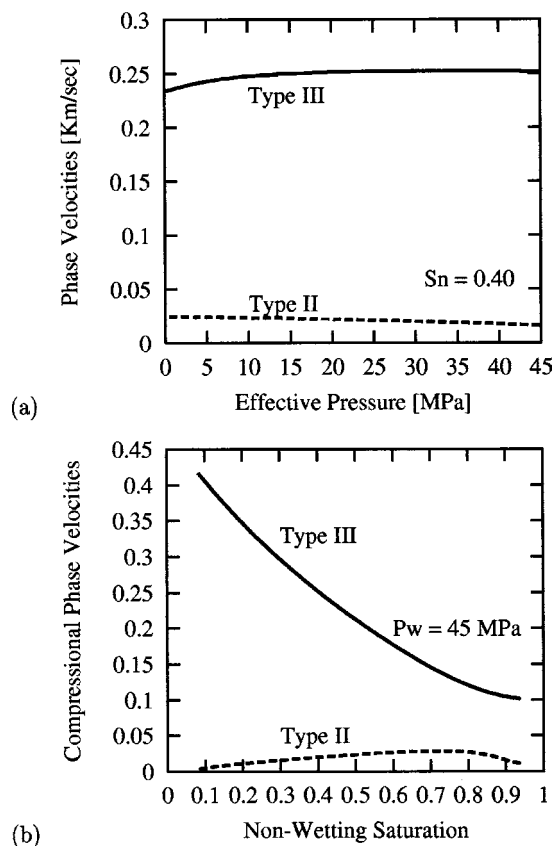


FIG. 6. Type-II and type-III P-wave phase velocities versus effective pressure (a), and versus nonwetting fluid saturation (b).

velocities versus effective pressure, which are almost constant in the whole pressure range. Figure 6(b) shows the phase velocities versus saturation. We observe for this sandstone that the type-III P wave is very sensitive to oil saturation. The velocity of the type-II P wave has a maximum at intermediate saturation values, approaching zero near the residual saturations. This suggests that the mobility of the fluid phases plays an important role. The behavior of these velocities versus frequency was analyzed by Santos *et al.*¹⁰ using a poroelastic formulation.

E. Quality factors

We use Eq. (3.6) to compute the quality factors Q_I , Q_{II} , Q_{III} , and Q_S associated with the different wave modes. The quality factors of the type-I P and S waves are shown in Figs. 7(a), (b), and (c). They show a pronounced monotonic increase with increasing effective pressure, in agreement with the observations made by different authors in other sandstones.^{15,16,44,47}

The behavior of Q_I versus saturation is more complex, showing a relative minimum in the range from low to intermediate nonwetting (oil) saturation. The shear-wave quality factor Q_S tends to increase with oil saturation. Since an increase in oil saturation increases the viscosity of the fluid mixture—and hence the viscous dissipation—these results suggest a nonlinear relationship between quality factors and viscosities showing the influence of viscous and inertial forces. This behavior was also observed when using a simple

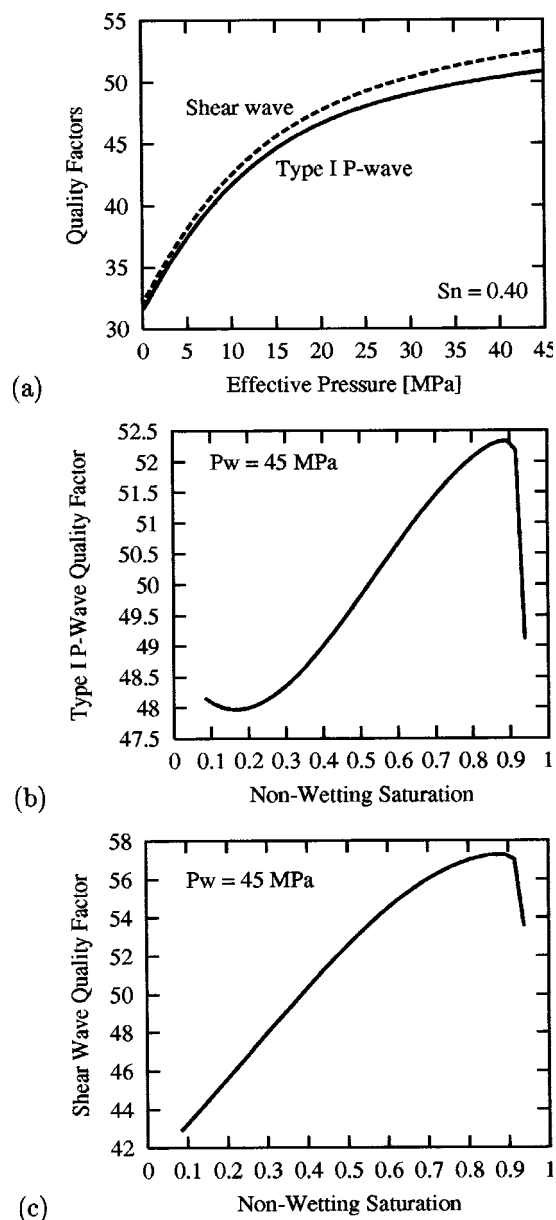


FIG. 7. Type-I P- and S-wave quality factors versus effective pressure (a), and versus nonwetting fluid saturation [(b) and (c)].

isostress mixture model (i.e., with zero capillary forces), as described in Berryman *et al.*⁷ Both quality factors show an abrupt decrease near the irreducible brine saturation, as can be appreciated in the corresponding wave velocities.

To our knowledge, no laboratory or field data of attenuation versus saturation for water–oil saturated sandstones are available. However, we can compare the predictions of our model with the observations made by different researchers in water–air (or water–gas) saturated rocks, considering water as the wetting phase. A careful analysis of the results published by Murphy III⁴⁵ for Massillon sandstone, of the averages obtained by Jones *et al.*⁴⁶ using a set of nine sandstones, and of the discussion in Schön⁴⁷ (Sec. 7.3.4.1), allows us to conclude that Q_S increases with \bar{S}_n , and that Q_I has a minimum at intermediate saturation states.

The quality factors for type-II and type-III compressional waves are shown in Figs. 8(a) and (b). Although they

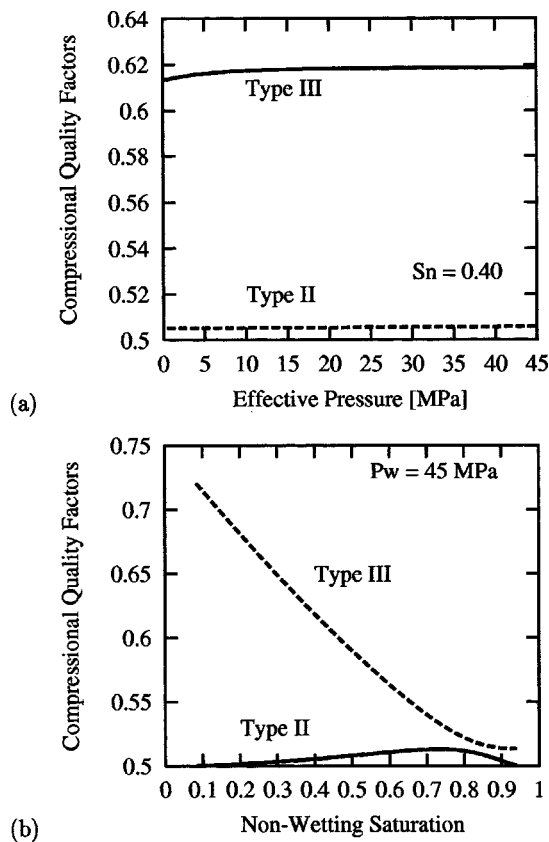


FIG. 8. Type-II and type-III P-wave quality factors versus effective pressure (a), and versus nonwetting fluid saturation (b).

are not very sensitive to effective pressure, they show appreciable variations in the whole saturation range.

V. CONCLUSIONS, REMARKS, AND FUTURE WORK

We have investigated the effects of effective pressure and fluid saturation on different static and dynamic properties of reservoir rocks. We have analyzed the behavior of a hydrocarbon (oil–brine) transition zone in a well consolidated sandstone, incorporating a realistic capillary pressure curve. The computations include the estimation of bulk volume compressibilities, effective pressure, ultrasonic phase velocities, and quality factors of the different wave modes, under different fluid pressures and saturations. Our main results can be summarized as follows.

- (1) The effective pressure versus saturation curve is almost constant, showing an abrupt decrease near the maximum admissible value of the nonwetting saturation (i.e., the irreducible brine saturation). Such a decrease is reflected in both the fast P- and S-wave velocities and quality factors. Unfortunately, nothing can be concluded from the available published data, since it requires detailed information about the fluid pressure conditions during the experiments.
- (2) For normal–hydrostatic–to intermediate wetting fluid pressures, the results obtained with classical formulas for the effective pressure (such as the differential pressure law of Gangi–Carlson³¹), may differ substantially from our estimations. For highly overpressured reservoirs

(high \bar{P}_w , low P_{wd}), however, the differential pressure of the wetting phase gives a good estimate of the effective pressure, extending a result valid also for single-phase fluid saturated rocks.³¹

- (3) The behavior of the bulk compressibilities versus effective pressure are similar to those observed in fully saturated rocks, and are controlled by the decrease of the matrix compressibility with pressure. In particular, it must be noted the magnitude of the new compressibility C_{bca} , indicating the influence that capillary pressure variations may have on bulk volume changes. This is particularly important in the range of intermediate to high nonwetting saturations, and for abnormally high pore pressures (low effective pressures).
- (4) Regarding the type-I (fast) P and S waves, our model simulates the increase in velocities and quality factors with effective pressure observed in laboratory experiments. Consequently, these quantities can be good indicators of abnormally high fluid pressures. The behavior of Q_1 and Q_5 for variable nonwetting (oil) saturation is, in general, coherent with observations made in similar situations (except at the residual brine saturation), and was also verified using an isostress mixture model for the fluids. Our results suggest that, unlike the corresponding phase velocity, the shear wave quality factor can be a good fluid indicator, particularly for normal (hydrostatic) to intermediate pressure conditions.
- (5) The velocities and quality factors of the type-II and type-III compressional waves are not very sensitive to pressure variations, but show a strong dependence on saturation.
- (6) The influence of capillary pressure on the acoustic properties of the different wave modes, was analyzed by comparing our results with the simplified isostress model—analogue to Biot's classical theory, in which only type-I, type-II P and S waves can propagate. We observe that, except at the irreducible brine saturation (where capillary pressure reaches its maximum), there is almost no difference between type-I P- and S-wave velocities. For type-II P waves, however, important differences can be observed. This seems reasonable since the first two waves involve coupled solid–fluid motions, while the latter is affected by the fluid mobility. Appreciable differences are observed in the quality factors given by the two models, indicating that the quality factors are more sensitive to capillary pressure than the phase velocities. This suggests that quality factor analysis can be used as a tool for discriminating full saturation from two-phase or partial saturation.

We must point out some limitations of the model, since due to the lack of data, we ignored the possible dependence of frame moduli with capillary pressure, an effect that should be carefully investigated. The same holds for the coefficients \hat{Q}_{Kc} and \hat{Q}_N , which for simplicity were taken independent of saturation. We also neglected the variations of petrophysical quantities such as porosity and permeability with effective pressure. Consequently, these results should be compared to appropriate laboratory experiments or field data.

Moreover, since unconsolidated reservoir sandstones are expected to show large sensitivity to fluid type and effective pressure,³³ a similar theoretical and applied analysis should be performed for such cases.

Finally, a rock in the subsurface is subjected to a non-hydrostatic state of stress: in general, the vertical stress is greater than the horizontal stress, and this situation induces P- and S-wave anisotropy in an otherwise isotropic rock.⁴⁸ Therefore, the theory should be generalized to describe anisotropy, and its related pressure effects.

ACKNOWLEDGMENTS

This work was funded by the Agencia Nacional de Promoción Científica y Tecnológica, Argentina (Contract No. BID-802/OC-AR), the European Union under the project "Detection of overpressured zones from seismic and well data" and CONICET, Argentina (PIP 0363/98).

- ¹M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low frequency range," *J. Acoust. Soc. Am.* **28**, 168–171 (1956).
- ²M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. II. High frequency range," *J. Acoust. Soc. Am.* **28**, 179–191 (1956).
- ³M. A. Biot, "Mechanics of deformation and acoustic propagation in porous media," *J. Appl. Phys.* **33**, 1482–1498 (1962).
- ⁴N. C. Dutta and H. Odé, "Attenuation and dispersion of compressional waves in fluid-filled porous rocks with partial gas saturation (White model)—Part I: Biot theory," *Geophysics* **44**, 1777–1788 (1979).
- ⁵J. E. White, "Computed seismic speed and attenuation in rocks with partial gas saturation," *Geophysics* **40**, 224–232 (1975).
- ⁶S. Mochizuki, "Attenuation in partially saturated rocks," *J. Geophys. Res.* **87**, 8598–8604 (1982).
- ⁷J. Berryman, L. Thigpen, and R. Chin, "Bulk elastic wave propagation in partially saturated porous solids," *J. Acoust. Soc. Am.* **84**, 360–373 (1988).
- ⁸M. N. Toksöz, C. H. Cheng, and A. Timur, "Velocities of seismic waves in porous rocks," *Geophysics* **41**, 621–645 (1976).
- ⁹J. E. Santos, J. Douglas, Jr., J. M. Corberó, and O. M. Lovera, "A model for wave propagation in a porous medium saturated by a two-phase fluid," *J. Acoust. Soc. Am.* **87**, 1439–1448 (1990).
- ¹⁰J. E. Santos, J. M. Corberó, and J. Douglas, Jr., "Static and dynamic behavior of a porous solid saturated by a two-phase fluid," *J. Acoust. Soc. Am.* **87**, 1428–1438 (1990).
- ¹¹J. M. Carcione and U. Tinivella, "Bottom-simulating reflectors: seismic velocities and AVO effects," *Geophysics* **65**, 54–67 (2000).
- ¹²U. Tinivella and J. M. Carcione, "Estimation of gas-hydrate concentration and free-gas saturation from log and seismic data," *The Leading Edge* **20**, 200–203 (2001).
- ¹³*Advances in Porous Media*, edited by M. Yavuz Corapcioglu (Elsevier, Amsterdam, 1996).
- ¹⁴H. B. Helle, "Detection of overpressure zones from seismic and well data, A summary of activities," Norsk Hydro Report No. R-089689, Norsk Hydro E & P Research Centre, N-5020 Bergen, Norway, 1999.
- ¹⁵J. M. Carcione and A. Gangi, "Gas generation and overpressure: effects on seismic attributes," *Geophysics* **65**, 1769–1779 (2000).
- ¹⁶J. M. Carcione and A. Gangi, "Non-equilibrium compaction and abnormal pore-fluid pressures: effects on rock properties," *Geophys. Prospect.* **48**, 521–537 (2000).
- ¹⁷M. S. King, "Wave velocities in rocks as a function of changes in overburden pressure and pore fluid saturants," *Geophysics* **31**, 50–73 (1966).
- ¹⁸J. Bear, *Dynamics of Fluids in Porous Media* (Dover, New York, 1972).
- ¹⁹R. E. Collins, *Flow of Fluids Through Porous Materials* (Reinhold, New York, 1961).
- ²⁰A. Scheidegger, *The Physics of Flow Through Porous Media* (MacMillan Company, Toronto, 1974).
- ²¹D. W. Peaceman, *Fundamentals of Numerical Reservoir Simulation* (Elsevier, Amsterdam, 1977).
- ²²R. W. Zimmerman, *Compressibility of Sandstones* (Elsevier Science, Amsterdam, 1991).
- ²³About the differences in notation between this paper and the former (Ref. 10), it must be mentioned that the old variables ΔP , S_0 , and e now correspond to ΔP_c , S_n , and e_b .
- ²⁴J. L. Auriault, O. Lebaigue, and G. Bonnet, "Dynamics of two immiscible fluids flowing through deformable porous media," *Transp. Porous Media* **4**, 105–128 (1989).
- ²⁵J. E. Santos, J. M. Corberó, C. L. Ravazzoli, and J. L. Hensley, "Reflection and transmission coefficients in fluid saturated porous media," *J. Acoust. Soc. Am.* **91**, 1911–1923 (1992).
- ²⁶F. Gassmann, "Über die elastizität poröser medien," *Vierteljahrsschr Naturforsch. Ges. Zurich* **96**, 1–23 (1951).
- ²⁷R. L. Mann and I. Fatt, "Effect of pore fluids on the elastic properties of sandstone," *Geophysics* **25**, 433–444 (1960).
- ²⁸J. Douglas, Jr., J. L. Hensley, and P. J. Paes Leme, "A study of the effect of inhomogeneities on immiscible flow in naturally fractured reservoirs," *International Series of Numerical Mathematics* (Birkhauser-Verlag, Basel, 1993), Vol. 114, pp. 59–74.
- ²⁹J. Berryman, "Elastic wave propagation in fluid-saturated porous solids," *J. Acoust. Soc. Am.* **69**, 416–424 (1981).
- ³⁰J. M. Hovem and G. D. Ingram, "Viscous attenuation of sound in saturated sand," *J. Acoust. Soc. Am.* **66**, 1807–1812 (1979).
- ³¹A. F. Gangi and R. L. Carlson, "An asperity-deformation model for effective pressure," *Tectonophysics* **256**, 241–251 (1996).
- ³²C. L. Ravazzoli and J. E. Santos, "Compressibility analysis of Berea sandstone versus saturation and effective pressure," *EAGE 62nd Conference and Technical Exhibition*, Glasgow, 29 May–2 June 2000, Extended Abstracts Book, Paper D-36, 2000.
- ³³R. Guerra, J. Meyer, A. Sibbit, and R. Van Delden, "Integrated time-lapse rock physics measurements—the key to understanding reservoir elastic moduli changes with hydrocarbon production," *Proceedings of PARIS 2000, Petrophysics Meets Geophysics*, 2000, Paper A-2.
- ³⁴M. A. Biot, "Theory of deformation of a porous viscoelastic anisotropic solid," *J. Appl. Phys.* **27**, 459–467 (1956).
- ³⁵R. D. Stoll, "Acoustic waves in saturated sediments," in *Physics of Sound in Marine Sediments*, edited by L. Hampton (Plenum, New York, 1974), pp. 19–39.
- ³⁶R. D. Stoll and G. M. Bryan, "Wave attenuation in saturated sediments," *J. Acoust. Soc. Am.* **47**, 1440–1447 (1970).
- ³⁷J. D. Keller, "Acoustic wave propagation in composite fluid-saturated media," *Geophysics* **54**, 1554–1563 (1989).
- ³⁸P. Rasolofosaon, "Plane acoustic waves in linear viscoelastic porous media: energy, particle displacement and physical interpretation," *J. Acoust. Soc. Am.* **89**, 1532–1550 (1991).
- ³⁹*Seismic Wave Attenuation*, edited by M. Nafi Toksöz and D. Johnston, Geophysics Reprint Series No. 2, Society of Exploration Geophysicists, 1981.
- ⁴⁰H. P. Liu, D. L. Anderson, and H. Kanamori, "Velocity dispersion due to anelasticity; implications for seismology and mantle composition," *Geophys. J. R. Astron. Soc.* **47**, 41–58 (1976).
- ⁴¹C. L. Ravazzoli, "Modelling of wave propagation phenomena in dissipative media," Ph.D. Thesis, Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, Argentina, 1995 (in Spanish).
- ⁴²J. M. Carcione, *Wave Fields in Real Media: Wave Propagation in Anisotropic, Anelastic and Porous Media* (Pergamon, Amsterdam, 2001).
- ⁴³K. Winkler and A. Nur, "Pore fluids and seismic attenuation in rocks," *Geophys. Res. Lett.* **6**, 1–4 (1979).
- ⁴⁴G. Tao, M. S. King, and M. Nabi-Bidhendi, "Ultrasonic wave propagation in dry and brine-saturated sandstones as a function of effective stress: laboratory measurements and modelling," *Geophys. Prospect.* **43**, 299–327 (1995).
- ⁴⁵W. F. Murphy III, "Effects of partial water saturation on attenuation in Massillon sandstone and Vycor porous glass," *J. Acoust. Soc. Am.* **71**, 1458–1468 (1982).
- ⁴⁶S. Jones, T. Astin, and C. Mc Caan, "The effect of degree of saturation on ultrasonic velocity and attenuation in sandstones," *Annual Meeting Abstracts*, Society of Exploration Geophysicists, pp. 1023–1026.
- ⁴⁷J. H. Schön, *Physical Properties of Rocks: Fundamentals and Principles of Petrophysics*, Handbook of Geophysical Exploration (Elsevier Science, Amsterdam, 1996), Vol. 18.
- ⁴⁸J. M. Carcione, "Energy balance and fundamental relations in dynamic anisotropic poroviscoelasticity," *Proc. R. Soc. London, Ser. A* **457**, 331–348 (2000).

Traction-free vibrations of finite trigonal elastic cylinders

Paul R. Heyliger

Department of Civil Engineering, Colorado State University, Fort Collins, Colorado 80523

Ward L. Johnson^{a)}

Materials Reliability Division, National Institute of Standards and Technology, Boulder, Colorado 80305

(Received 15 March 2002; revised 14 December 2002; accepted 20 December 2002)

The unrestrained, traction-free vibrations of finite elastic cylinders with trigonal material symmetry are studied using two approaches, based on the Ritz method, which formulate the weak form of the equations of motion in cylindrical and rectangular coordinates. Elements of group theory are used to divide approximation functions into orthogonal subsets, thus reducing the size of the computational problem and classifying the general symmetries of the vibrational modes. Results for the special case of an isotropic cylinder are presented and compared with values published by other researchers. For the isotropic case, the relative accuracy of the formulations in cylindrical and rectangular coordinates can be evaluated, because exact analytical solutions are known for the torsional modes. The calculation in cylindrical coordinates is found to be more accurate for a given number of terms in the series approximation functions. For a representative trigonal material, langatate, calculations of the resonant frequencies and the sensitivity of the frequencies on each of the elastic constants are presented. The dependence on geometry (ratio of length to diameter) is briefly explored. The special case of a transversely isotropic cylinder (with the elastic stiffness C_{14} equal to zero) is also considered. © 2003 Acoustical Society of America.
[DOI: 10.1121/1.1548159]

PACS numbers: 43.20.Ks, 43.40.Cw [ANN]

I. INTRODUCTION

Studies of the free vibration of finite elastic cylinders have been directed primarily toward those with isotropic material symmetry. Pochhammer¹ and Chree^{2,3} developed exact solutions to the equations of motion for isotropic cylinders of infinite length with stress-free radial faces. Hutchinson^{4–6} extensively studied the more complicated problem of a finite cylinder having all faces stress-free, solving the differential equations of motion exactly with respect to three of the six boundary conditions and approximately with respect to the remaining three. Visscher and co-workers,⁷ So and Leissa,⁸ Heyliger,⁹ and Heyliger and Jilani¹⁰ used the Ritz method to study the free vibration of isotropic and orthotropic cylinders of finite length, providing a number of useful results and observations.

In this study, this latter group of works involving the Ritz method is extended to include cases where the material belongs to one of the higher-symmetry trigonal crystal classes, C_{3v} , D_3 , or D_{3d} in the Schoenflies notation ($3m$, 32 , or $\bar{3}m$ in the international notation). The problem is considered in both rectangular and cylindrical coordinates. The only past treatments of trigonal crystals using the Ritz method are those of Ohno *et al.*,¹¹ who considered specimens in the form of rectangular parallelepipeds, and Mochizuki,¹² who considered spherical specimens. Our approach using rectangular coordinates is similar to that of Ohno with respect to the symmetries of the approximating functions that are employed in the calculation. For the cylindrical geometry, this approach does not separate the vibrational modes

into orthogonal symmetrically similar sets of minimal size (corresponding to the group-theoretical irreducible representations). Because of the relatively complicated symmetry transformations of rectangular coordinates associated with the threefold crystal axis, there does not appear to be any useful scheme for separating these approximating functions into minimally sized sets. Thus, the matrices involved in solving the vibrational problem in rectangular coordinates are larger than they need to be, and the symmetries of the modes are not fully characterized. For this reason, an alternate approach is pursued using cylindrical coordinates. This approach allows approximating functions to be separated into minimally sized sets using group-theoretical projection operators and, thus, provides a complete categorization of the symmetries of the resultant vibrational modes.

One common application of the Ritz method is in the least-squares determination of elastic constants from measured vibrational frequencies of crystals (the method of resonant ultrasound spectroscopy¹³). With respect to this application, the cylindrical geometry has advantages over other geometries for trigonal crystals, and this provides much of the motivation for pursuing the present work. In comparison to parallelepipeds, the cylindrical geometry eliminates the uncertainty associated with alignment of a crystal axis perpendicular to the trigonal axis (only one axis needs to be aligned, instead of two), thus reducing uncertainty in extracted elastic constants. For very accurate measurements, the fabrication of specimens with tight dimensional tolerances is more economical for cylinders than for spheres.

In Sec. II, the weak formulation of the equations of motion is presented in both rectangular and cylindrical coordinates. Then, Ritz-type approximations for the three displace-

^{a)}Electronic mail: wjohnson@boulder.nist.gov

ment components in the three independent coordinate directions are introduced and grouped according to symmetry to split the resulting eigenvalue problem. In Sec. III, calculations are performed for the special case of an isotropic cylinder and compared with those of So and Leissa.⁸ Then, the central problem of trigonal crystals is considered, using langatate as an example. Representative frequencies and modal plots are presented for specific geometries. The two formulations of the problem give nearly identical results, but there are important differences that are discussed. Calculations are also performed for the special case of transverse isotropy and compared with those for trigonal and isotropic symmetry.

II. THEORY

A. Geometry

The object of interest here is a finite circular cylinder composed of elastic material with trigonal symmetry. The principal geometric directions of the cylinder align with those of the cylindrical coordinate system ($x_1=r$, $x_2=\theta$, $x_3=z$), and the three displacements associated with these directions are denoted as $u_1=u_r$, $u_2=u_\theta$, and $u_3=u_z$. In an alternate rectangular Cartesian coordinate system, the standard coordinates ($x_1=x$, $x_2=y$, $x_3=z$) are used along with the displacement components ($u_1=u_x$, $u_2=u_y$, and $u_3=u_z$). The height of the cylinder in the axial direction z is H . The cylinder has radius R , diameter D , and density ρ and is treated as a homogeneous material. The boundary conditions are such that all external faces of the cylinder are traction free.

B. Variational formulation

The starting point for the variational formulation is Hamilton's principle for a linear elastic medium free of body forces, expressed as¹⁴

$$\delta \int_{t_0}^t dt \int_V \left[\frac{1}{2} \rho \dot{u}_j \dot{u}_j - U_0 \right] dV + \int_{t_0}^t dt \int_S \bar{t}_k \delta u_k dS = 0, \quad (1)$$

where t is time, V and S are the volume and surface occupied by and bounding the solid, \bar{t}_k are the components of the specified surface tractions, δ is the variational operator, and the overdot represents differentiation with respect to time. U_0 represents the strain energy density and is given by

$$U_0 = \frac{1}{2} C_{ijkl} \epsilon_{ij} \epsilon_{kl} = \frac{1}{2} \bar{C}_{ijkl} \bar{\epsilon}_{ij} \bar{\epsilon}_{kl}. \quad (2)$$

As usual, summations over repeated indices are implicit. The components of the elastic stiffness tensor are denoted as C_{ijkl} for the rectangular coordinate system and \bar{C}_{ijkl} for the cylindrical coordinate system. Similarly, ϵ_{ij} and $\bar{\epsilon}_{ij}$ are the strain components in the rectangular and cylindrical systems, respectively. The relationship between stress σ_{ij} and strain in rectangular coordinates is

$$\sigma_{ij} = C_{ijkl} \epsilon_{kl}, \quad (3)$$

and the corresponding relationship in cylindrical coordinates is

$$\bar{\sigma}_{ij} = \bar{C}_{ijkl} \bar{\epsilon}_{kl}. \quad (4)$$

At this point, the rectangular and cylindrical formulations of the problem diverge. The variational equations in the two coordinate systems are presented in Secs. II B1 and II B2. The Ritz approximation functions for the two formulations and symmetry-based groupings of these functions are presented in Sec. II C.

1. Rectangular coordinates

The general constitutive relation (Eq. (3)) for the higher trigonal crystal classes can be expressed in rectangular Cartesian coordinates as

$$\begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{Bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & 0 & 0 \\ C_{12} & C_{22} & C_{23} & C_{24} & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ C_{14} & C_{24} & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & C_{56} \\ 0 & 0 & 0 & 0 & C_{56} & C_{66} \end{bmatrix} \begin{Bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{Bmatrix}, \quad (5)$$

where

$$C_{22} = C_{11}, \quad (6a)$$

$$C_{23} = C_{13}, \quad (6b)$$

$$C_{24} = -C_{14}, \quad (6c)$$

$$C_{55} = C_{44}, \quad (6d)$$

$$C_{56} = C_{14}, \quad (6e)$$

$$C_{66} = \frac{1}{2}(C_{11} - C_{12}) \quad (6f)$$

(with six independent elastic constants). The conventional contracted notation ($\sigma_1 = \sigma_{11}$, $\sigma_4 = \sigma_{23}$, $\epsilon_1 = \epsilon_{11}$, $\epsilon_4 = 2\epsilon_{23}$, $C_{1111} = C_{11}$, $C_{1123} = C_{14}$, etc.) is used here, and the 1, 2, and 3 directions are \hat{x} , \hat{y} , and \hat{z} .

Hamilton's principle in the contracted notation becomes

$$\begin{aligned} 0 = & - \int_0^t \int_V \{ \sigma_1 \delta \epsilon_1 + \sigma_2 \delta \epsilon_2 + \sigma_3 \delta \epsilon_3 \\ & + \sigma_4 \delta \epsilon_4 + \sigma_5 \delta \epsilon_5 + \sigma_6 \delta \epsilon_6 \} dV dt \\ & + \frac{1}{2} \delta \int_0^t \int_V \rho (\dot{u}_x^2 + \dot{u}_y^2 + \dot{u}_z^2) dV dt. \end{aligned} \quad (7)$$

For the problem of unrestrained free vibration, there are no applied external loads and each face of the solid is assumed to be stress free. Hence, the surface integral in Eq. (1) is zero in this case. The strain-displacement relations are

$$\epsilon_1 = \epsilon_{11} = \frac{\partial u_x}{\partial x}, \quad (8a)$$

$$\epsilon_2 = \epsilon_{22} = \frac{\partial u_y}{\partial y}, \quad (8b)$$

$$\epsilon_3 = \epsilon_{33} = \frac{\partial u_z}{\partial z}, \quad (8c)$$

$$\epsilon_4 = 2\epsilon_{23} = \frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}, \quad (8d)$$

$$\epsilon_5 = 2\epsilon_{13} = \frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}, \quad (8e)$$

$$\epsilon_6 = 2\epsilon_{12} = \frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y}. \quad (8f)$$

Invoking time dependence of the displacement fields consistent with periodic motion and substituting the strain–displacement relations [Eq. (8)] and the stress–strain relations [Eq. (5)] into Eq. (7) lead to

$$\begin{aligned} 0 = \int_V \left\{ \left[C_{11} \frac{\partial u_x}{\partial x} + C_{12} \frac{\partial u_y}{\partial y} + C_{13} \frac{\partial u_z}{\partial z} + C_{14} \left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y} \right) \right] \frac{\partial \delta u_x}{\partial x} + \left[C_{12} \frac{\partial u_x}{\partial x} + C_{22} \frac{\partial u_y}{\partial y} + C_{23} \frac{\partial u_z}{\partial z} + C_{24} \left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y} \right) \right] \frac{\partial \delta u_y}{\partial y} \right. \\ \left. + \left[C_{13} \frac{\partial u_x}{\partial x} + C_{23} \frac{\partial u_y}{\partial y} + C_{33} \frac{\partial u_z}{\partial z} \right] \frac{\partial \delta u_z}{\partial z} + C_{44} \left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y} \right) \left(\frac{\partial \delta u_y}{\partial z} + \frac{\partial \delta u_z}{\partial y} \right) + \left[C_{55} \left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x} \right) + C_{56} \left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \right] \right. \\ \left. \times \left(\frac{\partial \delta u_x}{\partial z} + \frac{\partial \delta u_z}{\partial x} \right) + C_{66} \left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \left(\frac{\partial \delta u_x}{\partial y} + \frac{\partial \delta u_y}{\partial x} \right) - \rho \omega^2 (u_x \delta u_x + u_y \delta u_y + u_z \delta u_z) \right\} dV. \end{aligned} \quad (9)$$

There are an infinite number of solutions to this weak form of the equations of motion, each with a unique displacement function \mathbf{u} and a frequency ω that may be unique or may match that of one other mode. To determine approximate solutions, a general series form of the displacement components must be assumed. This step is discussed in the following for rectangular and cylindrical coordinates, after first deriving the weak form of the equations of motion in cylindrical coordinates.

2. Cylindrical coordinates

The strain–displacement relations in cylindrical coordinates are¹⁵

$$\bar{\epsilon}_1 = \epsilon_{rr} = \frac{\partial u_r}{\partial r}, \quad (10a)$$

$$\bar{\epsilon}_2 = \epsilon_{\theta\theta} = \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r}, \quad (10b)$$

$$\bar{\epsilon}_3 = \epsilon_{zz} = \frac{\partial u_z}{\partial z}, \quad (10c)$$

$$\bar{\epsilon}_4 = 2\epsilon_{\theta z} = \frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_z}{\partial \theta}, \quad (10d)$$

$$\bar{\epsilon}_5 = 2\epsilon_{rz} = \frac{\partial u_z}{\partial r} + \frac{\partial u_r}{\partial z}, \quad (10e)$$

$$\bar{\epsilon}_6 = 2\epsilon_{r\theta} = \frac{1}{r} \frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{\theta}{r}. \quad (10f)$$

The constitutive equations must be written in terms of the stress and strain components in the (r, θ, z) coordinates, and this involves transforming the components of the elastic stiffness tensor such that they form the link between the stress and strain in the cylindrical system, as represented in Eq. (4). This can be accomplished by using the standard transformation relations between fourth-order tensors:

$$\bar{C}_{ijkl} = a_{ip} a_{jq} a_{kr} a_{ls} C_{pqrs}, \quad (11)$$

where a_{ij} represents the elements of the direction cosines between the primed (cylindrical) and unprimed (rectangular) system. However, it is somewhat easier to use the Bond matrix¹⁵ $[B]$ to perform the transformation:

$$[\bar{C}] = [B][C][B]^T, \quad (12)$$

where

$$[B] = \begin{bmatrix} \cos^2 \theta & \sin^2 \theta & 0 & 0 & 0 & \sin 2\theta \\ \sin^2 \theta & \cos^2 \theta & 0 & 0 & 0 & -\sin 2\theta \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos \theta & -\sin \theta & 0 \\ 0 & 0 & 0 & \sin \theta & \cos \theta & 0 \\ -\frac{1}{2} \sin 2\theta & \frac{1}{2} \sin 2\theta & 0 & 0 & 0 & \cos 2\theta \end{bmatrix}. \quad (13)$$

θ is the angular position of the point being considered in the cylinder. From this transformation, the \bar{C}_{ij} in terms of the C_{ij} are given by

$$\bar{C}_{14} = \bar{C}_{56} = C_{14} \cos 3\theta, \quad (14a)$$

$$\bar{C}_{15} = \bar{C}_{25} = \bar{C}_{46} = C_{14} \sin 3\theta, \quad (14b)$$

$$\bar{C}_{24} = -\bar{C}_{14}. \quad (14c)$$

Note that these stiffnesses vary with position in the cylinder, since $[B]$ is a function of θ . The other C_{ij} are unchanged by the transformation.

The final matrix relationship between stress and strain in cylindrical coordinates in terms of the rotated elastic stiffnesses is

$$\begin{Bmatrix} \bar{\sigma}_1 \\ \bar{\sigma}_2 \\ \bar{\sigma}_3 \\ \bar{\sigma}_4 \\ \bar{\sigma}_5 \\ \bar{\sigma}_6 \end{Bmatrix} = \begin{Bmatrix} \sigma_{rr} \\ \sigma_{\theta\theta} \\ \sigma_{zz} \\ \sigma_{\theta z} \\ \sigma_{rz} \\ \sigma_{r\theta} \end{Bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & \bar{C}_{14} & \bar{C}_{15} & 0 \\ C_{12} & C_{22} & C_{23} & \bar{C}_{24} & \bar{C}_{25} & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ \bar{C}_{14} & \bar{C}_{24} & 0 & C_{44} & 0 & \bar{C}_{46} \\ \bar{C}_{15} & \bar{C}_{25} & 0 & 0 & C_{55} & \bar{C}_{56} \\ 0 & 0 & \bar{C}_{46} & \bar{C}_{56} & C_{66} & 0 \end{bmatrix} \begin{Bmatrix} \epsilon_{rr} \\ \epsilon_{\theta\theta} \\ \epsilon_{zz} \\ \gamma_{\theta z} \\ \gamma_{rz} \\ \gamma_{r\theta} \end{Bmatrix}. \quad (15)$$

The insertion of the strain–displacement and stress–strain relations [Eqs. (10) and (15)] into Hamilton’s principle yields the final variational form in terms of the elastic stiffnesses, density, geometry, and displacement components:

$$\begin{aligned} \int_V \left\{ \left[C_{11} \frac{\partial u_r}{\partial r} + C_{12} \left(\frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} \right) + C_{13} \frac{\partial u_z}{\partial z} + \bar{C}_{14} \left(\frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_r}{\partial \theta} \right) + \bar{C}_{15} \left(\frac{\partial u_z}{\partial r} + \frac{\partial u_r}{\partial z} \right) \right] \frac{\partial \delta u_r}{\partial r} + \left[C_{12} \frac{\partial u_r}{\partial r} + C_{22} \left(\frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} \right) \right. \right. \\ \left. \left. + C_{23} \frac{\partial u_z}{\partial z} + \bar{C}_{24} \left(\frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_r}{\partial \theta} \right) + \bar{C}_{25} \left(\frac{\partial u_z}{\partial r} + \frac{\partial u_r}{\partial z} \right) \right] \left(\frac{1}{r} \frac{\partial \delta u_\theta}{\partial \theta} + \frac{\delta u_r}{r} \right) + \left[C_{13} \frac{\partial u_r}{\partial r} + C_{23} \left(\frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} \right) + C_{33} \frac{\partial u_z}{\partial z} \right] \frac{\partial \delta u_z}{\partial z} \right. \\ \left. + \left[\bar{C}_{14} \frac{\partial u_r}{\partial r} + \bar{C}_{24} \left(\frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} \right) + C_{44} \left(\frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_r}{\partial \theta} \right) + \bar{C}_{46} \left(\frac{1}{r} \frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} \right) \right] \left(\frac{\partial \delta u_\theta}{\partial z} + \frac{1}{r} \frac{\partial \delta u_z}{\partial \theta} \right) \right. \\ \left. + \left[\bar{C}_{15} \frac{\partial u_r}{\partial r} + \bar{C}_{25} \left(\frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} \right) + C_{55} \left(\frac{\partial u_z}{\partial r} + \frac{\partial u_r}{\partial z} \right) + \bar{C}_{56} \left(\frac{1}{r} \frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} \right) \right] \left(\frac{\partial \delta u_z}{\partial r} + \frac{\partial \delta u_r}{\partial z} \right) + \left[\bar{C}_{46} \left(\frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_z}{\partial \theta} \right) \right. \right. \\ \left. \left. + \bar{C}_{56} \left(\frac{\partial u_z}{\partial r} + \frac{\partial u_r}{\partial z} \right) + C_{66} \left(\frac{1}{r} \frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} \right) \right] \left(\frac{1}{r} \frac{\partial \delta u_r}{\partial \theta} + \frac{\partial \delta u_\theta}{\partial r} - \frac{\delta u_\theta}{r} \right) - \rho \omega^2 (u_r \delta u_r + u_\theta \delta u_\theta + u_z \delta u_z) \right\} r dr d\theta dz = 0. \end{aligned} \quad (16)$$

As with the weak formulation in rectangular coordinates, it is this equation for which approximate solutions are sought for ω , u_r , u_θ , and u_z . This is discussed in Sec. II C.

C. Ritz approximations

In the Ritz method, the three displacements are approximated by finite linear combinations of the form

$$u_1(x_1, x_2, x_3) = \sum_{j=1} a_j \Psi_j^1(x_1, x_2, x_3), \quad (17a)$$

$$u_2(x_1, x_2, x_3) = \sum_{j=1} b_j \Psi_j^2(x_1, x_2, x_3), \quad (17b)$$

$$u_3(x_1, x_2, x_3) = \sum_{j=1} d_j \Psi_j^3(x_1, x_2, x_3), \quad (17c)$$

where the a_j , b_j , and d_j are unknown constants and the Ψ_j^i are known functions of the spatial coordinates used in the specific formulation. The corresponding approximations for the variations are¹⁶

$$\delta u_1(x_1, x_2, x_3) = \Psi_i^1(x_1, x_2, x_3), \quad (18a)$$

$$\delta u_2(x_1, x_2, x_3) = \Psi_i^2(x_1, x_2, x_3), \quad (18b)$$

$$\delta u_3(x_1, x_2, x_3) = \Psi_i^3(x_1, x_2, x_3). \quad (18c)$$

Substitution of Eqs. (17) and (18) into the weak forms given in Eq. (9) or (16) yields the matrix equation:

$$\begin{aligned} \begin{bmatrix} [M^{11}] & [0] & [0] \\ [0] & [M^{22}] & [0] \\ [0] & [0] & [M^{33}] \end{bmatrix} \begin{Bmatrix} \{a\} \\ \{b\} \\ \{d\} \end{Bmatrix} \omega^2 \\ + \begin{bmatrix} [K^{11}] & [K^{12}] & [K^{13}] \\ [K^{21}] & [K^{22}] & [K^{23}] \\ [K^{31}] & [K^{32}] & [K^{33}] \end{bmatrix} \begin{Bmatrix} \{a\} \\ \{b\} \\ \{d\} \end{Bmatrix} = \begin{Bmatrix} \{0\} \\ \{0\} \\ \{0\} \end{Bmatrix}, \end{aligned} \quad (19)$$

where

$$M_{ij}^{11} = \int_V \rho \Psi_i^1 \Psi_j^1 dV, \quad (20a)$$

$$M_{ij}^{22} = \int_V \rho \Psi_i^2 \Psi_j^2 dV, \quad (20b)$$

$$M_{ij}^{33} = \int_V \rho \Psi_i^3 \Psi_j^3 dV. \quad (20c)$$

General expressions for the coefficient matrices K^{nm} are given in Appendix A for rectangular and cylindrical coordinates, without restrictions on the specimen geometry or form of the approximation functions. The matrices in rectangular coordinates reduce to those presented by other authors for specific forms of the approximation functions. The matrices presented for cylindrical coordinates include, for the first time, the explicit dependence on θ and, therefore, are applicable to basis functions having any form of θ dependence.

In the following, the approximation functions used in rectangular and cylindrical coordinates are presented. In both

cases, the resulting terms are evaluated using a combination of exact and numerical integration, and the resulting eigenvalue problem is solved using the QR¹⁶ algorithm.

The eigenvalue problem, Eq. (19), generally can be block-diagonalized, if the approximation functions are grouped into appropriately symmetrized orthogonal sets. This accomplishes at least two useful tasks. First, it reduces the size of the computational problem, yielding identical results with only a fraction of the computational cost. Second, it allows consideration and study of the different types of vibrational modes based on their general symmetry properties.

The grouping of the approximation functions into orthogonal sets is dependent on the geometry and crystal symmetry. In this study, the focus is on the trigonal crystals classes C_{3v} , D_3 , and D_{3d} . The macroscopic cylindrical shape maintains all the symmetry operations of a trigonal crystal when the threefold axis is oriented along the cylinder axis. Therefore, the total symmetry of the specimen (including both shape and crystal structure) is the same as that of the crystal.

Inversion symmetry is not included in the C_{3v} or D_3 crystal classes. However, linear elastic vibrations are insensitive to a lack of inversion, so that the effective symmetry for linear vibrational problems does include inversion, and C_{3v} and D_3 become equivalent to D_{3d} . Therefore, the following discussion refers only to D_{3d} . The fact that inversion always can be included in the effective set of symmetry operations is seen by considering that, in Hamilton's principle [Eq. (1)], the symmetry of the crystal enters only through the general form of the elastic constant matrix. It is easy to show that the elastic-constant matrix for any crystal is unchanged by inversion regardless of whether the crystal has inversion symmetry. (See Ref. 15, Vol. I, Chap 7.B, for a description of matrix operations corresponding to crystal-symmetry transformations.)

1. Rectangular coordinates

Visscher and co-workers⁷ developed a powerful approach for the general solution of the weak form of nontrigonal material systems formulated in rectangular Cartesian coordinates. For each of the three displacement components, the following class of basis functions was selected:

$$\Psi(x, y, z) = x^i y^j z^k, \quad (21)$$

where i , j , and k are integers. The primary advantage of casting the problem in this form is that the integrals required by the weak form are simple to evaluate analytically. For a cylindrical specimen, the volume integrals that arise from inserting Eq. (21) into Eq. (9) have the form⁷

$$\int_V x^i y^j z^k dV = 4\pi \left(\frac{H}{2}\right)^{i+1} R^{j+k+2} \frac{(i-1)!!(j-1)!!}{(i+1)(j+k+2)!!}. \quad (22)$$

When formulating the problem for a rectangular parallelepiped of a trigonal crystal, Ohno *et al.*¹¹ separated ap-

TABLE I. Parities of indices of approximation functions in rectangular Cartesian coordinates [Eq. (21)] separated into subsets having the symmetry properties of the irreducible representations of C_{2h} . E =even. O =odd.

Subset	Component	i	$j+k$
A_g	u_x	O	E
	u_y	E	O
	u_z	E	O
A_u	u_x	E	E
	u_y	O	O
	u_z	O	O
B_g	u_x	E	O
	u_y	O	E
	u_z	O	E
B_u	u_x	O	O
	u_y	E	E
	u_z	E	E

proximation functions into four orthogonal subsets corresponding to the irreducible representations of group theory. This analysis began by identifying the effective symmetry of the specimen as C_{2h} ($2/m$, monoclinic). Their division of approximating functions is valid for any specimen possessing the symmetry operations of C_{2h} , including the cylindrical D_{3d} specimens considered here (which possess additional symmetry operations). In our treatment of D_{3d} cylinders in rectangular coordinates, approximating functions with the same general symmetry properties as those of Ohno *et al.* are employed, but with different specific functional form [the power-series form of Eq. (21), instead of Legendre polynomials]. The scheme for dividing these functions into subsets is shown in Table I. The subset labels A_g , B_g , A_u , and B_u correspond to the those of Ohno *et al.* and follow the conventional group-theoretical designation of the irreducible representations of C_{2h} . The subscripts "g" and "u" indicate that the corresponding approximating functions are even (*gerade*) or odd (*ungerade*) under inversion. In Table I, the column labeled " i " indicates whether this index in the approximating functions [Eq. (21)] is even (E) or odd (O). The column labeled " $j+k$ " indicates whether the sum of these indices in Eq. (21) is even or odd.

The scheme presented in Table I does not divide the approximating functions into minimally sized subsets corresponding to the irreducible representations of D_{3d} , because it does not consider differences in the way the functions transform under the threefold rotations of D_{3d} (see Appendix B). The difficulty in dividing the functions into subsets arises from the fact that the threefold rotations transform both the positions and the directions of the displacement vectors, so that the conditions on the power-series indices imposed by symmetry involve cross terms that are very complicated. This limitation is overcome by formulating the problem in cylindrical coordinates.

2. Cylindrical coordinates

The approximation functions employed for cylindrical coordinates incorporate power series in r , Fourier series in θ , and power series in z :

$$u_r(r, \theta, z) = \sum_{n, \alpha, \beta} a_{n\alpha\beta}^c r^\alpha \cos(n\theta) z^\beta + \sum_{m, \kappa, \gamma} a_{m\kappa\gamma}^s r^\kappa \sin(m\theta) z^\gamma, \quad (23a)$$

$$u_\theta(r, \theta, z) = \sum_{n, \alpha, \beta} b_{n\alpha\beta}^c r^\alpha \cos(n\theta) z^\beta + \sum_{m, \kappa, \gamma} b_{m\kappa\gamma}^s r^\kappa \sin(m\theta) z^\gamma, \quad (23b)$$

$$u_z(r, \theta, z) = \sum_{n, \alpha, \beta} d_{n\alpha\beta}^c r^\alpha \cos(n\theta) z^\beta + \sum_{m, \kappa, \gamma} d_{m\kappa\gamma}^s r^\kappa \sin(m\theta) z^\gamma, \quad (23c)$$

where the summation for each index is over non-negative integers, and $a_{n\alpha\beta}^c$, $a_{m\kappa\gamma}^s$, $b_{n\alpha\beta}^c$, $b_{m\kappa\gamma}^s$, $d_{n\alpha\beta}^c$, and $d_{m\kappa\gamma}^s$ are constants to be determined from the Ritz analysis. For terms with $\alpha=0$, the constants $a_{n\alpha\beta}^c$, $a_{m\kappa\gamma}^s$, $b_{n\alpha\beta}^c$, $b_{m\kappa\gamma}^s$, $d_{n\alpha\beta}^c$, and $d_{m\kappa\gamma}^s$ must generally be zero to avoid singular strains at the origin. However, functions of the form $z^\beta \cos \theta$ ($n=1, \alpha=0$) and $z^\gamma \cos \theta$ ($m=1, \alpha=0$) are not singular at the origin, since they represent rigid-body motion of the cylinder and result in zero strain. These terms must be included in the analysis to ensure convergence of this type of series and to represent the rigid-body modes of the cylinder. So and Leissa⁸ discuss this subject for the case of isotropic cylinders.

Simultaneously considering the complete range of indices over r , θ , and z would give the entire frequency spectrum of the cylinder. However, the goal is again to split the physical and mathematical problem into a collection of smaller problems. In group-theoretical terminology, the subsets of functions presented in Table I for the vibrational problem in rectangular coordinates provide bases for reducible representations of the point group D_{3d} . However, as indicated by the term “reducible,” these subsets are not of minimal size. D_{3d} has six irreducible representations with the conventional designations¹⁷ A_{1g} , A_{2g} , E_g , A_{1u} , A_{2u} , and E_u . The relation of these irreducible representations to those presented earlier for C_{2h} is described in Appendix B. A_{1g} , A_{2g} , A_{1u} , and A_{2u} are singly degenerate. That is, barring accidental degeneracies (unrelated to general symmetry properties), each vibrational mode belonging to any of these irreducible representations has a unique frequency. E_g and E_u are each doubly degenerate.

The set of cylindrical-coordinate approximating functions [Eq. (23)] can be separated into subsets corresponding to the six irreducible representations of D_{3d} using group-theoretical projection operators. This technique, which is fairly straightforward but laborious, is described in a separate paper.¹⁸ The results of the symmetry analysis are restrictions on the indices of the θ and z dependence of u_r , u_θ , and u_z , which are listed in Table II. There are no restrictions on the powers of r [α and κ in Eq. (23)]. For E_g , only one mode of each degenerate pair is included in Table II. The other mode

TABLE II. Parities of indices of approximation functions in cylindrical coordinates [Eq. (23)] separated into subsets having the symmetry properties of the irreducible representations of D_{3d} . $E \equiv$ even. $O \equiv$ odd. The values of n and m are additionally restricted to multiples of 3 for the nondegenerate irreducible representations (A_{1g} , A_{2g} , A_{1u} , A_{2u}) and are restricted to integers that are not multiples of 3 for the doubly degenerate irreducible representations (E_g and E_u).

Subset	Component	n	β	m	γ
A_{1g}	u_r	E	E	O	O
	u_θ	O	O	E	E
	u_z	E	O	O	E
A_{2g}	u_r	O	O	E	E
	u_θ	E	E	O	O
	u_z	O	E	E	O
E_g	u_r	O	O	E	E
	u_θ	E	E	O	O
	u_z	O	E	E	O
A_{1u}	u_r	O	E	E	O
	u_θ	E	O	O	E
	u_z	O	O	E	E
A_{2u}	u_r	E	O	O	E
	u_θ	O	E	E	O
	u_z	E	E	O	O
E_u	u_r	E	O	O	E
	u_θ	O	E	E	O
	u_z	E	E	O	O

in each E_g pair has the parities of all the indices reversed relative to those listed for E_g in Table II. The same is true for E_u modes.

The coefficients in the series expansions for the displacement patterns of two degenerate modes belonging to E_g or E_u are related. If one mode in a degenerate pair has the parities given in Table II and has a radial component given by Eq. (23a), then the radial component of the second function in this pair is given by¹⁸

$$u'_r(r, \theta, z) = \sum_{n, \alpha, \beta} \xi a_{n\alpha\beta}^c r^\alpha \sin(n\theta) z^\beta + \sum_{m, \kappa, \gamma} \eta a_{m\kappa\gamma}^s r^\kappa \cos(m\theta) z^\gamma, \quad (24)$$

where the $a_{n\alpha\beta}^c$ and $a_{m\kappa\gamma}^s$ are the coefficients of the first function, and

$$\xi = -\frac{2}{\sqrt{3}} \sin\left(\frac{2n\pi}{3}\right) = \begin{cases} -1 & \text{if } n = 1 + 3h \\ 1 & \text{if } n = 2 + 3h, \end{cases} \quad (25a)$$

$$\eta = \frac{2}{\sqrt{3}} \sin\left(\frac{2m\pi}{3}\right) = \begin{cases} 1 & \text{if } m = 1 + 3k \\ -1 & \text{if } m = 2 + 3k, \end{cases} \quad (25b)$$

where h and k are integers. The transformations of the u_θ and u_z components from the first function to the second function have the same form.

The volume integrals that appear in Eq. (16) can be evaluated analytically with approximation functions having the form of Eq. (23). The integrations in powers of r and z are straightforward, but those over the sine and cosine functions are less so when one considers that these functions must be multiplied by the rotated elastic stiffnesses, which

are themselves functions of θ . However, these integrals reduce to simple formulas when evaluated over the domain of the cylinder:

$$\int_0^{2\pi} \bar{C}_{14} \cos n\theta \cos m\theta d\theta = C_{14} \begin{cases} \pi & \text{if } m+n=3 \text{ and } |m-n|=3 \\ \frac{\pi}{2} & \text{if } m+n=3 \text{ and } |m-n|=1 \\ \frac{\pi}{2} & \text{if } m+n \neq 3 \text{ and } |m-n|=3 \\ 0 & \text{otherwise,} \end{cases} \quad (26a)$$

$$\int_0^{2\pi} \bar{C}_{14} \sin n\theta \sin m\theta d\theta = C_{14} \begin{cases} -\frac{\pi}{2} & \text{if } m+n=3 \text{ and } |m-n|=1 \\ \frac{\pi}{2} & \text{if } m+n \neq 3 \text{ and } |m-n|=3 \\ 0 & \text{otherwise,} \end{cases} \quad (26b)$$

$$\int_0^{2\pi} \bar{C}_{14} \sin n\theta \cos m\theta d\theta = 0, \quad (26c)$$

$$\int_0^{2\pi} \bar{C}_{15} \sin n\theta \cos m\theta d\theta = C_{15} \begin{cases} \pi & \text{if } m=0 \text{ and } n=3 \\ \frac{\pi}{2} & \text{if } m+n=3 \text{ and } |m-n|=1 \\ \frac{\pi}{2} & \text{if } n-m=3 \\ -\frac{\pi}{2} & \text{if } m-n=3 \\ 0 & \text{otherwise,} \end{cases} \quad (26d)$$

$$\int_0^{2\pi} \bar{C}_{15} \sin n\theta \sin m\theta d\theta = 0, \quad (26e)$$

$$\int_0^{2\pi} \bar{C}_{15} \cos n\theta \cos m\theta d\theta = 0. \quad (26f)$$

III. RESULTS

A. Isotropic cylinders

For the special case of homogeneous isotropic cylinders, the material symmetry allows for a separation of the approximating functions that simplifies the analysis. Although this problem is not of central interest here, it is considered to test the accuracy of the approach and demonstrate the difference between the two formulations. So and Leissa⁸ developed a highly accurate model that is a special case of our approach and applied it to comprehensively study the behavior of isotropic disks. Their results serve as a basis for comparison.

So and Leissa⁸ assert that the displacement field of a resonant mode of an isotropic cylinder involves only a single azimuthal wave number. In other words, the summations over n and m are eliminated in approximation functions having the form of Eq. (23), leaving only a single value of n . For $n > 0$, all modes are twofold degenerate, with displacement fields of each pair of modes having the form

$$\mathbf{u}_1(r, \theta, z) = U(r, z) \cos(n\theta) \hat{r} + V(r, z) \sin(n\theta) \hat{\theta} + W(r, z) \cos(n\theta) \hat{z}, \quad (27a)$$

$$\mathbf{u}_2(r, \theta, z) = U(r, z) \sin(n\theta) \hat{r} + V(r, z) \cos(n\theta) \hat{\theta} + W(r, z) \sin(n\theta) \hat{z}, \quad (27b)$$

with the same $U(r, z)$, $V(r, z)$, and $W(r, z)$ for both modes. For $n = 0$, the modes are singly degenerate, with either $U(r, z) = W(r, z) = 0$ (torsional modes) or $V(r, z) = 0$ (axisymmetric modes), since $\sin(n\theta)$ terms are zero.

The general form of Eq. (27) can also be deduced relatively easily from group theory [starting from the approximation functions given in Eq. (23)], but this is not pursued here. The symmetry operations of an isotropic cylinder belong to the Lie group $D_{\infty h}$, which has an infinite number of rotation operations about the z axis (since rotation by any angle about the cylinder axis leaves the object unchanged) and a correspondingly infinite number of irreducible representations. Functions of the form of Eq. (27) having different values of n belong to different irreducible representations. For $n > 0$, these irreducible representations are conventionally labeled E_{ng} and E_{nu} (with g and u again denoting the symmetry under inversion).¹⁷ For $n = 0$, there are four irreducible representations, A_{1g} , A_{2g} , A_{1u} , and A_{2u} , which correspond to the irreducible representations that have the same labels in the trigonal group D_{3d} .¹⁷

Because approximation functions with different values of n belong to different group-theoretical irreducible representations, the determination of resonant frequencies and displacements using the Ritz method can be pursued using subsets differentiated by their values of n . When these functions are substituted into the weak form of the equations of motion for isotropic cylinders, the sine and cosine functions cancel, leaving equations with only radial and axial variables and n as a parameter. For each value of n , there are a number of modes with different dependences on the radial and axial coordinates and generally different frequencies.

In addition to the value of n , So and Leissa used the reflection symmetry across the $z = 0$ plane to separate their approximation functions into subsets (corresponding to the character of the operator IC_{π} in the character table for $D_{\infty h}$, as presented by Cornwell¹⁷). After performing the calculations, they also characterized the singly degenerate modes according to the reflection symmetry across a plane that includes the cylinder axis (corresponding to the character of the operator $IC_{2\phi}$ in the character table). Modes that are unchanged by this reflection are axisymmetric, and those that are reversed in sign are torsional. The parity of reflection across these two planes, along with the values of n , serve to

completely characterize the symmetries of the modes of an isotropic cylinder (dividing them into irreducible representations).

So and Leissa studied several cases of isotropic cylinders with various dimensions and values of the Poisson ratio, providing numerous results for comparison. They claim an accuracy of four digits for values of the dimensionless frequency $\bar{\omega}$, which is defined to be $\omega R \sqrt{\rho/G}$, where G is the shear modulus.

Table III shows the dimensionless frequencies of the first 50 modes of a cylinder with $H/D=0.5$ and a Poisson ratio ν of 0.3 calculated using our formulations in cylindrical and rectangular coordinates (dimensionless frequencies $\bar{\omega}_c$ and $\bar{\omega}_r$, respectively). For the calculation in cylindrical coordinates, six terms in both the radial and axial coordinates and two terms in the angular coordinate have been employed in the approximation functions of Eq. (23). These results exactly match those of So and Leissa. This agreement is not surprising, since our model contains the approximations of Eq. (27) as a special case. For the calculation in rectangular coordinates, the approximation functions [Eq. (21)] include all terms with indices up to and including $i+j+k=10$.

The column labeled “mirror symmetry” in Table III gives the symmetry of reflection across the $z=0$ plane [symmetric (S) or antisymmetric (A)]. The second entry in this column (only appearing for the singly degenerate modes) gives the symmetry across a plane containing the cylinder axis (with S and A, in this case, corresponding to axisymmetric and torsional modes, respectively).

For each frequency, the column labeled “ D_{3d} subset” shows the label of the corresponding subset that was employed in the calculation in cylindrical coordinates, following the restrictions on the indices listed in Table II. Similarly, the column labeled “ C_{2h} ” shows the subsets used for the calculation in rectangular coordinates, following the restrictions on indices listed in Table I. The irreducible representations of the actual symmetry group of the isotropic cylinder, $D_{\infty h}$, are listed in the column labeled “ $D_{\infty h}$ subset.” These subsets were not employed in the calculations. They correspond to the classifications presented by So and Leissa (second and third columns).

The agreement between our results for the frequencies in cylindrical and rectangular coordinates ($\bar{\omega}_c$ and $\bar{\omega}_r$, respectively) is good, with the cylindrical formulation yielding slightly lower frequencies for the higher modes. Since the Ritz method converges to frequencies from above, these differences in values suggest higher accuracy for the calculation in cylindrical coordinates. Exact analytical solutions are known for the torsional modes listed in Table III (modes 12, 42, 56, and 64). The dispersion relation for these modes is¹⁹

$$\bar{\omega}^2 = \eta^2 + (\zeta R)^2, \quad (28)$$

where ζ is the wave number along the z axis, and η can have the values 0, 5.135 62, 8.417 24,... (roots of the Bessel function J_2). The modes with $\eta=0$ have undistorted cross sections (uniform twisting that varies with z), and those with higher values of η have angular displacements that vary with both r and, generally, z . ζ is an integral multiple of π/H (greater than or equal to 0). Therefore, for the cylinder con-

sidered in Table III (with $H/R=1$), the lowest values of $\bar{\omega}$ for the torsional modes are 3.1416, 5.1356, 6.0203, and 6.2832, in agreement with both sets of calculated values in the table.

The relative accuracy of the two approaches is shown in Tables IV and V, where the first four torsional modes are computed from the two formulations with various numbers of terms in the approximation functions. In the case of isotropic symmetry, the four original subsets used in the analysis in rectangular coordinates reduce to eight smaller subsets. The size of the eigenvalue problem still depends on the sum of the indices $i+j+k$. For example, when this sum is 5, the two modal subsets containing the torsional modes lead to 12 and 24 approximation functions. When the sum is 6, there are 30 and 24 terms (i.e., one of the subsets does not gain any additional terms). This continues up through $i+j+k=7$ (30 and 50 terms) and $i+j+k=8$ (60 and 50 terms). For the cylindrical formulation, the circumferential modes are orthogonal, and the problem effectively becomes two-dimensional in terms of power series in r and z . Using equal numbers of terms in r and z in the approximation leads to a total of 4 terms (2×2), 9 terms (3×3), and so on. It is clear from the results shown in Tables IV and V that the cylindrical formulation provides better accuracy for a given number of terms in the approximation.

B. Trigonal cylinders

The primary problem of interest is a finite cylinder with trigonal material symmetry. A representative cylinder is considered with its geometry defined by $H=6.880$ mm and $D=5.495$ mm. The density is 6.12 g/cm³, and the six independent elastic constants are $C_{11}=186.58$ GPa, $C_{33}=263.00$ GPa, $C_{13}=99.04$ GPa, $C_{12}=104.72$ GPa, $C_{14}=13.40$ GPa, and $C_{44}=50.77$ GPa. These constants are chosen to simulate the elastic properties of a material such as langatate²⁰ ($\text{La}_3\text{Ga}_{5.5}\text{Ta}_{0.5}\text{O}_{14}$), with the relatively small effects of piezoelectric and dielectric coupling on the vibrational frequencies neglected.

For the results presented in this section, each approximation function for displacement-field components in cylindrical coordinates includes 6 radial terms, 6 axial terms, and 2 circumferential terms. Inclusion of terms beyond this point has very little effect on most frequencies and leads to ill-conditioning of the eigenvalue problem. For the calculation in rectangular coordinates, indices up to $i+j+k=10$ are included.

The frequencies of this cylinder computed from the rectangular and cylindrical formulations are shown in Table VI. The entries in the last six columns give the relative change in the frequency arising from a change in each of the C_{ij} . For each row, these entries are scaled such that the sum of the derivatives is equal to 1. There are six modes with $\bar{\omega}=0$ that are not shown in Table VI. These frequencies correspond to rigid-body displacement or rotation with zero strain. For example, in rectangular coordinates, approximation functions with ($u_x=\text{constant}$, $u_y=0$, $u_z=0$) or ($u_x=z$, $u_y=0$, $u_z=-x$) are two such combinations. In cylindrical coordinates, the rigid-body modes include two frequencies each from subsets E_u and E_g , one from A_{2g} , corresponding to a rotation about the z axis, and one from A_{2u} , corresponding

TABLE III. Dimensionless frequencies, $\omega R \sqrt{\rho/G}$, for an isotropic cylinder with $H/D=0.5$ and $\nu=0.3$. Cylindrical formulation: $\bar{\omega}_c$ and D_{3d} subset. Rectangular formulation: $\bar{\omega}_r$ and C_{2h} subset. The entries in the third column indicate the symmetry (S =symmetric, A =antisymmetric) with respect to reflections; the first entry is for reflection across $z=0$, and the second entry, where it appears, is for reflection across a plane containing the cylinder axis.

Mode	n	Mirror symmetry	$D_{\infty h}$ subset	D_{3d} subset	$\bar{\omega}_c$	C_{2h} subset	$\bar{\omega}_r$
1,2	2	A	E_{2u}	E_u	1.591	A_u, B_u	1.591
3,4	2	S	E_{2g}	E_g	2.345	A_g, B_g	2.345
5	0	A, S	A_{2u}	A_{2u}	2.402	B_u	2.402
6,7	1	S	E_{1u}	E_u	2.705	A_u, B_u	2.705
8,9	3	A	E_{3g}	A_{1g}, A_{2g}	2.805	A_g, B_g	2.806
10,11	1	A	E_{1g}	E_g	3.088	A_g, B_g	3.088
12	0	A, A	A_{1u}	A_{1u}	3.142	A_u	3.142
13	0	S, S	A_{1g}	A_{1g}	3.238	A_g	3.238
14,15	3	S	E_{3u}	A_{1u}, A_{2u}	3.591	A_u, B_u	3.591
16,17	4	A	E_{4u}	E_u	3.945	A_u, B_u	3.946
18,19	2	S	E_{2g}	E_g	3.966	A_g, B_g	3.966
20,21	2	A	E_{2u}	E_u	4.040	A_u, B_u	4.040
22,23	1	A	E_{1g}	E_g	4.053	A_g, B_g	4.053
24,25	1	S	E_{1u}	E_u	4.595	A_u, B_u	4.595
26,27	1	A	E_{1g}	E_g	4.595	A_g, B_g	4.598
28,29	3	S	E_{3u}	A_{1u}, A_{2u}	4.612	A_u, B_u	4.612
30,31	4	S	E_{4g}	E_g	4.640	A_g, B_g	4.640
32	0	S, S	A_{1g}	A_{1g}	4.643	A_g	4.643
33	0	A, S	A_{2u}	A_{2u}	4.677	B_u	4.678
34,35	1	S	E_{1u}	E_u	4.836	A_u, B_u	4.836
36,37	2	S	E_{2g}	E_g	4.867	A_g, B_g	4.867
38,39	3	A	E_{3g}	A_{1g}, A_{2g}	4.962	A_g, B_g	4.966
40,41	5	A	E_{5g}	E_g	5.030	A_g, B_g	5.046
42	0	S, A	A_{2g}	A_{2g}	5.136	B_g	5.136
43,44	4	S	E_{4g}	E_g	5.230	A_g, B_g	5.231
45,46	2	A	E_{2u}	E_u	5.243	A_u, B_u	5.243
47,48	5	S	E_{5u}	E_u	5.571	A_u, B_u	5.572
49	0	S, S	A_{1g}	A_{1g}	5.617	A_g	5.618
50,51	2	S	E_{2g}	E_g	5.623	A_g, B_g	5.625
52,53	2	A	E_{2u}	E_u	5.850	A_u, B_u	5.858
54,55	4	A	E_{4u}	E_u	5.850	A_u, B_u	5.875
56	0	A, A	A_{1u}	A_{1u}	6.020	A_u	6.020
57,58	5	S	E_{5u}	E_u	6.034	A_u, B_u	6.035
59,60	3	S	E_{3u}	A_{1u}, A_{2u}	6.045	A_u, B_u	6.046
61,62	6	A	E_{6u}	A_{1u}, A_{2u}	6.072	A_u, B_u	6.104
63	0	A, S	A_{2u}	A_{2u}	6.145	B_u	6.156
64	0	S, A	A_{2g}	A_{2g}	6.283	B_g	6.283

to a translation along the z axis. Also shown in Table VI are the modal subsets associated with each of the formulations.

The results for the two formulations shown in Table VI are in good agreement (with differences less than a few parts

in 10^4 for the lower frequencies) and complement one another. The rectangular formulation provides a more natural perspective from which to consider the dependence of the frequencies on the individual elastic constants, since the C_{ij}

TABLE IV. Dimensionless frequencies $\bar{\omega}_r (= \omega R \sqrt{\rho/G})$ for an isotropic cylinder with $H/D=0.5$ and $\nu=0.3$ computed in the rectangular formulation for the first four torsional modes as a function of total number of terms in the approximation.

Number of terms	Mode 1	Mode 2	Mode 3	Mode 4
12	3.1425		7.7460	
24		5.2008		6.3059
30	3.1416		6.1137	
50		5.1368		6.2833
60	3.1416		6.0232	
Exact	3.1416	5.1356	6.0203	6.2832

TABLE V. Dimensionless frequencies $\bar{\omega}_c (= \omega R \sqrt{\rho/G})$ for an isotropic cylinder with $H/D=0.5$ and $\nu=0.3$ computed in the cylindrical formulation for the first four torsional modes as a function of total number of terms in the approximation.

Number of terms	Mode 1	Mode 2	Mode 3	Mode 4
4	3.1425	6.1237	6.8830	7.7456
9	3.1416	5.2003	6.0756	6.3059
16	3.1416	5.1466	6.0297	6.2833
25	3.1416	5.1360	6.0206	6.2832
36	3.1416	5.1356	6.0203	6.2832
Exact	3.1416	5.1356	6.0203	6.2832

TABLE VI. Frequencies and the relative dependence of the frequencies on the C_{ij} of a cylinder with the C_{ij} and ρ representative of langatate. $H = 6.880$ mm. $D = 5.495$ mm. Cylindrical formulation: $\bar{\omega}_c$ and D_{3d} subset. Rectangular formulation: $\bar{\omega}_r$ and C_{2h} subset.

Mode	$\omega_c/(2\pi)$ (MHz)	D_{3d} subset	$\omega_r/(2\pi)$ (MHz)	C_{2h} subset	$\partial\omega/\partial C_{11}$	$\partial\omega/\partial C_{33}$	$\partial\omega/\partial C_{13}$	$\partial\omega/\partial C_{12}$	$\partial\omega/\partial C_{14}$	$\partial\omega/\partial C_{44}$
1	0.1994	A_{1u}	0.1994	A_u	0.25	0.00	0.00	-0.14	-0.20	1.09
2,3	0.2744	E_u	0.2744	A_u, B_u	0.86	0.54	-0.29	-0.37	-0.20	0.46
4,5	0.2918	E_g	0.2918	A_g, B_g	1.18	0.03	-0.01	-0.66	-0.36	0.81
6,7	0.3139	E_u	0.3139	A_u, B_u	1.39	0.32	-0.17	-0.72	-0.04	0.22
8,9	0.3554	E_g	0.3554	A_g, B_g	1.61	0.28	-0.17	-0.83	-0.09	0.20
10	0.3844	A_{1g}	0.3844	A_g	0.42	1.27	-0.89	0.21	0.00	0.00
11,12	0.3901	E_g	0.3900	A_g, B_g	1.27	0.29	-0.14	-0.66	0.07	0.18
13,14	0.3923	E_u	0.3922	A_u, B_u	1.97	0.06	-0.08	-1.01	-0.30	0.36
15	0.3933	A_{2g}	0.3933	B_g	0.45	0.01	0.00	-0.25	-0.25	1.05
16,17	0.4494	E_g	0.4493	A_g, B_g	1.72	0.12	-0.14	-0.82	-0.15	0.27
18	0.4653	A_{2u}	0.4653	B_u	1.54	0.07	-0.10	-0.72	-0.34	0.55
19,20	0.4808	E_u	0.4808	A_u, B_u	1.12	0.21	-0.08	-0.58	-0.11	0.44
21	0.4827	A_{1g}	0.4827	A_g	1.81	0.15	-0.10	-0.95	-0.20	0.30
22	0.4956	A_{2g}	0.4956	B_g	1.62	0.07	-0.05	-0.88	-0.13	0.37
23	0.5041	A_{1u}	0.5041	A_u	1.84	0.22	-0.15	-0.97	-0.20	0.25
24	0.5147	A_{2u}	0.5147	B_u	1.75	0.31	-0.21	-0.86	-0.12	0.13
25,26	0.5274	E_u	0.5274	A_u, B_u	0.39	0.43	-0.28	-0.08	-0.02	0.56
27	0.5397	A_{1g}	0.5398	A_g	0.80	0.54	-0.52	0.10	0.01	0.08
28,29	0.5453	E_g	0.5458	A_g, B_g	0.25	0.28	-0.15	-0.06	-0.08	0.77
30	0.5468	A_{1u}	0.5467	A_u	2.14	0.05	-0.01	-1.20	-0.01	0.02
31	0.5523	A_{2u}	0.5523	B_u	1.22	0.45	-0.62	-0.06	0.00	0.01
32,33	0.5735	E_u	0.5734	A_u, B_u	1.23	0.12	-0.10	-0.60	-0.11	0.45
34	0.5928	A_{2u}	0.5928	B_u	0.89	0.54	-0.48	-0.13	0.05	0.13
35	0.6043	A_{2g}	0.6043	B_g	1.64	0.09	-0.02	-0.91	-0.17	0.38
36	0.6049	A_{1u}	0.6051	A_u	0.32	0.11	-0.07	-0.15	-0.15	0.94
37	0.6077	A_{1g}	0.6079	A_g	1.32	0.28	-0.17	-0.62	-0.12	0.31
38,39	0.6117	E_g	0.6120	A_g, B_g	1.31	0.14	-0.13	-0.62	-0.15	0.46
40,41	0.6144	E_u	0.6146	A_u, B_u	0.85	0.26	-0.19	-0.34	-0.04	0.47
42,43	0.6381	E_g	0.6384	A_g, B_g	1.54	0.10	-0.04	-0.81	-0.27	0.48
44,45	0.6505	E_u	0.6482	A_u, B_u	1.77	0.09	-0.07	-0.95	-0.19	0.35
46,47	0.6644	E_g	0.6636	A_g, B_g	1.73	0.14	-0.12	-0.89	-0.11	0.25
48	0.6724	A_{1g}	0.6726	A_g	0.74	0.30	-0.29	0.06	-0.03	0.22
49	0.6777	A_{2g}	0.6789	B_g	1.71	0.15	-0.08	-0.92	-0.29	0.44
50,51	0.6791	E_u	0.6792	A_u, B_u	1.43	0.09	-0.10	-0.70	-0.12	0.39
52	0.6921	A_{1u}	0.6923	A_u	1.56	0.13	-0.09	-0.82	-0.26	0.48
53,54	0.6942	E_g	0.6937	A_g, B_g	1.86	0.09	-0.04	-1.01	-0.09	0.19
55,56	0.7162	E_g	0.7172	A_g, B_g	0.83	0.34	-0.33	-0.22	-0.08	0.45
57,58	0.7235	E_u	0.7233	A_u	0.75	0.29	-0.19	-0.28	-0.08	0.51
59	0.7243	A_{2u}	0.7253	B_u	0.89	0.27	-0.22	-0.33	-0.09	0.47
60,61	0.7383	E_g	0.7395	A_g, B_g	1.39	0.22	-0.16	-0.66	-0.13	0.33
62	0.7470	A_{1g}	0.7487	A_g	0.69	0.15	-0.10	-0.21	-0.05	0.52
63,64	0.7710	E_u	0.7705	A_u, B_u	1.24	0.14	-0.05	-0.66	-0.02	0.35
65	0.7757	A_{1u}	0.7771	B_u	0.51	0.28	-0.26	-0.03	0.06	0.45
66	0.7764	A_{2u}	0.7779	A_u	1.66	0.11	-0.04	-0.92	-0.18	0.37
67,68	0.7826	E_g	0.7842	B_g	1.07	0.27	-0.21	-0.44	-0.12	0.43

are given in this coordinate system. The cylindrical formulation provides greater accuracy and allows more insight to be gained into the characteristics of the modal deformations. The displacement patterns for the lowest modes of each of the six D_{3d} subsets are shown in Fig. 1.

C. Transversely isotropic cylinders

An approach that is sometimes taken when C_{14} is small is to model the material as being transversely isotropic (hexagonal, with five independent elastic constants and $C_{14} = 0$). Such an approach was applied to sapphire by Lusher and

Hardy²¹ when extending Hutchinson's method for transversely isotropic cylinders. The effects of this assumption are investigated by computing the frequencies, $\omega/(2\pi)$, that result from such an approach for the present geometry and material. The results are shown in Table VII, where the grouping nomenclature of the rectangular formulation (C_{2h}) is employed. The frequencies are significantly different from the corresponding trigonal case. Several general observations can be made. First, the first seven modes obtained by the hexagonal approximation have frequencies that are higher than those computed with the full set of constants, and these changes affect the ordering of the modes. Second, the fre-

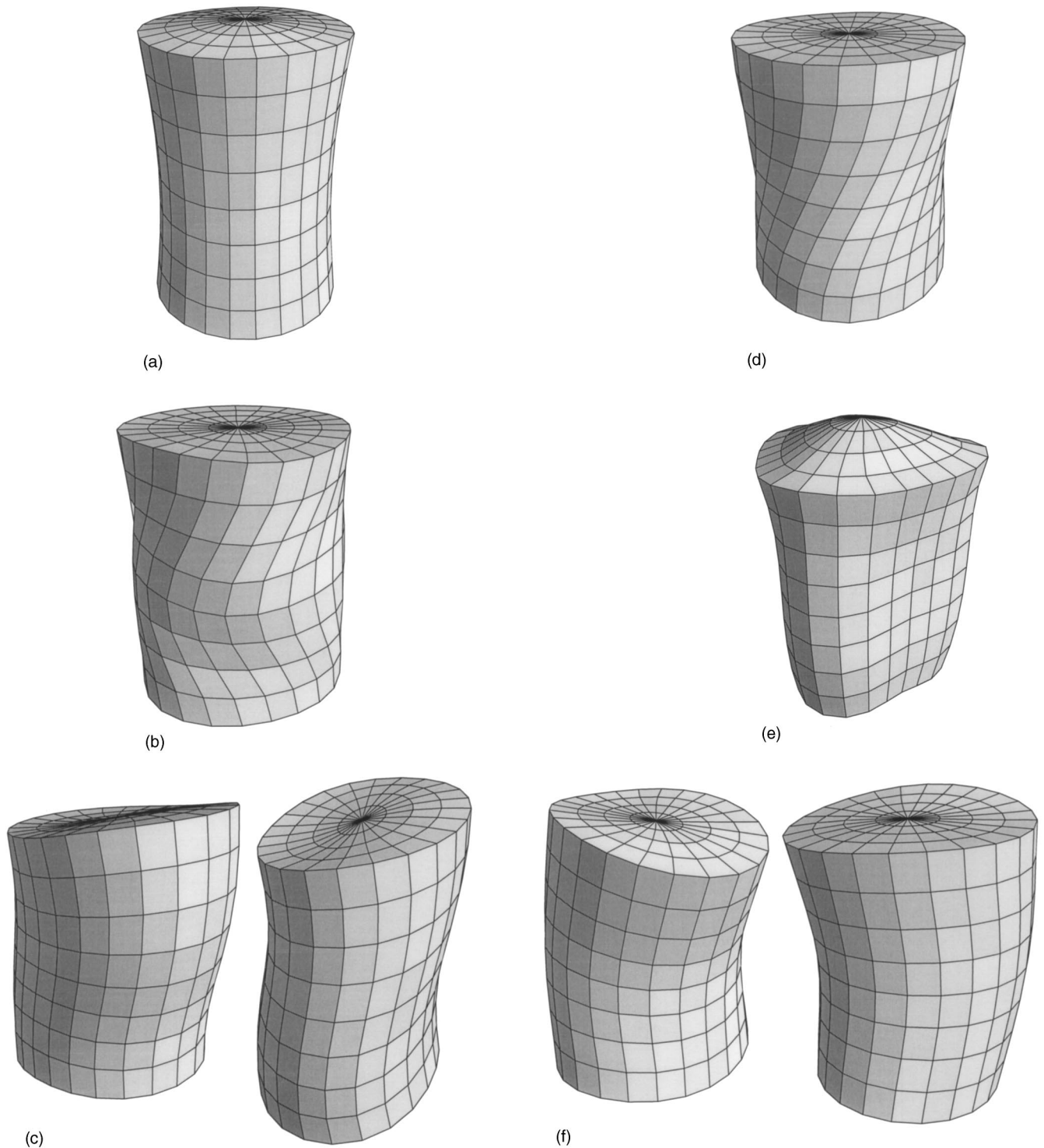


FIG. 1. Displacement patterns of the modes with the lowest frequencies from each of the six D_{3d} subsets: (a) A_{1g} , (b) A_{2g} , (c) E_g (doubly degenerate), (d) A_{1u} , (e) A_{2u} , and (f) E_u (doubly degenerate).

quencies for modes 1 and 13 depend only on the shear modulus C_{44} (the derivatives of ω with respect to all other elastic constants are zero), and ω_{13} is exactly twice ω_1 . In fact, there are an infinite number of modes that are multiples of the first frequency. These are purely torsional modes with the same dimensionless frequencies that are given by Eq. (28) for an isotropic cylinder. (This can be confirmed by calculating the dimensionless frequencies for modes 1 and 13 in

Table VII using the definition $\bar{\omega} \equiv \omega R \sqrt{\rho/G}$ with $G = C_{44}$ and the values of H , ρ , and $R (= D/2)$ given above for the langatate cylinder.) The presence of a nonzero C_{14} complicates the deformation and eliminates the purely torsional character of these modes. Although C_{14} is small compared with the on-axis constants in langatate, it is comparable in magnitude to C_{44} , and this results in a relatively large influence on the frequencies.

TABLE VII. Frequencies and the relative dependence of the frequencies on the C_{ij} of a cylinder with the C_{ij} and ρ representative of langatate, except that $C_{14}=0$. $H=6.880$ mm. $D=5.495$ mm.

Mode	$\omega/(2\pi)$ (MHz)	C_{2h} subset	$\partial\omega/\partial C_{11}$	$\partial\omega/\partial C_{33}$	$\partial\omega/\partial C_{13}$	$\partial\omega/\partial C_{12}$	$\partial\omega/\partial C_{14}$	$\partial\omega/\partial C_{44}$
1	0.2093	A_u	0.00	0.00	0.00	0.00	0.00	1.00
2,3	0.2875	A_u, B_u	0.18	0.85	-0.45	0.07	0.00	0.35
4,5	0.3171	A_u, B_u	1.86	0.02	-0.01	-1.04	0.00	0.17
6,7	0.3275	A_g, B_g	0.01	0.06	-0.02	0.00	0.00	0.95
8,9	0.3515	A_g, B_g	2.26	0.01	-0.01	-1.26	0.00	0.00
10,11	0.3844	A_g, B_g	1.37	0.56	-0.30	-0.65	0.00	0.03
12	0.3846	A_g	0.41	1.27	-0.89	0.21	0.00	0.00
13	0.4186	B_g	0.00	0.00	0.00	0.00	0.00	1.00

D. Effect of geometry

The ratio of the height to diameter of the specimen affects the frequencies and the sensitivity of the frequencies to each of the elastic constants. Table VIII shows calculations of the frequencies for a langatate cylinder with the height/diameter ratio increased to 5 (and with the material properties and D unchanged). An examination of the frequencies and derivatives with respect to the elastic constants shows that changing the length of the cylinder completely changes the ordering of the modes and the dependence of the frequencies on each of the elastic constants. For example, if these frequencies are measured for the original cylinder, 8 of the first 12 measurements correspond to doubly degenerate modes, whereas, for the longer cylinder, only 3 of the first 12 have this character. The modes that are most strongly dependent on C_{44} are closer to being purely torsional (the dependence on other elastic constants is reduced) for the longer cylinder. The strong dependence of the derivatives on the aspect ratio suggest that a parametric study of this effect could determine an optimal specimen aspect ratio of a cylinder for use in estimating elastic constants using resonant ultrasound spectroscopy or other techniques.

IV. SUMMARY

This study has considered the traction-free vibration of finite trigonal cylinders using Ritz algorithms in both cylindrical and rectangular Cartesian coordinates and employing group theory to reduce the matrix equations to block-diagonal form. The special case of an isotropic cylinder has been explored to compare the calculated frequencies ob-

tained from the two formulations and from the calculations of So and Leissa.⁸ Values of the frequencies calculated in cylindrical coordinates are in excellent agreement with those of So and Leissa. Those calculated in rectangular coordinates are slightly higher for some modes. An analysis of calculations for torsional modes, for which exact solutions are known, has shown that the formulation in cylindrical coordinates is more accurate than that in rectangular coordinates, for a given number of terms in the series approximation functions. For a cylinder of a representative trigonal crystal, langatate, approximate solutions for the frequencies and modal shapes have been determined. The relative dependence of the frequencies on each of the elastic constants of this material is also determined. Results for a different cylinder geometry show significant changes in ordering of the modes and dependencies on the elastic constants and, therefore, suggest that the aspect ratio of the specimen can be chosen to optimize the accuracy of the determination of elastic constants from measured resonant frequencies. The special case of transverse isotropy ($C_{14}=0$) also shows significant changes in frequencies, even though the trigonal material to which it is compared (langatate) has a C_{14} that is small relative to the other elastic constants. Purely torsional modes appear in the case of transverse isotropy.

The methods presented here have several advantages over previous Ritz formulations for trigonal materials. The cylindrical geometry has the advantage, over rectangular (parallelepiped) geometries, of eliminating uncertainties associated with the alignment of an additional crystallographic axis relative to the specimen faces. Therefore, it can provide greater accuracy in the determination of elastic constants.

TABLE VIII. Frequencies and the relative dependence of the frequencies on the C_{ij} of a cylinder with the C_{ij} and ρ representative of langatate. $H=27.47$ mm. $D=5.495$ mm. ($H/D=5$.)

Mode	$\omega/(2\pi)$ (MHz)	C_{2h} subset	$\partial\omega/\partial C_{11}$	$\partial\omega/\partial C_{33}$	$\partial\omega/\partial C_{13}$	$\partial\omega/\partial C_{12}$	$\partial\omega/\partial C_{14}$	$\partial\omega/\partial C_{44}$
1,2	0.1307	A_u, B_u	0.22	1.20	-0.61	0.09	-0.02	0.12
3	0.2000	A_u	0.22	0.00	0.00	-0.12	-0.19	1.09
4,5	0.3045	A_g, B_g	0.21	0.94	-0.47	0.04	-0.06	0.33
6	0.3998	B_g	0.22	0.00	0.00	-0.13	-0.19	1.09
7	0.4097	A_g	0.23	1.34	-0.70	0.13	0.00	0.00
8,9	0.5034	A_u, B_u	0.22	0.73	-0.35	0.00	-0.09	0.50
10	0.5991	A_u	0.23	0.00	0.00	-0.13	-0.20	1.09
11,12	0.7093	A_g	0.24	0.57	-0.27	-0.04	-0.12	0.63
13	0.7987	B_g	0.24	0.00	0.00	-0.14	-0.20	1.09
14	0.8116	B_u	0.25	1.34	-0.74	0.14	0.00	0.00
15,16	0.9105	A_u	0.29	0.45	-0.21	-0.09	-0.15	0.72
17	0.9995	A_u	0.26	0.00	0.00	-0.14	-0.21	1.09

The formulation of the problem in cylindrical coordinates provides a complete block-diagonalization of the matrix problem and a straightforward description of the symmetries of the vibrational displacements.

ACKNOWLEDGMENTS

This work was initiated when P.R.H. was an Alexander von Humboldt Research Fellow at the University of Stuttgart. The support of the Alexander von Humboldt Research Foundation and the hospitality of his co-workers at the Staatliche Materialprüfungsanstalt (Stuttgart) are gratefully acknowledged.

APPENDIX A: COEFFICIENT MATRICES

1. Rectangular Cartesian coordinates

$$K_{ij}^{11} = \int_V \left(C_{11} \frac{\partial \Psi_i^x}{\partial x} \frac{\partial \Psi_j^x}{\partial x} + C_{55} \frac{\partial \Psi_i^x}{\partial z} \frac{\partial \Psi_j^x}{\partial z} + C_{56} \frac{\partial \Psi_i^x}{\partial z} \frac{\partial \Psi_j^x}{\partial y} + C_{56} \frac{\partial \Psi_i^x}{\partial y} \frac{\partial \Psi_j^x}{\partial z} + C_{66} \frac{\partial \Psi_i^x}{\partial y} \frac{\partial \Psi_j^x}{\partial y} \right) dV,$$

$$K_{ij}^{12} = K_{ji}^{21} = \int_V \left(C_{12} \frac{\partial \Psi_i^x}{\partial x} \frac{\partial \Psi_j^y}{\partial y} + C_{14} \frac{\partial \Psi_i^x}{\partial x} \frac{\partial \Psi_j^y}{\partial z} + C_{56} \frac{\partial \Psi_i^x}{\partial z} \frac{\partial \Psi_j^y}{\partial x} + C_{66} \frac{\partial \Psi_i^x}{\partial y} \frac{\partial \Psi_j^y}{\partial x} \right) dV,$$

$$K_{ij}^{13} = K_{ji}^{31} = \int_V \left(C_{13} \frac{\partial \Psi_i^x}{\partial x} \frac{\partial \Psi_j^z}{\partial z} + C_{14} \frac{\partial \Psi_i^x}{\partial x} \frac{\partial \Psi_j^z}{\partial y} + C_{55} \frac{\partial \Psi_i^x}{\partial z} \frac{\partial \Psi_j^z}{\partial x} + C_{56} \frac{\partial \Psi_i^x}{\partial y} \frac{\partial \Psi_j^z}{\partial x} \right) dV,$$

$$K_{ij}^{22} = \int_V \left(C_{22} \frac{\partial \Psi_i^y}{\partial y} \frac{\partial \Psi_j^y}{\partial y} + C_{24} \frac{\partial \Psi_i^y}{\partial y} \frac{\partial \Psi_j^y}{\partial z} + C_{24} \frac{\partial \Psi_i^y}{\partial z} \frac{\partial \Psi_j^y}{\partial y} + C_{44} \frac{\partial \Psi_i^y}{\partial z} \frac{\partial \Psi_j^y}{\partial z} + C_{66} \frac{\partial \Psi_i^y}{\partial x} \frac{\partial \Psi_j^y}{\partial x} \right) dV,$$

$$K_{ij}^{23} = K_{ji}^{32} = \int_V \left(C_{23} \frac{\partial \Psi_i^y}{\partial y} \frac{\partial \Psi_j^z}{\partial z} + C_{24} \frac{\partial \Psi_i^y}{\partial y} \frac{\partial \Psi_j^z}{\partial y} + C_{44} \frac{\partial \Psi_i^y}{\partial z} \frac{\partial \Psi_j^z}{\partial y} + C_{56} \frac{\partial \Psi_i^y}{\partial x} \frac{\partial \Psi_j^z}{\partial x} \right) dV,$$

$$K_{ij}^{33} = \int_V \left(C_{33} \frac{\partial \Psi_i^z}{\partial z} \frac{\partial \Psi_j^z}{\partial z} + C_{44} \frac{\partial \Psi_i^z}{\partial y} \frac{\partial \Psi_j^z}{\partial y} + C_{55} \frac{\partial \Psi_i^z}{\partial x} \frac{\partial \Psi_j^z}{\partial x} \right) dV.$$

2. Cylindrical coordinates

$$K_{ij}^{11} = \int_V \left[C_{11} \frac{\partial \Psi_i^r}{\partial r} \frac{\partial \Psi_j^r}{\partial r} + \frac{C_{12}}{r} \Psi_i^r \frac{\partial \Psi_j^r}{\partial r} + C_{15} \frac{\partial \Psi_i^r}{\partial z} \frac{\partial \Psi_j^r}{\partial r} + \frac{C_{12}}{r} \frac{\partial \Psi_i^r}{\partial r} \Psi_j^r + \frac{C_{22}}{r^2} \Psi_i^r \Psi_j^r + \frac{C_{25}}{r} \frac{\partial \Psi_i^r}{\partial z} \Psi_j^r + C_{15} \frac{\partial \Psi_i^r}{\partial r} \frac{\partial \Psi_j^r}{\partial z} + \frac{C_{25}}{r} \Psi_i^r \frac{\partial \Psi_j^r}{\partial z} + C_{55} \frac{\partial \Psi_i^r}{\partial z} \frac{\partial \Psi_j^r}{\partial z} \frac{C_{56}}{r} + \frac{\partial \Psi_i^r}{\partial \theta} \frac{\partial \Psi_j^r}{\partial z} + \frac{C_{56}}{r} \frac{\partial \Psi_i^r}{\partial z} \frac{\partial \Psi_j^r}{\partial \theta} + \frac{C_{66}}{r^2} \frac{\partial \Psi_i^r}{\partial \theta} \frac{\partial \Psi_j^r}{\partial \theta} \right] r dr d\theta dz,$$

$$K_{ij}^{12} = \int_V \left[\frac{C_{12}}{r} \frac{\partial \Psi_i^r}{\partial r} \frac{\partial \Psi_j^\theta}{\partial \theta} + \frac{C_{22}}{r^2} \Psi_i^r \frac{\partial \Psi_j^\theta}{\partial \theta} + \frac{C_{25}}{r} \frac{\partial \Psi_i^r}{\partial z} \frac{\partial \Psi_j^\theta}{\partial \theta} + C_{14} \frac{\partial \Psi_i^r}{\partial r} \frac{\partial \Psi_j^\theta}{\partial z} + \frac{C_{24}}{r} \Psi_i^r \frac{\partial \Psi_j^\theta}{\partial z} + \frac{C_{46}}{r} \frac{\partial \Psi_i^r}{\partial \theta} \frac{\partial \Psi_j^\theta}{\partial z} + C_{56} \frac{\partial \Psi_i^r}{\partial z} \left(\frac{\partial \Psi_j^\theta}{\partial r} - \frac{\Psi_j^\theta}{r} \right) + \frac{C_{66}}{r} \frac{\partial \Psi_i^r}{\partial \theta} \right] \times \left(\frac{\partial \Psi_j^\theta}{\partial r} - \frac{\Psi_j^\theta}{r} \right) r dr d\theta dz,$$

$$K_{ij}^{13} = \int_V \left[C_{13} \frac{\partial \Psi_i^r}{\partial r} \frac{\partial \Psi_j^z}{\partial z} + \frac{C_{23}}{r} \Psi_i^r \frac{\partial \Psi_j^z}{\partial z} + \frac{C_{14}}{r} \frac{\partial \Psi_i^r}{\partial r} \frac{\partial \Psi_j^z}{\partial \theta} + \frac{C_{24}}{r^2} \Psi_i^r \frac{\partial \Psi_j^z}{\partial \theta} + \frac{C_{46}}{r^2} \frac{\partial \Psi_i^r}{\partial \theta} \frac{\partial \Psi_j^z}{\partial \theta} + C_{15} \frac{\partial \Psi_i^r}{\partial r} \frac{\partial \Psi_j^z}{\partial r} + \frac{C_{25}}{r} \Psi_i^r \frac{\partial \Psi_j^z}{\partial r} + C_{55} \frac{\partial \Psi_i^r}{\partial z} \frac{\partial \Psi_j^z}{\partial r} + \frac{C_{56}}{r} \frac{\partial \Psi_i^r}{\partial \theta} \frac{\partial \Psi_j^z}{\partial r} \right] r dr d\theta dz,$$

$$K_{ij}^{22} = \int_V \left[\frac{C_{22}}{r^2} \frac{\partial \Psi_i^\theta}{\partial \theta} \frac{\partial \Psi_j^\theta}{\partial \theta} + \frac{C_{24}}{r} \frac{\partial \Psi_i^\theta}{\partial z} \frac{\partial \Psi_j^\theta}{\partial \theta} + \frac{C_{24}}{r} \frac{\partial \Psi_i^\theta}{\partial \theta} \frac{\partial \Psi_j^\theta}{\partial z} + C_{46} \left(\frac{\partial \Psi_i^\theta}{\partial r} - \frac{\Psi_i^\theta}{r} \right) \frac{\partial \Psi_j^\theta}{\partial z} + C_{46} \frac{\partial \Psi_i^\theta}{\partial z} \left(\frac{\partial \Psi_j^\theta}{\partial r} - \frac{\Psi_j^\theta}{r} \right) + C_{66} \left(\frac{\partial \Psi_i^\theta}{\partial r} - \frac{\Psi_i^\theta}{r} \right) \left(\frac{\partial \Psi_j^\theta}{\partial r} - \frac{\Psi_j^\theta}{r} \right) \right] r dr d\theta dz,$$

$$K_{ij}^{23} = \int_V \left[\frac{C_{23}}{r} \frac{\partial \Psi_i^\theta}{\partial \theta} \frac{\partial \Psi_j^z}{\partial z} + \frac{C_{24}}{r^2} \frac{\partial \Psi_i^\theta}{\partial \theta} \frac{\partial \Psi_j^z}{\partial \theta} + \frac{C_{44}}{r} \frac{\partial \Psi_i^\theta}{\partial z} \frac{\partial \Psi_j^z}{\partial \theta} + \frac{C_{46}}{r} \left(\frac{\partial \Psi_i^\theta}{\partial r} - \frac{\Psi_i^\theta}{r} \right) \frac{\partial \Psi_j^z}{\partial \theta} + \frac{C_{25}}{r} \frac{\partial \Psi_i^\theta}{\partial \theta} \frac{\partial \Psi_j^z}{\partial r} + C_{56} \left(\frac{\partial \Psi_i^\theta}{\partial r} - \frac{\Psi_i^\theta}{r} \right) \frac{\partial \Psi_j^z}{\partial r} \right] r dr d\theta dz,$$

TABLE IX. Superimposed character tables of D_{3d} and C_{2h} . Numbers in boldface are characters common to both C_{2h} and D_{3d} . Other numbers are characters only for D_{3d} .

D_{3d}	C_{2h}	χ_1 χ'_1	χ_2	χ_3 χ'_2	χ_4 χ'_3	χ_5	χ_6 χ'_4
A_{1g}	A_g	1	1	1	1	1	1
A_{2g}	B_g	1	1	-1	1	1	-1
E_g		2	-1	0	2	-1	0
A_{1u}	A_u	1	1	1	-1	-1	-1
A_{2u}	B_u	1	1	-1	-1	-1	1
E_u		2	-1	0	-2	1	0

$$K_{ij}^{33} = \int_V \left[C_{33} \frac{\partial \Psi_i^z}{\partial z} \frac{\partial \Psi_j^z}{\partial z} + \frac{C_{44}}{r^2} \frac{\partial \Psi_i^z}{\partial \theta} \frac{\partial \Psi_j^z}{\partial \theta} + C_{55} \frac{\partial \Psi_i^z}{\partial r} \frac{\partial \Psi_j^z}{\partial r} \right] r dr d\theta dz.$$

APPENDIX B: SYMMETRY OPERATIONS OF D_{3d} AND C_{2h}

A specimen of a D_{3d} crystal in the form of a cylinder with the trigonal axis oriented along the cylinder axis has D_{3d} symmetry. That is, the cylindrical geometry maintains all the symmetry elements of the D_{3d} point group, including the identity operation, two threefold rotations, three twofold rotations, and the product of all these operations with inversion.

The group-theoretical characters of the elements of D_{3d} are shown superimposed with those of C_{2h} in Table IX. The classes of symmetry operations of D_{3d} (column labels) include the following elements.¹⁷

$$\chi_1 \equiv E,$$

$$\chi_2 \equiv C_{3z}, C_{3z}^{-1},$$

$$\chi_3 \equiv C_{2x}, C_{2y}, C_{2z},$$

$$\chi_4 \equiv I,$$

$$\chi_5 \equiv IC_{3z}, IC_{3z}^{-1},$$

$$\chi_6 \equiv IC_{2x}, IC_{2y}, IC_{2z},$$

where C_{3z} and C_{3z}^{-1} are, respectively, rotations by $2\pi/3$ and $-2\pi/3$ about the cylinder axis (threefold crystal axis), C_{2x} , C_{2y} , and C_{2z} are rotations by $\pi/2$ about twofold axes perpendicular to \hat{z} , and I is the inversion operation. $I (= IE)$, IC_{3z} , etc., are the same operations followed by inversion. The labels for the irreducible representations (row labels) follow the usual conventions for D_{3d} and C_{2h} .¹⁷ The valid characters for C_{2h} are written in boldface type in Table IX. The classes for this point group include only a fraction of the symmetry operations of D_{3d} :

$$\chi'_1 \equiv E, \quad \chi'_2 \equiv C_{2x}, \quad \chi'_3 \equiv I, \quad \chi'_4 \equiv IC_{2x}.$$

The two threefold rotations are not valid for C_{2h} , so these columns are absent, and there are correspondingly fewer irreducible representations (the doubly degenerate E_g and E_u are absent).

If approximating functions for the Ritz method are separated into subsets without considering their properties under

the threefold rotations (as they are above in the formulation for the trigonal cylinder in rectangular coordinates), then these subsets transform according to the irreducible representations of C_{2h} . If the threefold rotations are considered, then the approximating functions can be separated more completely according to the irreducible representations of D_{3d} . The correspondence of the C_{2h} and D_{3d} subsets can be understood by an inspection of the character tables. If the threefold rotations (χ_2 and χ_5) are ignored, then the characters of E_g are the sums of those of A_g and B_g . Therefore, E_g is a reducible representation in C_{2h} , and approximating functions that transform according to E_g cannot be separated from A_g or B_g functions on the basis of symmetry in this point group. Similarly, the characters of E_u are the sums of those of A_u and B_u , so that there are no symmetry-based criteria for separating E_u functions from A_u and B_u functions in C_{2h} .

- ¹L. Pochhammer, "Über die fortpflanzungsgeschwindigkeiten kleiner schwingungen in einem unbegrenzten isotropen kreiszylinder" ("On the propagation velocities of small vibrations in an infinite isotropic circular cylinder"), *J. Rein. Angew. Math.* **81**, 324–376 (1876).
- ²C. Chree, "Longitudinal waves of a solid bar," *Q. J. Math.* **21**, 287–298 (1886).
- ³C. Chree, "The equations of an isotropic elastic solid in polar and cylindrical coordinates, Their solutions and applications," *Trans. Cambridge Philos. Soc.* **14**, 250–369 (1889).
- ⁴J. R. Hutchinson, "Axisymmetric vibrations of a free finite-length rod," *J. Acoust. Soc. Am.* **51**, 233–240 (1972).
- ⁵J. R. Hutchinson, "Vibration of solid cylinders," *J. Appl. Mech.* **47**, 901–907 (1980).
- ⁶J. R. Hutchinson, "Transverse vibrations of beams: Exact versus approximate solutions," *J. Appl. Mech.* **48**, 923–928 (1981).
- ⁷W. M. Visscher, A. Migliori, T. M. Bell, and R. A. Reinert, "On the normal modes of free vibration of inhomogeneous and anisotropic elastic objects," *J. Acoust. Soc. Am.* **90**, 2154–2162 (1991).
- ⁸J. So and A. W. Leissa, "Three-dimensional vibrations of thick circular and annular plates," *J. Sound Vib.* **209**, 15–41 (1998).
- ⁹P. R. Heyliger, "Axisymmetric vibration of finite anisotropic cylinders," *J. Sound Vib.* **148**, 507–520 (1991).
- ¹⁰P. R. Heyliger and A. Jilani, "The free vibrations of inhomogeneous anisotropic cylinders and spheres," *Int. J. Solids Struct.* **29**, 2689–2708 (1992).
- ¹¹I. Ohno, S. Yamamoto, O. L. Anderson, and J. Noda, "Determination of elastic constants of trigonal crystals by the rectangular parallelepiped resonance method," *J. Phys. Chem. Solids* **47**, 1103–1108 (1986).
- ¹²E. Mochizuki, "Sphere-resonance method to determine elastic constants of crystal," *J. Appl. Phys.* **63**, 5668–5673 (1988).
- ¹³A. Migliori and J. L. Sarrao, *Resonant Ultrasound Spectroscopy: Applications to Physics, Materials Measurements, and Nondestructive Evaluation* (Wiley-Interscience, New York, 1997).
- ¹⁴H. F. Tiersten, *Linear Piezoelectric Plate Vibrations* (Plenum, New York, 1969).
- ¹⁵B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger, Malabar, FL, 1990).
- ¹⁶L. Meirovitch, *Computational Methods in Structural Dynamics* (Sijthoff and Noordhoff, Rockville, MD, 1980).
- ¹⁷J. F. Cornwell, *Group Theory in Physics* (Academic, New York, 1984), Vol. I.
- ¹⁸W. L. Johnson and P. R. Heyliger, "Symmetrization of Ritz approximation functions for vibrational analysis of trigonal cylinders," *J. Acoust. Soc. Am.* **113**, 1826–1832 (2003).
- ¹⁹A. C. Eringen and E. S. Şuhubi, *Elastodynamics* (Academic, New York, 1975), Vol. II.
- ²⁰H. Ledbetter and S. Kim (unpublished).
- ²¹C. P. Lusher and W. N. Hardy, "Axisymmetric free vibrations of a transversely isotropic finite cylindrical rod," *ASME J. Appl. Mech.* **55**, 855–862 (1988).

Symmetrization of Ritz approximation functions for vibrational analysis of trigonal cylinders

Ward L. Johnson

Materials Reliability Division, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305

Paul R. Heyliger

Department of Civil Engineering, Colorado State University, Fort Collins, Colorado 80523

(Received 22 March 2002; revised 3 January 2003; accepted 5 January 2003)

In the Ritz method of calculating vibrational normal modes, a set of finite series approximation functions provides a matrix eigenvalue equation for the coefficients in the series and the resonant frequency. The matrix problem usually can be block-diagonalized by grouping the functions into subsets according to their properties under the symmetry operations that are common to the specimen geometry and crystal class. This task is addressed, in this study, for the case of cylindrical specimens of crystals belonging to one of the higher trigonal crystal classes. The existence of doubly degenerate resonant modes significantly complicates the analysis. Group-theoretical projection operators are employed to extract, from series approximation functions in cylindrical coordinates, the terms that transform according to each irreducible representation of the point group. This provides a complete symmetry-based block diagonalization and categorization of the modal symmetries. Off-diagonal projection operators are used to provide relations between the displacement patterns of degenerate modes. The method of analysis is presented in detail to assist in its application to other geometries, crystal structures, and/or forms of Ritz approximation functions. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1558372]

PACS numbers: 43.20.Ks, 43.40.Cw [ANN]

I. INTRODUCTION

The Ritz method is a variational analytic technique for calculating the vibrational normal-mode displacements and frequencies from the elastic constants and geometry of an object. This method is frequently employed in iterative inversion algorithms to determine unknown elastic constants from measured frequencies. By inserting a finite number of approximation functions into Hamilton's principle, the method leads to a matrix eigenvalue equation involving variable parameters of the approximation functions. The matrix equation usually can be block diagonalized by grouping the approximation functions into orthogonal subsets according to their general symmetry properties. In this way, computer code for solving the eigenvalue problem can be made significantly more efficient. Such categorization of approximation functions also provides insight into common symmetry characteristics of the various vibrational modes and facilitates the prediction of relative coupling strengths for a specified distribution of excitation.

The task of symmetrizing Ritz approximation functions for a given macroscopic geometry, crystal symmetry, and crystal orientation involves an analysis of the transformations of these functions under the symmetry operations of the object (the operations that leave both the crystal structure and the geometry unchanged). Such symmetrizations have been performed for a number of geometries and crystal systems.

Demarest¹ presented a classification of the symmetries of approximation functions for cubes of orthorhombic material (belonging to the crystallographic point group that is

denoted by D_{2h} in the Schoenflies system and mmm in the international system) with the twofold crystal axes perpendicular to the cube faces. (For an overview of the crystallographic point groups, including representative graphical objects, see Ashcroft and Mermin.²) Demarest derived restrictions on the parities of the three Cartesian components of displacement with respect to reflections across the three perpendicular mirror planes and, in this way, divided approximation functions into eight orthogonal sets. These eight sets correspond to the eight irreducible representations of D_{2h} derived from group theory and, therefore, provide a block diagonalization of the Ritz matrices with minimal sizes of the submatrices for this symmetry. (For a brief summary of some concepts from group theory, see Appendix A.) Reducing the macroscopic symmetry (geometry) of the specimen to orthorhombic (a rectangular parallelepiped with unequal sides) maintains the D_{2h} symmetry. Therefore, the same classification scheme is valid for this geometry, and it was employed by Ohno.³ Ohno also presented a similar classification of approximation functions for monoclinic crystals (C_{2h} in the Schoenflies notation; $2/m$ in the international notation) in the form of rectangular parallelepipeds, which divided functions into two sets according to their parity with respect to the single mirror plane. Since this approach does not consider the effects of all the symmetry operations C_{2h} , it does not minimize the size of the submatrices. In a subsequent paper, Ohno *et al.*⁴ extended the analysis of C_{2h} specimens to include restrictions derived from the two-fold rotation of this point group and, in this way, separated the approximation functions into four sets and completed the block diago-

nalization. This latter work actually focused on parallelepipeds of trigonal material, but the overall symmetry is the same as that of a parallelepiped of monoclinic material, because the rectangular geometry eliminates the threefold trigonal symmetry axis.

The symmetrization of approximation functions for D_{2h} and C_{2h} specimens is relatively easy to accomplish, because all of the normal-mode displacements are either unchanged or reversed in sign by the symmetry operations of the object. In other words, all vibrational normal modes are nondegenerate (no two modes have the same frequency), so that the application of symmetry operations does not transform the displacement pattern of one mode into that of another mode. The normal modes remain nondegenerate for specimens with fewer symmetry operations. In terms of group theory, D_{2h} and lower-symmetry point groups (including C_{2h}) have only one-dimensional irreducible representations. The symmetrization of approximation functions for specimens belonging to the cubic, tetragonal, and hexagonal point groups is much more difficult. These point groups have irreducible representations with dimension greater than one, so that degenerate normal modes are transformed into one another under some of the symmetry operations.

Mochizuki⁵ provided a solution to the problem of block diagonalizing Ritz matrices for all crystal classes in spherical coordinates, except for the lower tetragonal and trigonal classes. He took advantage of the fact that the complete analytical solution for the normal modes of an isotropic sphere are known.^{6,7} Using these symmetrized solutions as approximation functions, Mochizuki derived restrictions on indices of the functions for various crystallographic point groups based on transformation properties of the elastic constants. Although this derivation was presented in the context of spherical specimens, the same functions and classification scheme could be used for any specimen with an overall symmetry (combination of geometry and material symmetry) equal to a crystal symmetry considered by Mochizuki, with the limits of volume integrals in the Ritz calculation adjusted to match the geometry.

Although Mochizuki's analysis provides a solution to the problem of block-diagonalizing Ritz matrices for most of the crystal classes, this solution, since it is expressed in terms of complicated functions of spherical coordinates, is unattractive for nonspherical geometries. In another publication, Mochizuki⁸ considered the symmetry restrictions on approximation functions in Cartesian coordinates. Using group-theoretical projection operators, he categorized approximation functions for right square prisms of tetragonal crystals and cubes of cubic crystals, in addition to parallelepipeds of lower-symmetry crystals.

Trigonal and hexagonal crystals have received the least attention in past work. The complete symmetrization of approximation functions for such crystals in geometries that do not eliminate symmetry operations has been accomplished previously only by Mochizuki⁵ in spherical coordinates.

In the present study, we address the problem of symmetrizing approximation functions, expressed in cylindrical coordinates, for vibrational modes of cylindrical crystals belonging to one of the higher-symmetry trigonal classes, C_{3v} ,

D_3 , or D_{3d} ($3m$, 32 , or $\bar{3}m$ in the international notation). This analysis has been used in the development of a Ritz algorithm for trigonal cylinders, which is described elsewhere.⁹ Several factors have determined the direction of this work. The focus on trigonal crystal classes arises from the important technological role that piezoelectric materials with this symmetry (including quartz) have in electronic oscillators. Currently, there is a particular need for temperature-dependent measurements of the elastic constants of several trigonal crystals with the structure of langasite, which may provide characteristics superior to quartz in some oscillator applications.^{10,11} The cylindrical geometry of specimens was chosen, instead of a rectangular geometry, partly because it can provide measurements of elastic constants with less uncertainty arising from crystal alignment (only one crystal axis needs to be aligned, instead of two). This geometry also has all the symmetry elements of the crystal, so that the overall symmetry of the specimen is the same as that of the crystal and the size of Ritz submatrices can be minimized. The problem is formulated in cylindrical coordinates, because the conditions on approximation functions arising from a consideration of the threefold rotations appear to be intractable in Cartesian coordinates. Cylindrical coordinates allow for a complete categorization of the symmetries of approximation functions and the associated minimization of submatrices. The complete description of modal symmetries was also pursued because of an interest in characterizing the resonant modes that are excited with a direct inductive piezoelectric transduction method recently employed with cylindrical specimens.¹²

Group-theoretical projection operators are employed to extract, from series approximation functions, the terms that transform according to each irreducible representation of the point group. This method is presented here in some detail. The aim is to illustrate the method completely, so that other researchers may be assisted in its application to other specimen geometries, crystal structures, and/or forms of Ritz approximation functions.

II. SYMMETRY OF D_{3d} CYLINDERS

The objective of this work is to symmetrize Ritz approximation functions for resonant modes of specimens with the symmetry of one of the higher trigonal crystal classes, C_{3v} , D_3 , or D_{3d} . However, only the D_{3d} class needs to be considered. Since linear elastic vibrations are insensitive to a lack of inversion, the inversion operator can be added to the C_{3v} and D_3 point groups, making them equivalent to the trigonal point group with the greatest number of symmetry operations, D_{3d} . This effective equivalence of C_{3v} , D_3 , and D_{3d} for linear elasticity can also be seen from the fact that the elastic constant matrices for these crystal classes, which contain all of the elastic symmetry information, have the same form.¹³

The symmetry of a D_{3d} crystal in the form of a cylinder with the trigonal axis oriented along the cylinder axis is shown in Fig. 1. The cylindrical geometry maintains all the elements of the D_{3d} point group, including rotations of $2\pi/3$ and $-2\pi/3$ (threefold rotations) about the vertical axis (\hat{z}),

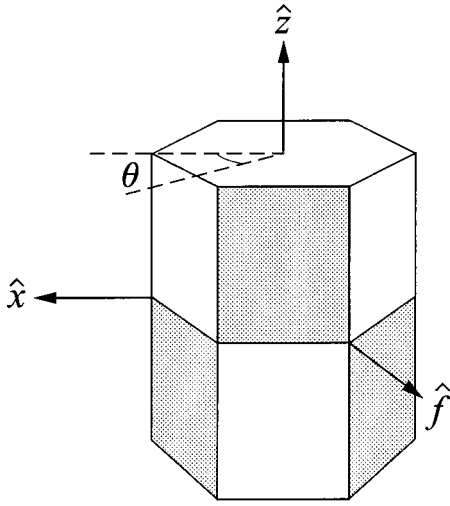


FIG. 1. Symmetry and definition of axes for an object belonging to the D_{3d} point group.

rotations of π (twofold rotations) about \hat{x} , \hat{f} , and \hat{d} (where \hat{d} , on the back side of the object, is in the same plane as \hat{x} and \hat{f} and halfway between them), inversion, and the products of inversion and the rotations. The labeling of the \hat{f} and \hat{d} axes corresponds to that of Cornwell's¹⁴ Appendix D. However, the axis label δ employed by Cornwell has been changed to \hat{z} , and \hat{b} has been changed to \hat{x} to match the conventions used by Heyliger and Johnson,⁹ which seem more appropriate for this geometry.

Corresponding to each of the symmetry operations T is a transformation operator $P(T)$ that acts on vector functions. These transformation operators, when applied to an arbitrary function, $\vec{u}(r, \theta, z) = u_r(r, \theta, z)\hat{r} + u_\theta(r, \theta, z)\hat{\theta} + u_z(r, \theta, z)\hat{z}$, yield the following functions:

$$P(E)\vec{u} = u_r(r, \theta, z)\hat{r} + u_\theta(r, \theta, z)\hat{\theta} + u_z(r, \theta, z)\hat{z}, \quad (1a)$$

$$P(C_{3z})\vec{u} = u_r\left(r, \theta + \frac{2\pi}{3}, z\right)\hat{r} + u_\theta\left(r, \theta + \frac{2\pi}{3}, z\right)\hat{\theta} + u_z\left(r, \theta + \frac{2\pi}{3}, z\right)\hat{z}, \quad (1b)$$

$$P(C_{3z}^{-1})\vec{u} = u_r\left(r, \theta - \frac{2\pi}{3}, z\right)\hat{r} + u_\theta\left(r, \theta - \frac{2\pi}{3}, z\right)\hat{\theta} + u_z\left(r, \theta - \frac{2\pi}{3}, z\right)\hat{z}, \quad (1c)$$

$$P(C_{2x})\vec{u} = u_r(r, -\theta, -z)\hat{r} - u_\theta(r, -\theta, -z)\hat{\theta} - u_z(r, -\theta, -z)\hat{z}, \quad (1d)$$

$$P(C_{2f})\vec{u} = u_r\left(r, -\theta + \frac{4\pi}{3}, -z\right)\hat{r} - u_\theta\left(r, -\theta + \frac{4\pi}{3}, -z\right)\hat{\theta} - u_z\left(r, -\theta + \frac{4\pi}{3}, -z\right)\hat{z}, \quad (1e)$$

$$P(C_{2d})\vec{u} = u_r\left(r, -\theta - \frac{4\pi}{3}, -z\right)\hat{r} - u_\theta\left(r, -\theta - \frac{4\pi}{3}, -z\right)\hat{\theta} - u_z\left(r, -\theta - \frac{4\pi}{3}, -z\right)\hat{z}, \quad (1f)$$

$$P(I)\vec{u} = u_r(r, \theta + \pi, -z)\hat{r} + u_\theta(r, \theta + \pi, -z)\hat{\theta} - u_z(r, \theta + \pi, -z)\hat{z}, \quad (1g)$$

where the unit vectors \hat{r} and $\hat{\theta}$ are in the radial and azimuthal directions at a given point. E is the identity operation. C_{3z} and C_{3z}^{-1} are rotations about \hat{z} by $2\pi/3$ and $-2\pi/3$, respectively. C_{2x} , C_{2f} , and C_{2d} are twofold rotations about \hat{x} , \hat{f} , and \hat{d} , respectively. I is the inversion operation. The effects of the additional transformation operators involving inversion (IC_{3z} , etc.) can be derived easily from the relation¹⁴

$$P(IT) = P(I)P(T). \quad (2)$$

III. RESTRICTIONS ON INDICES OF APPROXIMATION FUNCTIONS

Following Heyliger and Johnson,⁹ the components of a normal-mode displacement $\vec{u}(r, \theta, z)$ of a D_{3d} cylinder are approximated in cylindrical coordinates by finite series of the form

$$u_r(r, \theta, z) = \sum_{\alpha, n, \beta} a_{\alpha n \beta}^c r^\alpha \cos(n\theta) z^\beta + \sum_{\kappa, m, \gamma} a_{\kappa m \gamma}^s r^\kappa \sin(m\theta) z^\gamma, \quad (3a)$$

$$u_\theta(r, \theta, z) = \sum_{\alpha, n, \beta} b_{\alpha n \beta}^c r^\alpha \cos(n\theta) z^\beta + \sum_{\kappa, m, \gamma} b_{\kappa m \gamma}^s r^\kappa \sin(m\theta) z^\gamma, \quad (3b)$$

$$u_z(r, \theta, z) = \sum_{\alpha, n, \beta} c_{\alpha n \beta}^c r^\alpha \cos(n\theta) z^\beta + \sum_{\kappa, m, \gamma} c_{\kappa m \gamma}^s r^\kappa \sin(m\theta) z^\gamma, \quad (3c)$$

where n , m , α , β , κ , and γ are non-negative integers. For a given material density and set of elastic constants, the coefficients $a_{\alpha n \beta}^c$, $a_{\kappa m \gamma}^s$, $b_{\alpha n \beta}^c$, $b_{\kappa m \gamma}^s$, $c_{\alpha n \beta}^c$, and $c_{\kappa m \gamma}^s$ for each vibrational mode can be determined through Ritz analysis by inserting the approximation functions and the specimen geometry into a variational equation derived from Hamilton's principle.⁹

According to group theory, the displacement field \vec{u} of each of the normal modes is a basis function for one of the irreducible representations of D_{3d} , which are conventionally labeled A_{1g} , A_{2g} , A_{1u} , A_{2u} , E_g , and E_u .¹⁴ (See Appendixes A and B.) Barring accidental degeneracies, modes transforming as one of the first four of these irreducible representations are nondegenerate, and those transforming as E_g or E_u are doubly degenerate. Since functions transforming according to different irreducible representations are or-

thogonal (even if they are not solutions of the elastic eigenvalue problem), solutions of the vibrational problem can be sought using a limited set of approximation functions of the form given by Eq. (3) with the indices in the functions restricted to values that provide the symmetry associated with a given irreducible representation. In other words, submatrices can be considered separately in the Ritz calculation, if approximation functions are grouped in this way.

The restrictions on the indices of approximation functions are determined here by applying group-theoretical projection operators, which are defined as¹⁴

$$\mathcal{P}_{ij}^p \equiv \frac{d_p}{g} \sum_{T \in G} \Gamma^p(T)_{ij}^* P(T), \quad (4)$$

where $\Gamma^p(T)$ is the matrix for a symmetry operation T in an irreducible representation of dimension d_p of a group G that contains g symmetry operations (elements). (See Appendix A.) The summation is over all elements in the group, and the asterisk denotes the complex conjugate. For D_{3d} , g is equal to 12, the $P(T)$ are given by Eq. (1), p assumes six arbitrarily assigned values corresponding to the six irreducible representations, and d_p is either 1 (for the nondegenerate representations) or 2 (for the doubly degenerate representations).¹⁴ The matrices for D_{3d} are given in Appendix B. For this point group, all of the matrices are real, so that $\Gamma^p(T)_{ij}^* = \Gamma^p(T)_{ij}$.

A diagonal projection operator \mathcal{P}_{ii}^p with a given value of i , when applied to an arbitrary function, extracts the part of the function that transforms as the i th row of the representation Γ^p . When applied to series approximation functions such as those given by Eq. (3), inspection of the extracted functions leads to a determination of restrictions on the indices for functions belonging to a given irreducible representation. For the doubly degenerate representations, the projection operators with $i \neq j$ can be used to determine the relationship between coefficients in the expansions of two degenerate modes.

The application of a projection operator to a series approximation \vec{u} with the form of Eq. (3) yields a function \vec{u}' :

$$\mathcal{P}_{ij}^p \vec{u} = \vec{u}' = u'_r(r, \theta, z) \hat{r} + u'_\theta(r, \theta, z) \hat{\theta} + u'_z(r, \theta, z) \hat{z}, \quad (5)$$

with

$$\begin{aligned} u'_r(r, \theta, z) = & \frac{d_p}{g} \sum_{\alpha, n, \beta} [A_{n\beta}(z) a_{\alpha n \beta}^c r^\alpha \cos(n\theta) \\ & + B_{n\beta}(z) a_{\alpha n \beta}^c r^\alpha \sin(n\theta)] \\ & + \frac{d_p}{g} \sum_{\kappa, m, \gamma} [C_{m\gamma}(z) a_{\kappa m \gamma}^s r^\kappa \cos(m\theta) \\ & + D_{m\gamma}(z) a_{\kappa m \gamma}^s r^\kappa \sin(m\theta)], \end{aligned} \quad (6a)$$

$$\begin{aligned} u'_\theta(r, \theta, z) = & \frac{d_p}{g} \sum_{\alpha, n, \beta} [A_{n\beta}(z) b_{\alpha n \beta}^c r^\alpha \cos(n\theta) \\ & + B_{n\beta}(z) b_{\alpha n \beta}^c r^\alpha \sin(n\theta)] \\ & + \frac{d_p}{g} \sum_{\kappa, m, \gamma} [C_{m\gamma}(z) b_{\kappa m \gamma}^s r^\kappa \cos(m\theta) \\ & + D_{m\gamma}(z) b_{\kappa m \gamma}^s r^\kappa \sin(m\theta)], \end{aligned} \quad (6b)$$

$$\begin{aligned} u'_z(r, \theta, z) = & \frac{d_p}{g} \sum_{\alpha, n, \beta} [A_{n\beta}(z) c_{\alpha n \beta}^c r^\alpha \cos(n\theta) \\ & + B_{n\beta}(z) c_{\alpha n \beta}^c r^\alpha \sin(n\theta)] \\ & + \frac{d_p}{g} \sum_{\kappa, m, \gamma} [C_{m\gamma}(z) c_{\kappa m \gamma}^s r^\kappa \cos(m\theta) \\ & + D_{m\gamma}(z) c_{\kappa m \gamma}^s r^\kappa \sin(m\theta)], \end{aligned} \quad (6c)$$

where

$$\begin{aligned} A_{n\beta}(z) = & z^\beta + [\Gamma_{ij}(C_{3z})_{ij} + \Gamma(C_{3z}^{-1})_{ij}] \cos(2n\pi/3) z^\beta + \Delta_1 \Gamma(C_{2x})_{ij} (-z)^\beta + \Delta_1 [\Gamma(C_{2f})_{ij} + \Gamma(C_{2d})_{ij}] \cos(4n\pi/3) (-z)^\beta \\ & + \Delta_2 \Gamma(I)_{ij} \cos(n\pi) (-z)^\beta + \Delta_2 [\Gamma(IC_{3z})_{ij} + \Gamma(IC_{3z}^{-1})_{ij}] \cos(n\pi) \cos(2n\pi/3) (-z)^\beta + \Delta_1 \Delta_2 \Gamma(IC_{2x})_{ij} \cos(n\pi) z^\beta \\ & + \Delta_1 \Delta_2 [\Gamma(IC_{2f})_{ij} + \Gamma(IC_{2d})_{ij}] \cos(n\pi) \cos(4n\pi/3) z^\beta, \end{aligned} \quad (7a)$$

$$\begin{aligned} B_{n\beta}(z) = & [-\Gamma(C_{3z})_{ij} + \Gamma(C_{3z}^{-1})_{ij}] \sin(2n\pi/3) z^\beta + \Delta_1 [\Gamma(C_{2f})_{ij} - \Gamma(C_{2d})_{ij}] \sin(4n\pi/3) (-z)^\beta + \Delta_2 [-\Gamma(IC_{3z})_{ij} \\ & + \Gamma(IC_{3z}^{-1})_{ij}] \cos(n\pi) \sin(2n\pi/3) (-z)^\beta + \Delta_1 \Delta_2 [\Gamma(IC_{2f})_{ij} - \Gamma(IC_{2d})_{ij}] \cos(n\pi) \sin(4n\pi/3) z^\beta, \end{aligned} \quad (7b)$$

$$\begin{aligned} C_{m\gamma}(z) = & [\Gamma(C_{3z})_{ij} - \Gamma(C_{3z}^{-1})_{ij}] \sin(2m\pi/3) z^\gamma + \Delta_1 [\Gamma(C_{2f})_{ij} - \Gamma(C_{2d})_{ij}] \sin(4m\pi/3) (-z)^\gamma + \Delta_2 [\Gamma(IC_{3z})_{ij} \\ & - \Gamma(IC_{3z}^{-1})_{ij}] \cos(m\pi) \sin(2m\pi/3) (-z)^\gamma + \Delta_1 \Delta_2 [\Gamma(IC_{2f})_{ij} - \Gamma(IC_{2d})_{ij}] \cos(m\pi) \sin(4m\pi/3) z^\gamma, \end{aligned} \quad (7c)$$

$$\begin{aligned} D_{m\gamma}(z) = & z^\gamma + [\Gamma(C_{3z})_{ij} + \Gamma(C_{3z}^{-1})_{ij}] \cos(2m\pi/3) z^\gamma - \Delta_1 \Gamma(C_{2x})_{ij} (-z)^\gamma - \Delta_1 [\Gamma(C_{2f})_{ij} + \Gamma(C_{2d})_{ij}] \cos(4m\pi/3) (-z)^\gamma \\ & + \Delta_2 \Gamma(I)_{ij} \cos(m\pi) (-z)^\gamma + \Delta_2 [\Gamma(IC_{3z})_{ij} + \Gamma(IC_{3z}^{-1})_{ij}] \cos(m\pi) \cos(2m\pi/3) (-z)^\gamma - \Delta_1 \Delta_2 \Gamma(IC_{2x})_{ij} \cos(m\pi) z^\gamma \\ & - \Delta_1 \Delta_2 [\Gamma(IC_{2f})_{ij} + \Gamma(IC_{2d})_{ij}] \cos(m\pi) \cos(4m\pi/3) z^\gamma, \end{aligned} \quad (7d)$$

$$\Delta_1 \equiv \begin{cases} 1 & \text{for } u'_r, \\ -1 & \text{for } u'_\theta \text{ and } u'_z, \end{cases} \quad (8a)$$

$$\Delta_2 \equiv \begin{cases} 1 & \text{for } u'_r \text{ and } u'_\theta, \\ -1 & \text{for } u'_z. \end{cases} \quad (8b)$$

To illustrate the way restrictions on indices of the approximation functions are determined, consider the relatively simple case of vibrational normal modes that transform according to the irreducible representation A_{1g} . The values of the Γ_{11} (characters) of this one-dimensional representation are all equal to 1 (Table II), which corresponds to basis functions being unchanged by any of the symmetry operations of D_{3d} . With these values, Eq. (7) reduces to

$$A_{n\beta}(z) = [z^\beta + \Delta_1(-z)^\beta][1 + 2 \cos(2n\pi/3)] \\ + \Delta_2[\Delta_1 z^\beta + (-z)^\beta] \cos(n\pi) \\ \times [1 + 2 \cos(2n\pi/3)], \quad (9a)$$

$$B_{n\beta}(z) = 0, \quad (9b)$$

$$C_{m\gamma}(z) = 0, \quad (9c)$$

$$D_{m\gamma}(z) = [z^\gamma - \Delta_1(-z)^\gamma][1 + 2 \cos(2m\pi/3)] \\ - \Delta_2[\Delta_1 z^\gamma - (-z)^\gamma] \cos(m\pi) \\ \times [1 + 2 \cos(2m\pi/3)], \quad (9d)$$

using the identity $\cos(4n\pi/3) = \cos(2n\pi/3)$. For u'_r , Δ_1 and Δ_2 are equal to 1, and Eqs. (9a) and (9d) become

$$A_{n\beta}(z) = [z^\beta + (-z)^\beta][1 + 2 \cos(2n\pi/3) + \cos(n\pi) \\ + 2 \cos(n\pi) \cos(2n\pi/3)], \quad (10a)$$

$$D_{m\gamma}(z) = [z^\gamma - (-z)^\gamma][1 + 2 \cos(2m\pi/3) - \cos(m\pi) \\ - 2 \cos(m\pi) \cos(2m\pi/3)]. \quad (10b)$$

Inspection of Eq. (10a) reveals that $A_{n\beta}(z)$ is zero unless n and β are even and n is a multiple of 3. Similarly, $D_{m\gamma}(z)$ is zero unless m and γ are odd and m is a multiple of 3. Thus, only approximation functions with these indices for the radial component need to be considered when searching for solutions of the vibrational problem that transform as A_{1g} .

The restrictions on indices for the azimuthal and axial components of A_{1g} functions are obtained by inserting the corresponding values of Δ_1 and Δ_2 [Eq. (8)] into Eq. (9). Similarly, restrictions on each of the components of functions transforming according to the other irreducible representations of D_{3d} are obtained by inserting the corresponding values of Γ_{11} , Δ_1 , and Δ_2 into Eq. (7). The results of this analysis are summarized in Table I. For the doubly degenerate irreducible representations, E_g and E_u , the restrictions on the indices for the second mode, obtained by inserting Γ_{22} into Eq. (7), are also included in this table. The first and second modes of E_g and E_u are labeled "1" and "2" in the second column of the table, corresponding to the row definitions implicit in the Γ matrices, Eq. (B2).

The expressions for $B_{n\beta}(z)$ and $C_{m\gamma}(z)$ [Eqs. (7b) and (7c)] provide explicit information about the relationship be-

TABLE I. Parities of indices of approximation functions [Eq. (3)] for each of the irreducible representations (i.r.) of D_{3d} . ($E \equiv$ even, $O \equiv$ odd.) The values of n and m also must be multiples of 3 for A_{1g} , A_{2g} , A_{1u} , A_{2u} and must not be multiples of 3 for E_g and E_u . The second column lists the row indices of the doubly degenerate irreducible representations [Eq. (B2)].

i.r.	Row	Component	n	β	m	γ
A_{1g}		u_r	E	E	O	O
		u_θ	O	O	E	E
		u_z	E	O	O	E
A_{2g}		u_r	O	O	E	E
		u_θ	E	E	O	O
		u_z	O	E	E	O
E_g	1	u_r	O	O	E	E
		u_θ	E	E	O	O
		u_z	O	E	E	O
E_g	2	u_r	E	E	O	O
		u_θ	O	O	E	E
		u_z	E	O	O	E
A_{1u}		u_r	O	E	E	O
		u_θ	E	O	O	E
		u_z	O	O	E	E
A_{2u}		u_r	E	O	O	E
		u_θ	O	E	E	O
		u_z	E	E	O	O
E_u	1	u_r	E	O	O	E
		u_θ	O	E	E	O
		u_z	E	E	O	O
E_u	2	u_r	O	E	E	O
		u_θ	E	O	O	E
		u_z	O	O	E	E

tween coefficients in the expansions for the degenerate modes that transform as E_g or E_u . Once the displacement field \vec{u} for a mode transforming according to the first row of E_g or E_u is found, the displacement field \vec{u}' of the second mode of the degenerate pair is given by Eqs. (5)–(8) with $i = 2$ and $j = 1$. For example, inserting the values for Γ_{21} of E_g [Eq. (B2)] into Eq. (7),

$$A_{n\beta}(z) = 0, \quad (11a)$$

$$B_{n\beta}(z) = -\sqrt{3}[z^\beta - \Delta_1(-z)^\beta] \sin(2n\pi/3) + \sqrt{3}\Delta_2[\Delta_1 z^\beta \\ - (-z)^\beta] \cos(n\pi) \sin(2n\pi/3), \quad (11b)$$

$$C_{m\gamma}(z) = \sqrt{3}[z^\gamma + \Delta_1(-z)^\gamma] \sin(2m\pi/3) + \sqrt{3}\Delta_2[\Delta_1 z^\gamma \\ + (-z)^\gamma] \cos(m\pi) \sin(2m\pi/3), \quad (11c)$$

$$D_{m\gamma}(z) = 0. \quad (11d)$$

These equations reduce to the same expressions for all three of the components, u'_r , u'_θ , and u'_z :

$$B_{n\beta}(z) = -4\sqrt{3}z^\beta \sin(2n\pi/3) \\ = \begin{cases} -6z^\beta & \text{if } n = 1 + 3h, \\ 6z^\beta & \text{if } n = 2 + 3h, \end{cases} \quad (12a)$$

$$C_{m\gamma}(z) = 4\sqrt{3}z^\gamma \sin(2m\pi/3) = \begin{cases} 6z^\gamma & \text{if } m = 1 + 3k, \\ -6z^\gamma & \text{if } m = 2 + 3k, \end{cases} \quad (12b)$$

where h and k are non-negative integers, and n and m are still restricted to the parities indicated in Table I. Equation (12) also applies to the transformation between degenerate functions of E_u . Note that the prefactor d_p/g in the general expressions for u'_r , u'_θ , and u'_r [Eq. (6)] is equal to $1/6$ for E_g and E_u . Therefore, the expansion for the second function (\vec{u}') in a degenerate pair is equal to that for the first function (\vec{u}) with sine functions replacing cosine functions, cosine functions replacing sine functions, and coefficients multiplied by 1 or -1 [depending on the values of n or m , according to Eq. (12)].

IV. CONCLUSION

Group-theoretical projection operators provide a straightforward, although somewhat cumbersome, approach for symmetrizing Ritz approximation functions for cylinders of crystals belonging to one of the higher trigonal classes. For nondegenerate vibrational modes, this approach actually does not need to be employed, since restrictions on the functions can be derived by considering, in turn, each symmetry operation and the corresponding entries in the character table. However, the doubly degenerate modes introduce a complexity that has been addressed here by employing the full power of group theory through the projection operators. The final results for the specific set of series approximation functions given by Eq. (3) are simple restrictions on the indices of the series coefficients (Table I) that divide the functions into subsets corresponding to the irreducible representations of D_{3d} . These results are used by Heyliger and Johnson⁹ to formulate a Ritz algorithm for calculating the modal displacements and frequencies of trigonal cylinders.

APPENDIX A: GROUP-THEORETICAL TERMINOLOGY

A brief summary of several relevant concepts from group theory is presented here for the benefit of readers who have little familiarity with this theory. Readers interested in complete mathematically rigorous definitions of terms should refer to the cited literature.

A crystallographic point group G is the set of all symmetry operations (elements), not including translations, that leave a crystal structure unchanged. Depending on the crystal structure, these operations may include rotations about an axis, inversion through a fixed point, reflections across a plane, rotation reflections, and/or rotation inversions.²

A matrix representation Γ of G is a set of nonsingular square matrices with the properties that 1), for every element T_i of G , there is a corresponding matrix $\Gamma(T_i)$ and 2), for every pair of elements T_i and T_j , matrix multiplication corresponds to successive application of symmetry operations:¹⁴

$$\Gamma(T_i T_j) = \Gamma(T_i) \Gamma(T_j). \quad (\text{A1})$$

All of the matrices in a representation have the same number of rows, and this number is the dimension d of the representation. The character of each matrix is defined to be the trace.

A set of functions $\psi_1, \psi_2, \dots, \psi_d$ is a basis for a representation of G if

TABLE II. Character table for D_{3d} .

	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6
A_{1g}	1	1	1	1	1	1
A_{2g}	1	1	-1	1	1	-1
E_g	2	-1	0	2	-1	0
A_{1u}	1	1	1	-1	-1	-1
A_{2u}	1	1	-1	-1	-1	1
E_u	2	-1	0	-2	1	0

$$P(T) \psi_n(\vec{r}) = \sum_{m=1}^d \Gamma(T)_{mn} \psi_m(\vec{r}), \quad (\text{A2})$$

where $P(T)$ is an operator that transforms the coordinates of $\psi_n(\vec{r})$ according to the symmetry operation T . The function ψ_n is said to transform as the n th row of G .

There are an infinite number of representations of each crystallographic point group. However, almost all of these are reducible, which means that all the matrices in such a representation can be simultaneously block diagonalized through the application of an appropriate similarity transformation. For each crystallographic point group, there are only a few submatrices that appear in any completely block diagonalized representation (apart from similarity transformations of the submatrices). These submatrices are the irreducible representations of the point group.

For a given point group, basis functions which belong to different irreducible representations or different rows of the same irreducible representation are orthogonal. This group-theoretical result is central to the current paper, because the sorting of series approximation functions according to their irreducible representations (and rows, for two-dimensional representations) leads to a block diagonalization of the Ritz matrices.

APPENDIX B: IRREDUCIBLE REPRESENTATIONS OF D_{3d}

There are six irreducible representations of the crystallographic point group D_{3d} : four one-dimensional representations, which normally are labeled A_{1g} , A_{2g} , A_{1u} , A_{2u} , and two two-dimensional representations labeled E_g and E_u .¹⁴ The subscripts “ g ” and “ u ” indicate that the corresponding basis functions are even and odd, respectively, under inversion.

Abbreviated information on the symmetries of basis functions for the irreducible representations is given by the characters, which are presented in Table II. The column labels in this table designate the classes, which are defined to include the following elements:¹⁴

$$\chi_1 \equiv E, \quad (\text{B1a})$$

$$\chi_2 \equiv C_{3z}, C_{3z}^{-1}, \quad (\text{B1b})$$

$$\chi_3 \equiv C_{2x}, C_{2y}, C_{2z}, \quad (\text{B1c})$$

$$\chi_4 \equiv I, \quad (\text{B1d})$$

$$\chi_5 \equiv IC_{3z}, IC_{3z}^{-1}, \quad (\text{B1e})$$

$$\chi_6 \equiv IC_{2x}, IC_{2y}, IC_{2z}. \quad (\text{B1f})$$

The effect of each symmetry operation on basis functions of the one-dimensional representations can be read directly from the table: the functions are either unchanged or reversed in sign by an operation, depending on whether the corresponding character is 1 or -1 , respectively. For these representations, the single entry in the matrix $\Gamma(T)$ for each symmetry operation T is simply the character of that T .

The matrices for E_g , as presented by Cornwell,¹⁴ are

$$\Gamma(E) = \Gamma(I) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (\text{B2a})$$

$$\Gamma(C_{3z}) = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}, \quad (\text{B2b})$$

$$\Gamma(C_{3z}^{-1}) = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}, \quad (\text{B2c})$$

$$\Gamma(C_{2x}) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (\text{B2d})$$

$$\Gamma(C_{2f}) = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}, \quad (\text{B2e})$$

$$\Gamma(C_{2d}) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}. \quad (\text{B2f})$$

The additional matrices involving inversion [$\Gamma(IC_{3z})$, $\Gamma(IC_{3z}^{-1})$, etc.] are the same as the corresponding matrices above that do not involve inversion. The matrices for E_u are the same, except that those involving inversion have each matrix element reversed in sign.

- ¹H. H. Demarest, Jr., "Cube-resonance method to determine the elastic constants of solids," *J. Acoust. Soc. Am.* **49**, 768–775 (1971).
- ²N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart, and Winston, New York, 1976).
- ³I. Ohno, "Free vibration of a rectangular parallelepiped crystal and its application to determination of elastic constants of orthorhombic crystals," *J. Phys. Earth* **24**, 355–379 (1976).
- ⁴I. Ohno, Y. Shigeru, O. L. Anderson, and J. Noda, "Determination of elastic constants of trigonal crystals by the rectangular parallelepiped resonance method," *Phys. Chem. Solids* **47**, 1103–1108 (1986).
- ⁵E. Mochizuki, "Sphere-resonance method to determine elastic constants of crystal," *J. Appl. Phys.* **63**, 5668–5673 (1988).
- ⁶Y. Satô and T. Usami, "Basic study of the oscillations of a homogeneous elastic sphere, II. Distribution of Displacement," *Geophys. Mag.* **31**, 25–47 (1962).
- ⁷Y. Satô and T. Usami, "Basic study of the oscillations of a homogeneous elastic sphere, III. Boundary conditions and the generation of elastic waves," *Geophys. Mag.* **31**, 49–62 (1962).
- ⁸E. Mochizuki, "Application of group theory to free oscillations of an anisotropic rectangular parallelepiped," *J. Phys. Earth* **35**, 159–170 (1987).
- ⁹P. R. Heyliger and W. L. Johnson, "Traction-free vibrations of finite trigonal elastic cylinders," *J. Acoust. Soc. Am.* **113**, 1812–1825 (2003).
- ¹⁰D. C. Malocha, M. P. da Cunha, E. Adler, R. C. Smythe, S. Frederick, M. Chou, R. Helmbold, and Y. S. Zhou, "Recent measurements of material constants versus temperature for langatate, langanite, and langasite," *Proceedings of the 2000 IEEE/EIA International Frequency Control Symposium and Exhibition*, Kansas City (IEEE, Danvers, MA, 2000), pp. 200–205.
- ¹¹J. A. Kosinski and R. A. Pastore, Jr., "A review of langasite material constants from BAW and SAW data: toward an improved data set," *Proceedings of the 2001 IEEE International Frequency Control Symposium and PDA Exhibition*, Seattle (IEEE, Danvers, MA, 2000), pp. 278–286.
- ¹²W. Johnson, S. Kim, and D. Lauria, "Anelastic loss in langatate," in *Proceedings of the 2000 IEEE/EIA International Frequency Control Symposium and Exhibition*, Kansas City (IEEE, Danvers, MA, 2000), pp. 186–190.
- ¹³B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger, Malabar, FL, 1990), Vol. I.
- ¹⁴J. F. Cornwell, *Group Theory in Physics* (Academic, New York, 1984), Vol. I.

Investigation of the vibrational modes of edge-constrained fibrous samples placed in a standing wave tube

Bryan H. Song^{a)} and J. Stuart Bolton^{b)}

1077 Ray W. Herrick Laboratories, School of Mechanical Engineering, Purdue University, West Lafayette, Indiana 47907-1077

(Received 5 December 2001; revised 4 November 2002; accepted 23 December 2002)

In earlier work it was suggested that the frictional constraint of a porous sample around its circumference in a standing wave tube resulted in shearing resonances of the sample. In the present work that effect has been confirmed by direct measurement of the spatial distribution of the velocity of the solid phase of a fibrous sample placed in a rigidly terminated standing wave tube and driven into motion by a plane, incident sound field. The measurements were performed using a standing wave tube to which a transparent downstream section was attached. A laser Doppler velocimeter was then used to measure the velocity of the solid phase of acoustically driven samples. The materials considered here were two types of aviation-grade glass fiber. A poroelastic finite element model was used to simulate the response of the constrained fibrous samples. Good agreement between measured and predicted mode shapes was found both when the samples were constrained only around their edges, and when an additional constraint plane was inserted axially through the samples. The present results confirm that glass fiber samples placed in a standing wave tube exhibit shearing modes and that those modes are associated with previously observed transmission loss minima. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1548155]

PACS numbers: 43.20.Ks, 43.20.Mv, 43.55.Ev, 43.58.Bh [DEC]

I. INTRODUCTION

In earlier articles it was shown that a fibrous sample placed in an anechoically terminated standing wave tube exhibits a transmission loss minimum^{1,2} and that at frequencies below that minimum the transmission loss increases to a finite, low frequency limit proportional to the sample's flow resistance.² It was suggested that the transmission loss minimum resulted from a shearing resonance of the sample in which the sample mass reacted against the stiffness induced by the frictional constraint of the sample around its edges. The latter suggestion was supported by the results of finite element (FE) simulations of the acoustical behavior of edge-constrained fibrous samples.² In particular, it was shown that the anechoic transmission loss, reflection coefficient and surface normal impedance of fibrous samples placed in an anechoically terminated standing wave tube could be reproduced in detail by using a poroelastic finite element model to represent the edge-constrained samples. A parameter variation study conducted as part of that work showed that, all other parameters being held constant, the frequency of the transmission loss minimum was controlled by the shear stiffness of the edge-constrained fibrous sample, i.e., the Young's modulus and Poisson's ratio were found to have essentially no independent effect on the transmission loss so long as the value of the shear modulus was held constant. The latter observation is consistent with the suggestion that the sample responds in an almost purely shearing motion at the transmission loss minimum. That is, at that frequency the motion

of the solid phase of the porous material is primarily axial, it having zero velocity where the sample contacts the tube wall and the velocity being a maximum at the center of the sample.

Although the close agreement between measured and predicted results makes it reasonable to assume that the shearing mechanism operates in practice, no direct evidence has been presented to show that an edge-constrained sample responds in the manner suggested. Thus the intention of the work described in this paper was to visualize the vibratory motion of the solid phase of constrained glass fiber samples when driven by a plane wave field and so to verify the existence of the shearing motion that was previously inferred indirectly by matching acoustical measurements and finite element predictions.

In the present work, the frame motion of glass fiber materials placed in a standing wave tube and backed by a rigidly terminated airspace was measured directly by using a laser vibrometer. It will be shown that the measured spatial variation of the frame velocity is consistent with the shearing motion described above and that the measured results agree closely with corresponding finite element predictions.

II. BACKGROUND

Many papers related to the measurement of the surface normal velocity of vibrating panels or plates by the use of laser vibrometers have appeared in the literature: see Refs. 3 and 4 for two recent examples. However, only a very few papers have been published in which direct measurement of the vibratory motion of a porous material's solid phase is described.^{5,6}

In the first of these papers, Dubbelday describes a procedure for estimating the Poisson's ratio of foamed alumi-

^{a)}Now at NASA, Expendable Launch Vehicles, Mail Code VB-A3, Kennedy Space Center, FL 32899.

^{b)}Author to whom correspondence should be addressed. Electronic mail: bolton@purdue.edu

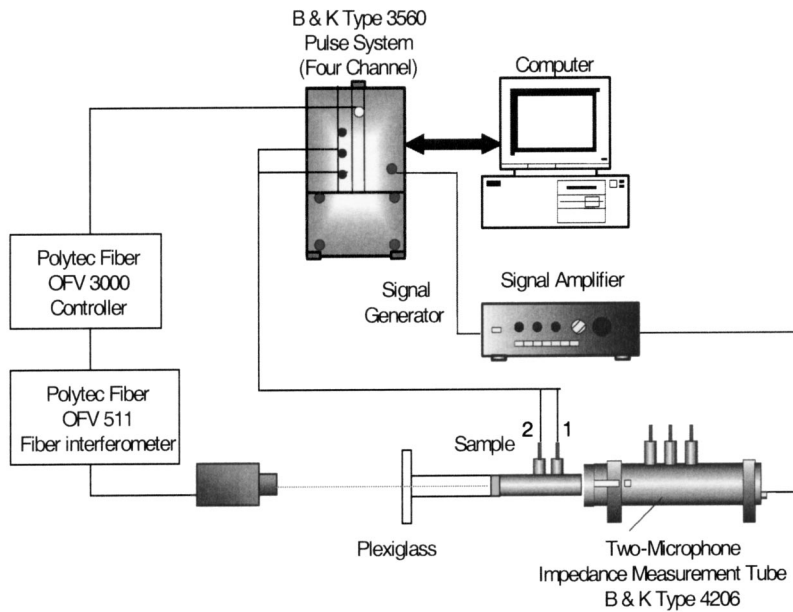


FIG. 1. Schematic of the laser measurement setup: the high frequency tube is shown here.

num samples.⁵ In his procedure, two laser beams were used to measure the lateral motion at single points on two sides of a foam sample that was forced into axial vibration. It was concluded that the foam's Young's modulus and Poisson's ratio could be estimated simultaneously in this way. However, as noted, the laser vibration measurements were made at only two fixed points rather than over an extended surface.

Later, Mariez *et al.* described ways of identifying the mechanical properties of polyurethane foams.⁶ In their procedure, a cube of polyurethane foam was mounted between rigid plates and was then prestressed. When the sample was excited by a shaker attached to one of the plates, the sample's lateral extension was measured at a single point on the side of the sample by using a laser vibrometer. At the same time, the input mechanical impedance of the system was measured by using a piezoelectric sensor mounted between the shaker and the upper plate. Finally, the transfer function between the lower plate's displacement and the laser vibrometer signal was estimated. Mariez *et al.* then used a finite element procedure to predict the mechanical impedance and transfer function with respect to a range of values of the sample's Young's modulus and Poisson's ratio and estimated the latter properties by matching measured and predicted re-

sults. In this way the mechanical properties of open cell foams were determined at low frequencies (1 Hz to 100 Hz).

More recently, Khirnykh and Cummings described an electromagnetic approach for estimating the space-averaged, vibrational velocity of the frame of an elastic porous material placed in a standing wave tube.⁷ In their measurements, thin conducting wires were inserted transversely into a porous sample. The measurement was then based on the fact that an electromotive force is associated with the integral of the cross-product of the wire's velocity vector and a magnetic field vector imposed along the length of the wire. When it is assumed that the amplitude of the magnetic field vector is constant over the length of the wire, the average vibrational velocity along the length of the wire can be calculated. It was assumed that the wire moved locally with the same velocity as the frame of the porous material and thus that the wire's average velocity represented the average of the frame velocity across the width of the sample. Since the wires that were inserted into the porous materials were relatively thin (their diameters were 40 μm and 140 μm) and light, it was assumed that the mass loading of the sample by the wires was negligible. In addition, it was assumed that the mechanical properties of the porous material (i.e., its stiffness) and hence

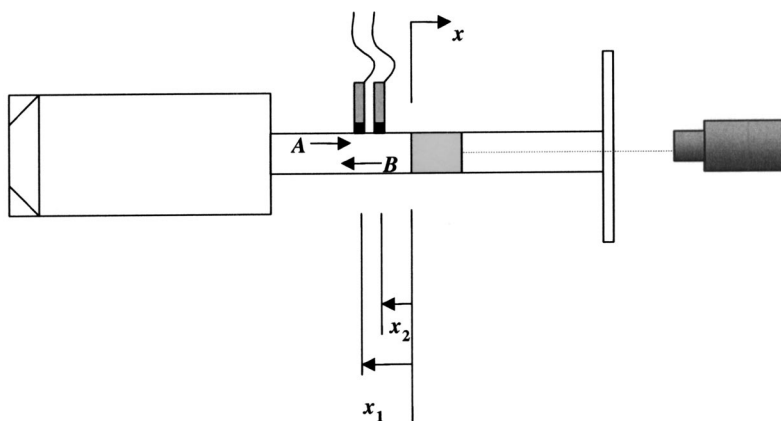


FIG. 2. Schematic of the two-microphone standing wave tube setup.

TABLE I. Poroelastic material properties used in FE calculations.

Material	Bulk density (Kg/m ³)	Porosity	Tortuosity	Flow resistivity (MKS Rayls/m)	Shear modulus (Pa)	Loss factor
Yellow	6.7	0.99	1.1	21000	1200	0.350
Green	9.6	0.99	1.1	31000	2800	0.275

the motion of the sample were not affected by a wire's presence. It was suggested that this procedure offered a reliable way of estimating the cross-sectionally averaged vibrational velocity of the solid skeleton of a porous sample placed in a standing wave tube; however no theoretical results were ad-

duced to verify the accuracy of the experimental results. Finally, note that the procedure, at least in the form described, does not allow the spatial distribution of the frame velocity to be visualized.

In the present work, a laser vibrometer was used to measure the surface normal velocity of the rear surface of glass fiber materials placed in a standing wave tube: i.e., the measurements were made directly on a point-by-point basis along a diametral line across the rear surface of a sample backed by an airspace and a rigid, transparent termination. A finite element code (COMET/Safe ver. 4.0) was used to calculate theoretical results to which the corresponding measured surface normal velocities were compared. The experimental apparatus, materials and procedures used in the present work are described in the next section. A brief description of the finite element models used here is presented in Sec. IV followed by a discussion of the measured and predicted results in Sec. V.

III. EXPERIMENTAL APPARATUS, MATERIALS AND PROCEDURES

A. Experimental apparatus

A schematic diagram of the experimental arrangement used to measure a sample's frame vibration is shown in Fig. 1. The measurement was based on the use of a Brüel and Kjær Two-Microphone Impedance Measurement Tube (Type 4206). Measurements were made in both the low frequency (50 Hz to 1600 Hz) and high frequency (100 Hz to 6400 Hz) ranges by using the large and small diameter tubes. The samples were placed at the downstream end of the standing wave tube, and then large and small diameter transparent, circular Plexiglas tubes were used as appropriate to terminate the downstream section so that a laser vibrometer could be used to measure the surface normal velocity on the rear surface of the samples. Note that a hard termination condition was adopted for these tests since the acoustical effect of an open termination is difficult to predict by using FE methods owing to the complexity of the acoustic loading that would be experienced by the downstream surface of the porous layer in that case.

The Plexiglas sections were handmade to fit tightly onto the end of the impedance tubes in order to avoid leakage and the tube wall thickness was made relatively large (both large and small diameter tubes had a wall thickness of 6 mm) so that the sample holder would not exhibit ovaling modes in the frequency range of interest. The tube termination was sealed tightly with a flat piece of plexiglass, also 6 mm thick. The inner diameter of the plexiglass tubes was made the same as the Brüel and Kjær impedance tube (i.e., 9.9 cm for the large tube and 2.9 cm for the small tube). In

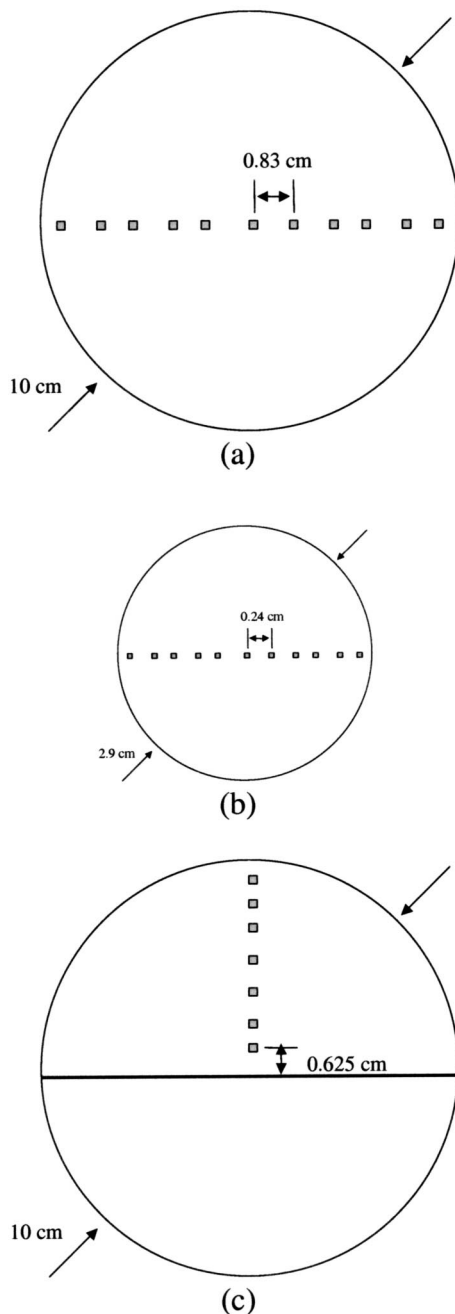


FIG. 3. Schematic of glass fiber samples with reflecting tape: (a) large sample, (b) small sample, and (c) large sample (plane-constraint case).

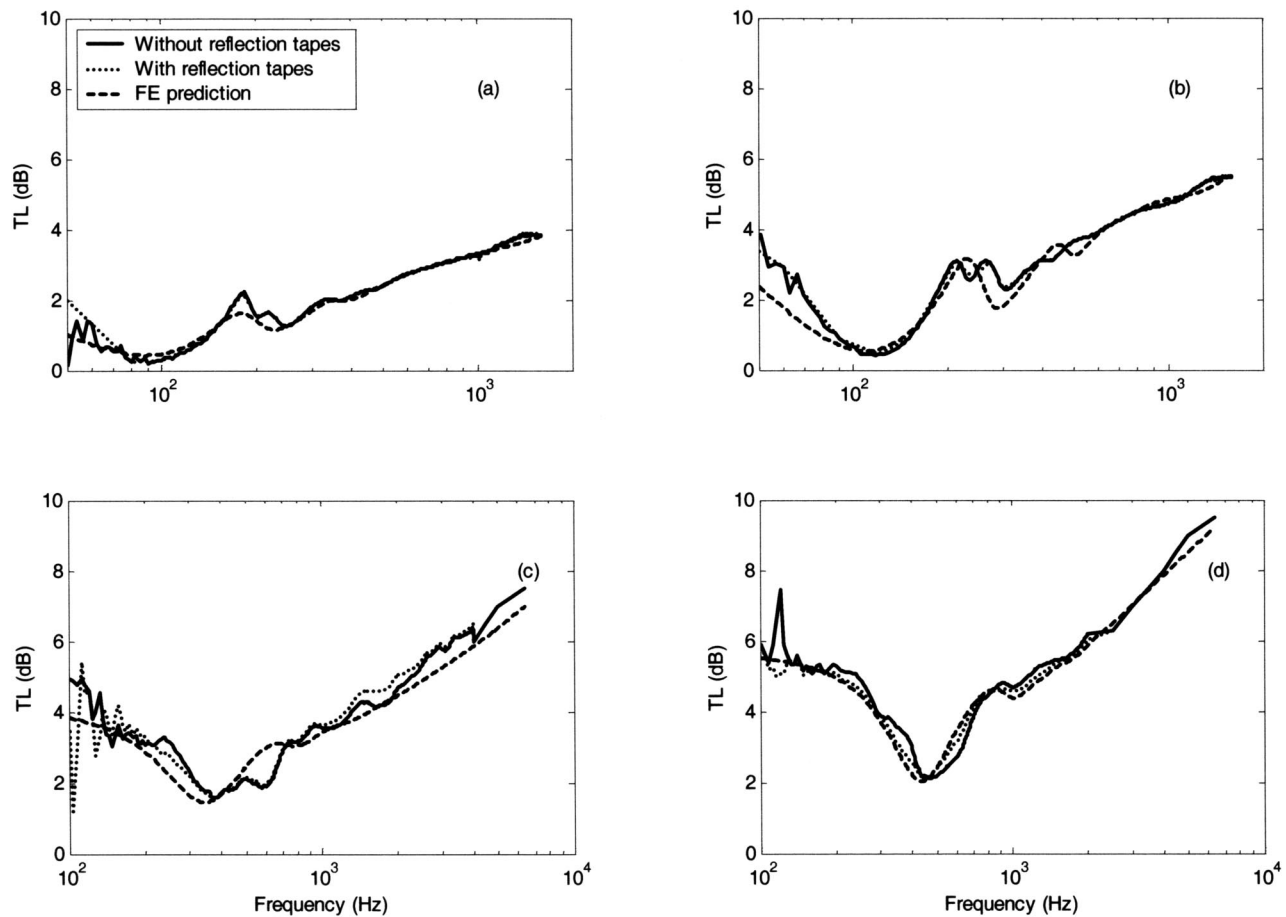


FIG. 4. Effect of reflecting tape on the transmission loss of 2.5 cm thick glass fiber samples: (a) yellow sample in the large tube, (b) green sample in the large tube, (c) yellow sample in the small tube, (d) green sample in the small tube.

the case of the large tube, the distance from the microphone at position 1 to the front surface of the glass fiber sample (in all cases 2.5 cm thick) was 35.0 cm. In the small tube case, the distance from the microphone at position 1 to the front surface of the glass fiber sample was 18.8 cm. The rigid termination of the Plexiglas tube was 9.7 cm and 10.0 cm behind the rear surface of the glass fiber samples in the large and small tube cases, respectively.

A Polytec laser vibrometer was used to measure the surface normal velocity as shown in Fig. 1. Its tracking filter was adjusted to “slow,” the velocity range was chosen to be HF 25 mm/s/V and the velocity filter was set to 5 kHz. The laser beam was re-focused every time it was moved from one location to another and the distance between the laser and the plexiglass termination was kept constant (3 cm) in all cases.

B. Two-microphone method

The reflection coefficient of each sample was measured using a two-microphone procedure at the same time that the laser measurements were performed. The two-microphone method is based on measuring the transfer function between microphones located at x_1 and x_2 : see Fig. 2. The complex sound pressure at those points can be expressed as

$$H_1 = [Ae^{-jkx_1} + Be^{jkx_1}], \quad (1a)$$

$$H_2 = [Ae^{-jkx_2} + Be^{jkx_2}], \quad (1b)$$

where A and B are the complex amplitudes of the positive- and negative-going plane wave components upstream of the sample, k is the complex wave number in the ambient fluid, and an $e^{+j\omega t}$ sign convention has been assumed. The transfer function between the two microphones is then denoted $H_{21} = H_2/H_1$. When Eqs. (1) are substituted into the latter expression it is possible to express the plane wave reflection coefficient, R , in terms of the transfer function, H_{21} , i.e.,

$$R = \frac{H_{21}e^{-ikx_1} - e^{-ikx_2}}{e^{ikx_2} - H_{21}e^{ikx_1}}. \quad (2)$$

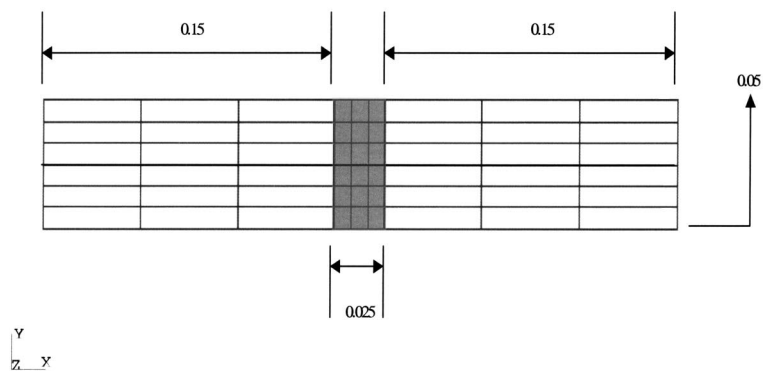
Then the normalized surface normal impedance of the sample can be calculated as

$$\frac{Z_n}{\rho_0 c} = \frac{1 + R}{1 - R}. \quad (3)$$

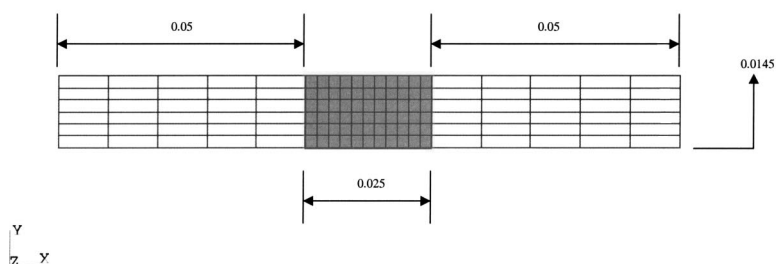
The complex wave number was calculated using a formula presented by Temkin (Eq. 6.6.9 in Ref. 8).

C. Materials tested and sample preparation

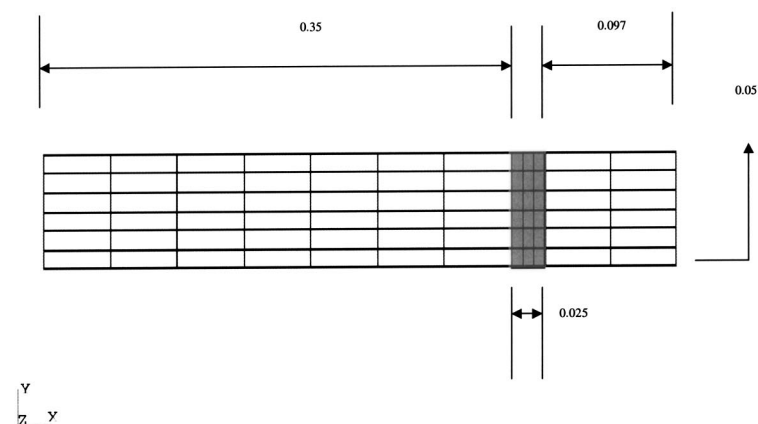
The glass fiber materials used in these measurements were the same as those described in earlier articles,^{1,2} i.e., yellow and green, aviation-grade glass fiber having the properties listed in Table I. The samples were cut to fit snugly into the impedance tube to avoid the creation of air-gaps



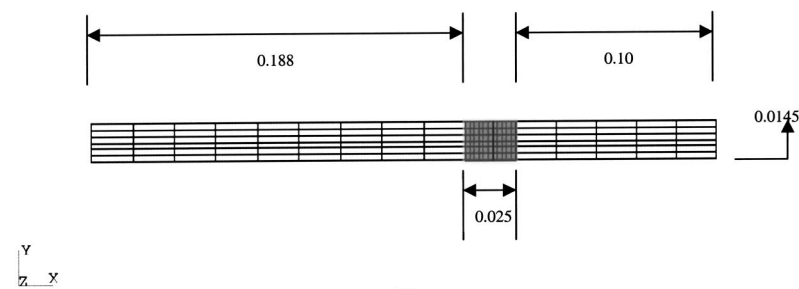
(a)



(b)



(c)



(d)

FIG. 5. Axisymmetric finite element models (all dimensions in meters): (a) low frequency tube for transmission loss prediction, (b) high frequency tube for transmission loss prediction, (c) low frequency tube for laser measurement, (d) high frequency tube for laser measurement.

around their circumference. Each sample, as noted previously, was 2.5 cm thick. A 2.5 cm sample depth was used here rather than the 7.5 cm sample depth used in earlier work^{1,2} to ensure that there were measurable levels of vibration at the rear surface of the sample.

To ensure good coherence between the reference micro-

phone signal and the laser output, it was found necessary to apply light-weight reflecting tape to the sample surface: see Fig. 3. Reflecting paints and retroreflective dust were also tried in this application, but the use of the reflecting tape resulted in the best coherence between the reference microphone signal and the surface normal velocity signal. It was

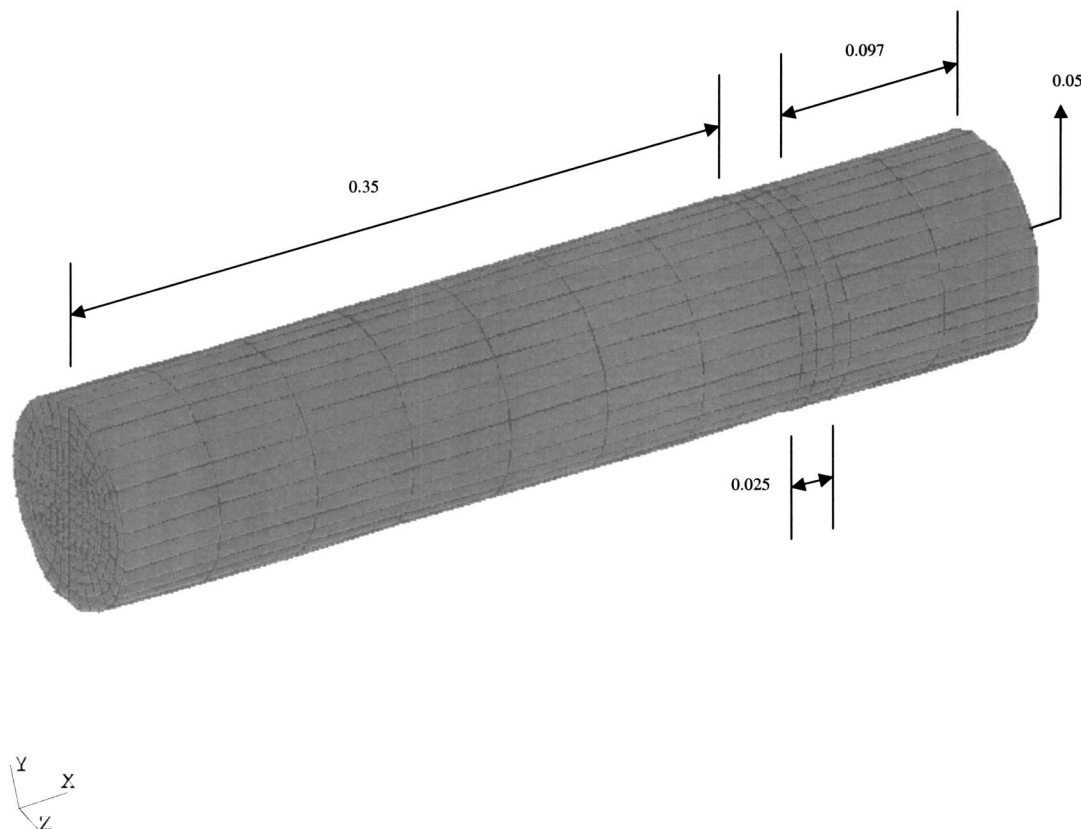


FIG. 6. Three-dimensional finite element model (all dimensions in meters): low frequency tube for laser measurement (plane-constraint case).

assumed that the pieces of tape moved with the same normal velocity as the fibrous material to which they were attached. The pieces of reflecting tape were cut to be smaller than 2.5 mm on each side to minimize sample mass-loading effects.

As shown in Figs. 3(a) and (b), measurements of the edge-constrained samples were made at 11 points along a diametral line: one measurement was made at the center of the sample, and the rest were made at five equally spaced radial points (for a 0.83 cm measurement interval for the large samples and a 0.24 cm sample interval for the small samples) on the left-hand and right-hand hand sides of the center. Measurements were also made for plane-constrained, large samples, as illustrated in Fig. 3(c): in that case, the sample spacing was 0.625 cm. The constraint in the latter case was provided by a 1 mm thick sheet of plexiglass running axially through the sample (as in Ref. 2) which divided the 10 cm diameter into two equal areas. The Plexiglas sheet was held rigidly in place by friction at the tube walls.

To reduce the effect of variations due to sample mounting, tape application inconsistency, and of material inhomogeneity from sample to sample, 20 different samples of the yellow and green glass fibers were measured. The results of the twenty individual tests were then averaged to give the results presented here. Note that in the case of the edge constrained samples it was assumed that the sample's motion was axisymmetric so that measurements could be restricted to a single radial line. The assumption of axisymmetric motion is reasonable in the light of the good agreement between the measured and impedances and the corresponding predic-

tions of the axisymmetric FE model to be discussed below. The effect of any nonaxisymmetric motion resulting from material inhomogeneities was further suppressed by measuring along a complete diametral line and then averaging the results at equal radii on opposite sides of the sample.

D. Mass-loading effect on the sample due to attachment of reflecting tape

Before performing the laser measurements, it was necessary to verify that the acoustical effect of the reflecting tape applied to the samples was small. It was assumed that the tape might alter the response of the glass fiber samples either by mass loading the relatively light-weight glass fiber samples or by partially blocking the rear surface of the sample. Thus the anechoic transmission losses through 10 cm and 2.9 cm diameter, 2.5 cm deep yellow and green glass fiber samples were measured (using the procedure described in Ref. 1) before and after attaching the reflecting tape to their rear surfaces. As shown in Fig. 4, the difference between the transmission loss results with and without reflecting tape was very small. That level of variation was considered to be acceptable, especially for a single sample case. Thus it was concluded that application of the reflecting tape to the samples did not have a significant effect on the latter's acoustical properties.

The transmission loss through the 2.5 cm thick glass fiber samples was also predicted by using the FE model described previously² in combination with the properties listed

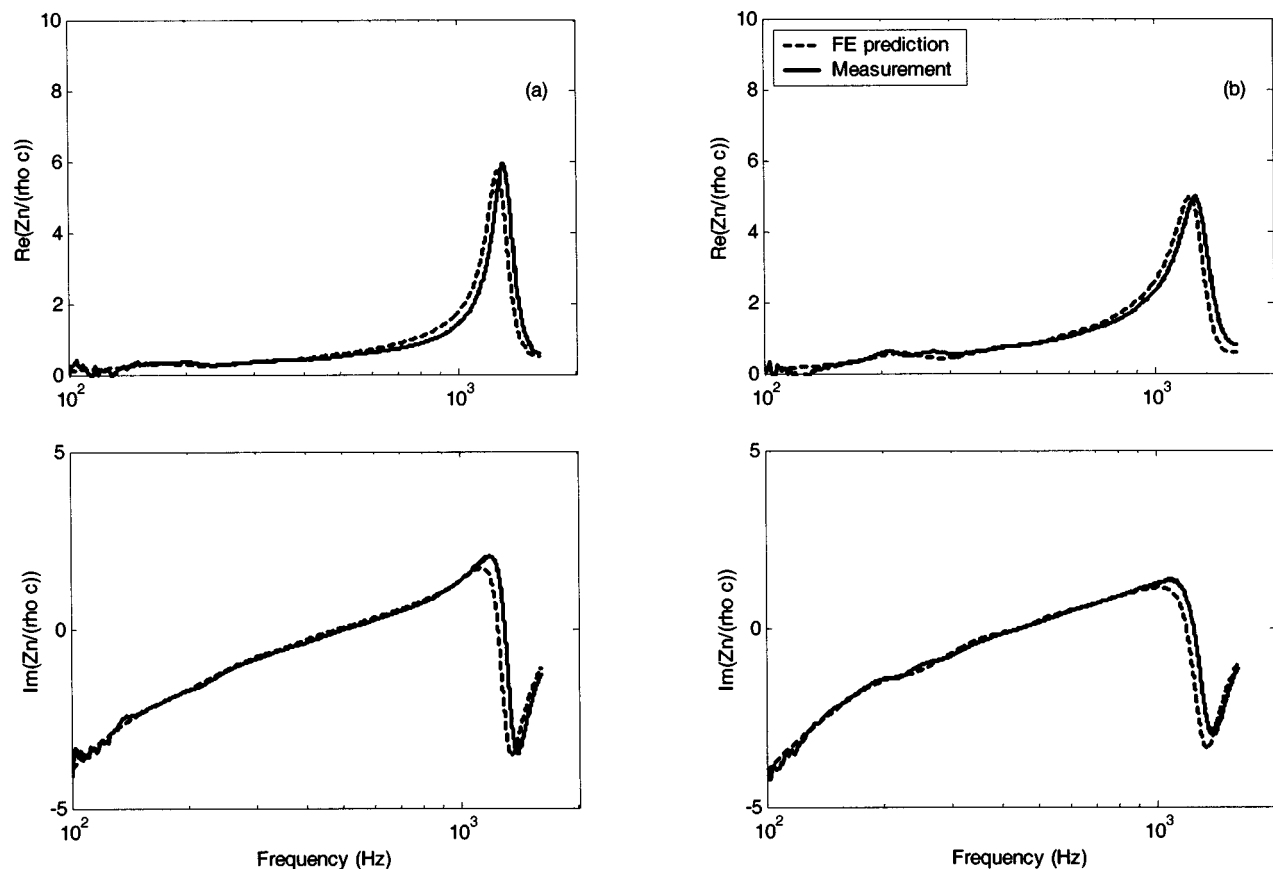


FIG. 7. Surface normal impedances: (a) yellow sample in the low frequency tube, (b) green sample in the low frequency tube, (c) yellow sample in the high frequency tube, (d) green sample in the high frequency tube, (e) yellow sample in the low frequency tube (plane-constraint case), (f) green sample in the low frequency tube (plane-constraint case).

in Table I: see Sec. IV for details of the FE model. It was found that the transmission loss predictions and the corresponding measurements with and without reflecting tape were in reasonable agreement for both yellow and green samples except at the lowest frequencies where the measurements become unreliable owing to the effects of inter-microphone mismatch.

E. Frame velocity of the glass fiber materials

The normal velocity of the solid phase of the glass fiber samples was measured when the sample was driven by a plane sound field. In the low frequency case, a broadband random signal was generated over the frequency range 50 Hz to 1600 Hz. In the high frequency case, the broadband random signal spanned the frequency range 100 Hz to 6400 Hz. The source signal was adjusted to a level that was low enough to avoid a nonlinear response of the sample (as determined by examining the coherence between the reference signal and the laser output). The transfer function between the microphone 1 signal (input) and the surface normal velocity signal (output) was measured by using a Brüel and Kjær Pulse system (Brüel and Kjær Type 3560). The associated coherence in the low frequency tube case for the two materials tested here was excellent from 100 Hz to approximately 800 Hz. However, the coherence above 800 Hz was poor, presumably due to low vibration levels since the excitation of the solid phase by coupling with interstitial fluid

motion becomes progressively weaker as the frequency increases. In the high frequency tube case, it was found that the coherence was excellent from approximately 150 Hz to 1 kHz. As a result of the relatively poor coherence at high frequencies, the vibration results presented here are considered to be accurate only up to frequencies in the neighborhood of 1 kHz.

The transfer function between the reference signal (microphone signal 1), H_1 , and the frame velocity, V_f , can be expressed as

$$V_{f1} = \frac{V_f}{H_1}. \quad (4)$$

It was found that $|H_1|$ was relatively smooth over the frequency range of primary interest (with the exception of a dip in the large tube case at approximately 150 Hz when microphone 1 is approximately one-quarter wavelength in front of the rigid termination). The velocity spectrum was also relatively smooth. In the large sample cases, the spectral level increased up to approximately 200 Hz, presumably since the sample was becoming less stiff with increasing frequency.² Above that frequency, the response fell off as the excitation of the solid phase by the sound field diminished. The same kind of behavior was observed in the small tube case, except that the transition frequency was much higher owing to the higher frequency of the transverse shearing resonance in this case.

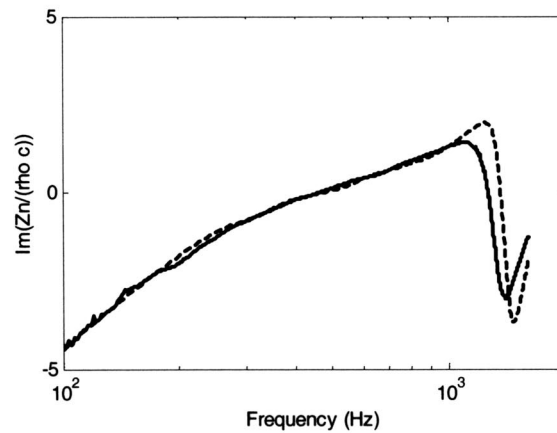
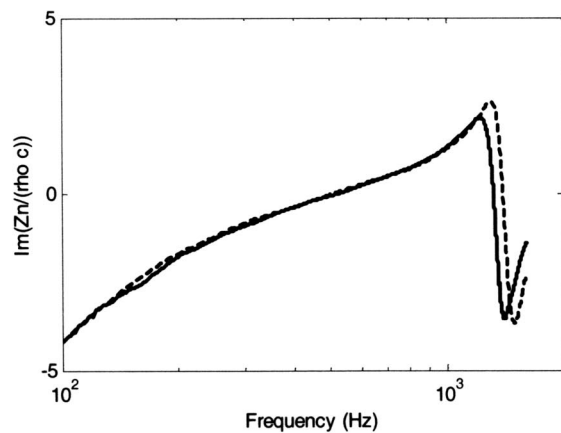
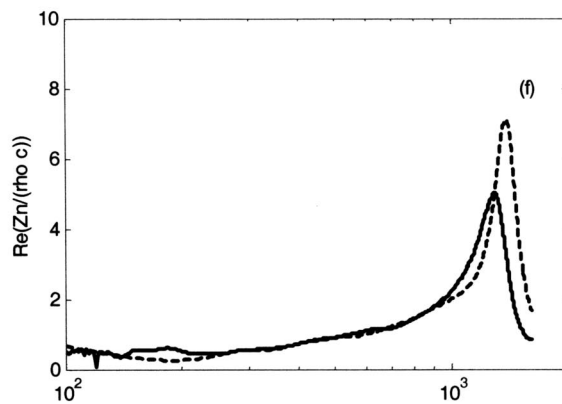
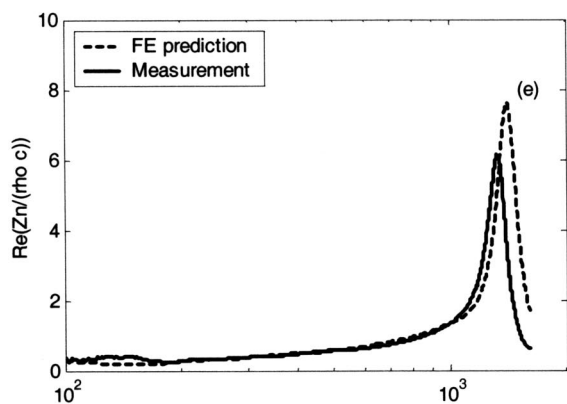
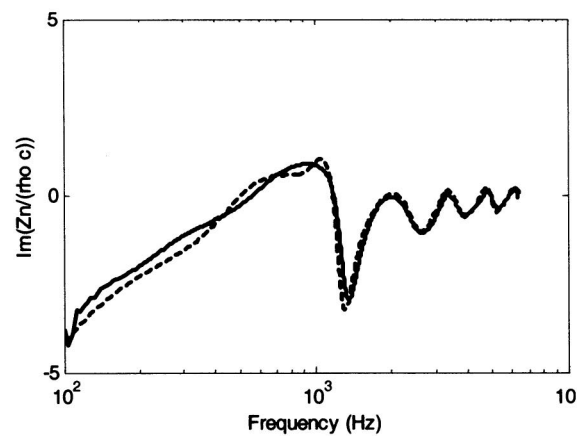
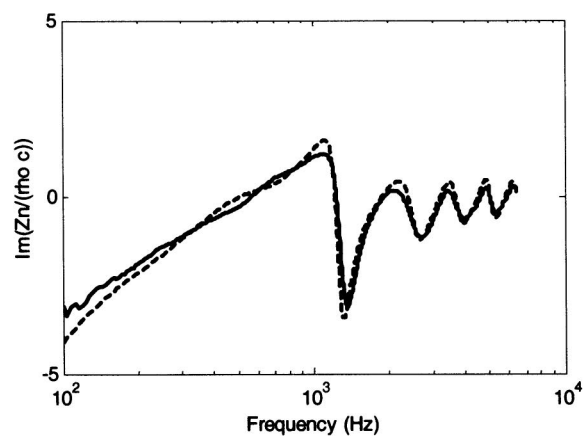
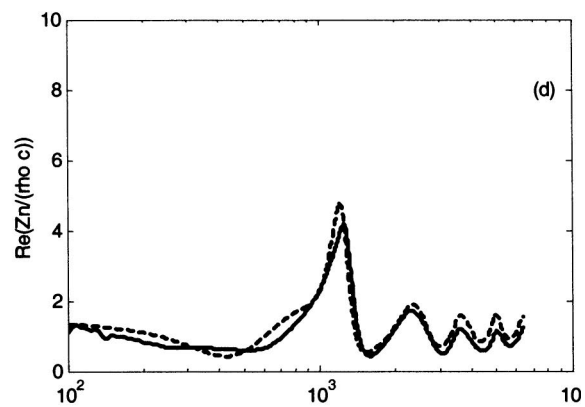
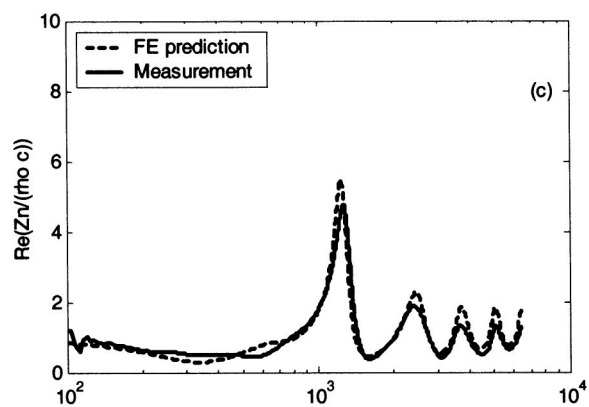


FIG. 7. (Continued.)

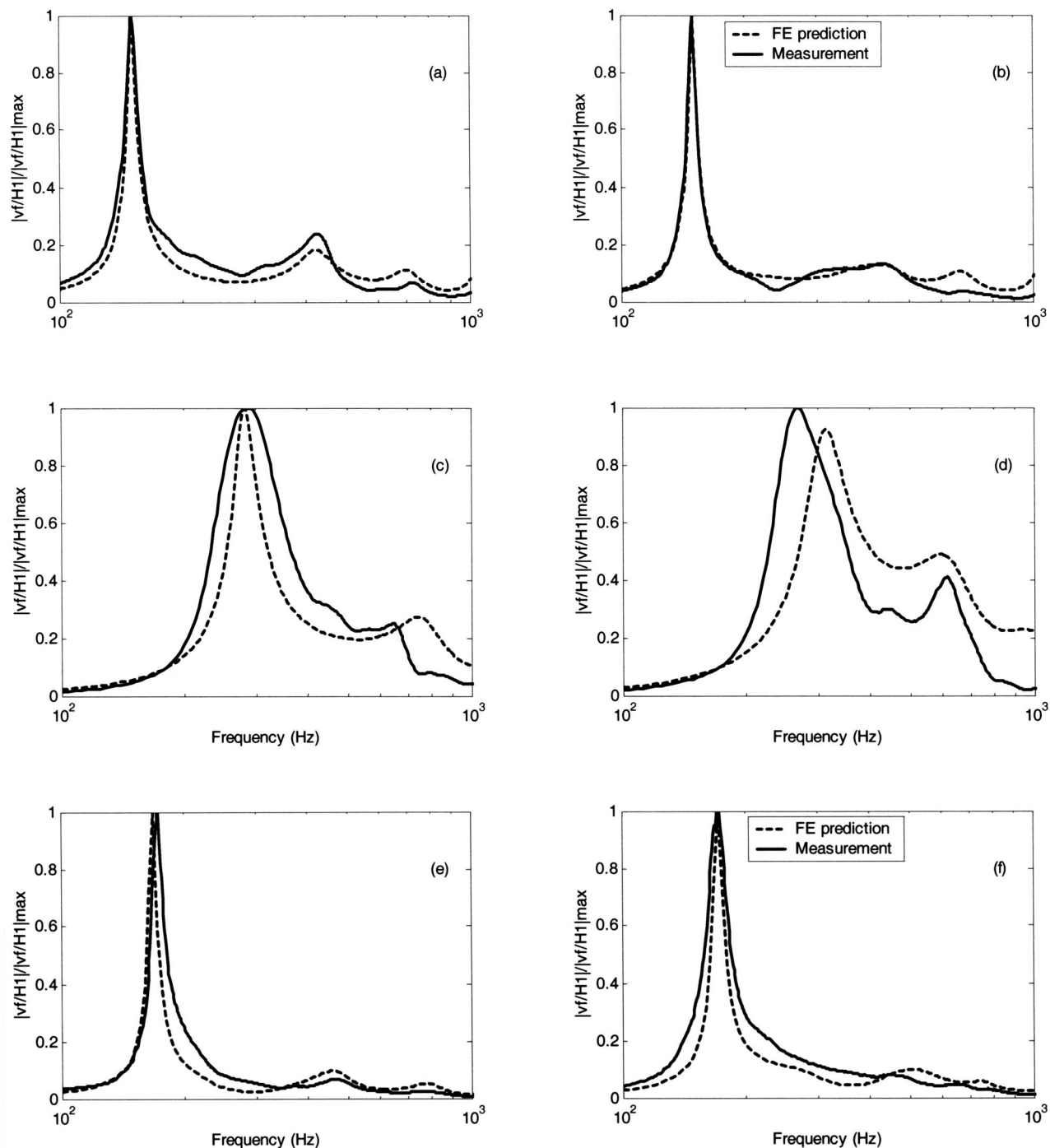


FIG. 8. Magnitude of V_{f1} space-averaged according to Eq. (5) and normalized by maximum value: (a) yellow sample in large tube, (b) green sample in large tube, (c) yellow sample in small tube, (d) green sample in small tube, (e) yellow sample in large tube (plane-constraint case), (f) green sample in large tube (plane-constraint case).

In the case of the edge-constrained samples, the area-weighted, space-averaged, frame root-mean-square velocity can be written as

$$V_{ave} = \frac{\sum_{i=1}^6 \sqrt{V_{f1i} V_{f1i}^*} / 2 S_i}{S_{tot}}. \quad (5)$$

In Eq. (5), each value of V_{f1} represents the average of the two measurements at the same radius on the left-hand and right-hand sides of the center (with the exception of the value at the center of the sample), S_i represents the radially segmented surface area associated with each V_{f1} , and S_{tot} is

the total surface area of the sample. In the case of the plane-constrained samples, a simple average of the seven measured velocities was calculated, i.e., no area weighting was used. When comparing the measured and FE predicted, space-averaged frame velocities, the results were normalized by the maximum value over the frequency range of interest.

IV. FINITE ELEMENT MODELS

Finite element models were created using the code COMET/safe ver. 4.0 to represent both the anechoic and rigid termination cases. As shown in Fig. 5(a), the axisym-

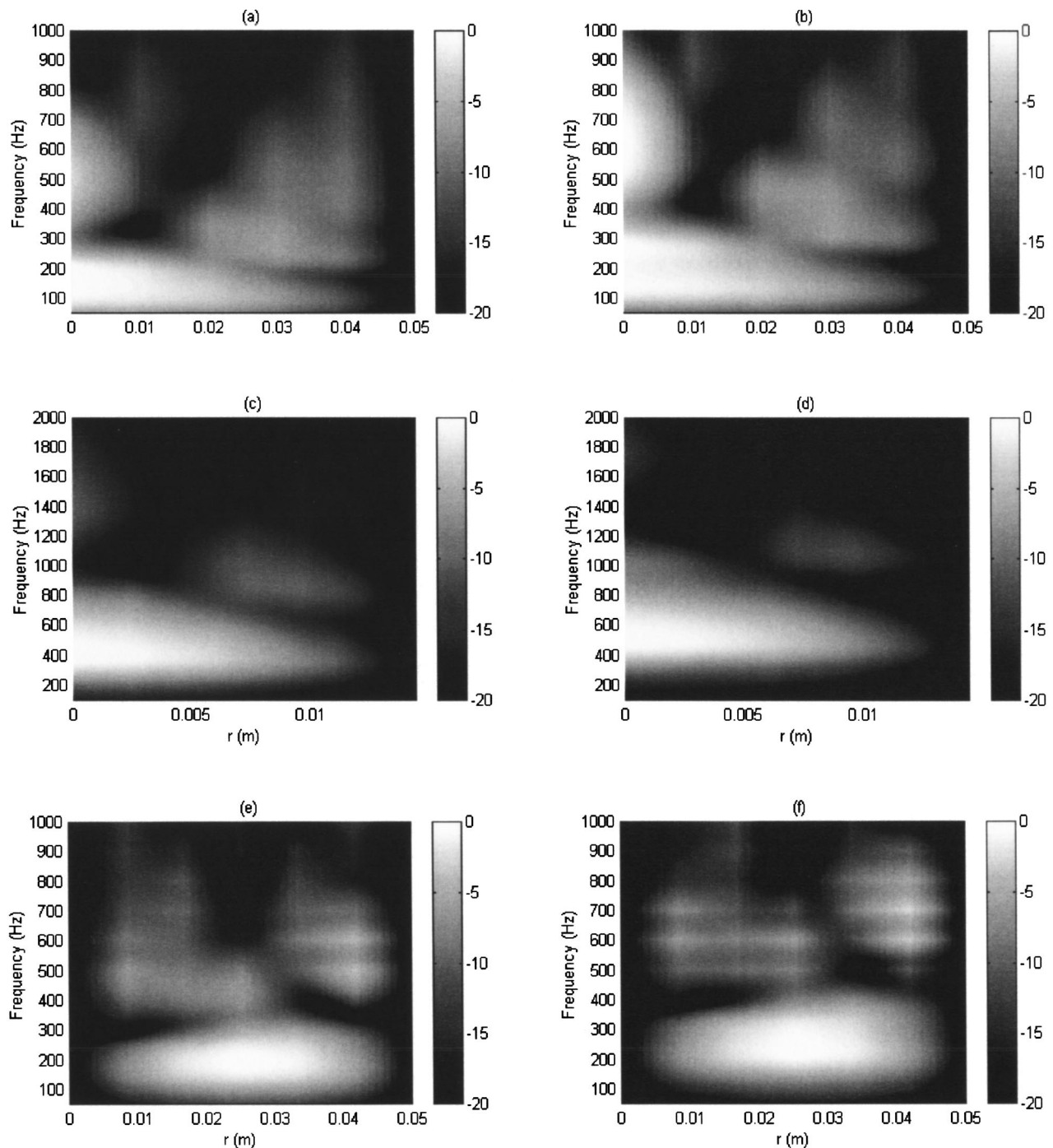


FIG. 9. FE-predicted normalized frame velocities of glass fiber materials in anechoic termination case: (a) yellow sample in large tube, (b) green sample in large tube, (c) yellow sample in small tube, (d) green sample in small tube, (e) yellow sample in large tube (plane-constraint case), (f) green sample in large tube (plane-constraint case).

metric FE model for the anechoic transmission loss case at low frequencies comprised 18 poroelastic elements and 36 air elements for a total of 54 elements having 70 nodes. The FE model in the high frequency anechoic transmission loss case comprised 66 poroelastic elements and 60 air elements for a total of 126 elements having 154 nodes: see Fig 5(b). For the purpose of meshing, all of the elements in the axisymmetric model were chosen to be linear, four-node, quadrilateral elements (ANSYS Shell 63). The impedance tube circumference was itself assumed to be a rigid. Thus in the

air regions, a zero normal velocity condition was applied in the radial direction at the duct wall. Within the poroelastic medium, the radial displacements of both the solid and fluid phases were set to zero at the duct circumference to model the hard wall boundary condition. In addition, to represent the constraint of the glass fiber by contact with the duct wall, the axial component of the solid phase displacement was also set equal to zero at the duct wall, i.e., the material was considered to be “bonded” to the duct wall. Note, however, that the tangential displacement of the fluid component displace-

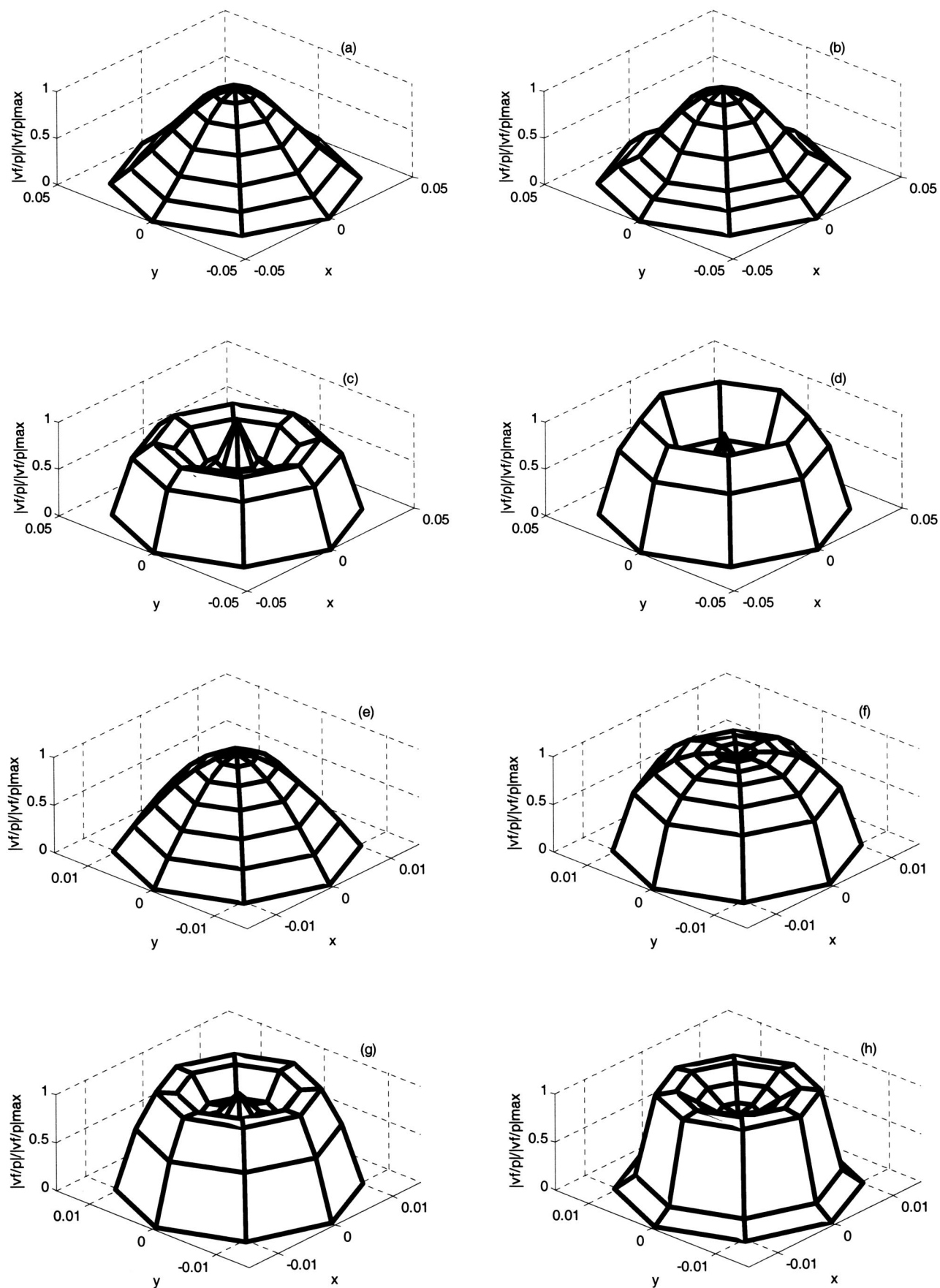


FIG. 10. The first and second mode shapes of the yellow material. Constrained in the large impedance tube: (a) FE prediction at 100 Hz, (b) measurement at 100 Hz, (c) FE prediction at 350 Hz, (d) measurement at 350 Hz. Constrained in the small impedance tube: (e) FE prediction at 390 Hz, (f) measurement at 390 Hz, (g) FE prediction at 830 Hz, (h) measurement at 830 Hz. Cross-sectional dimensions in meters.

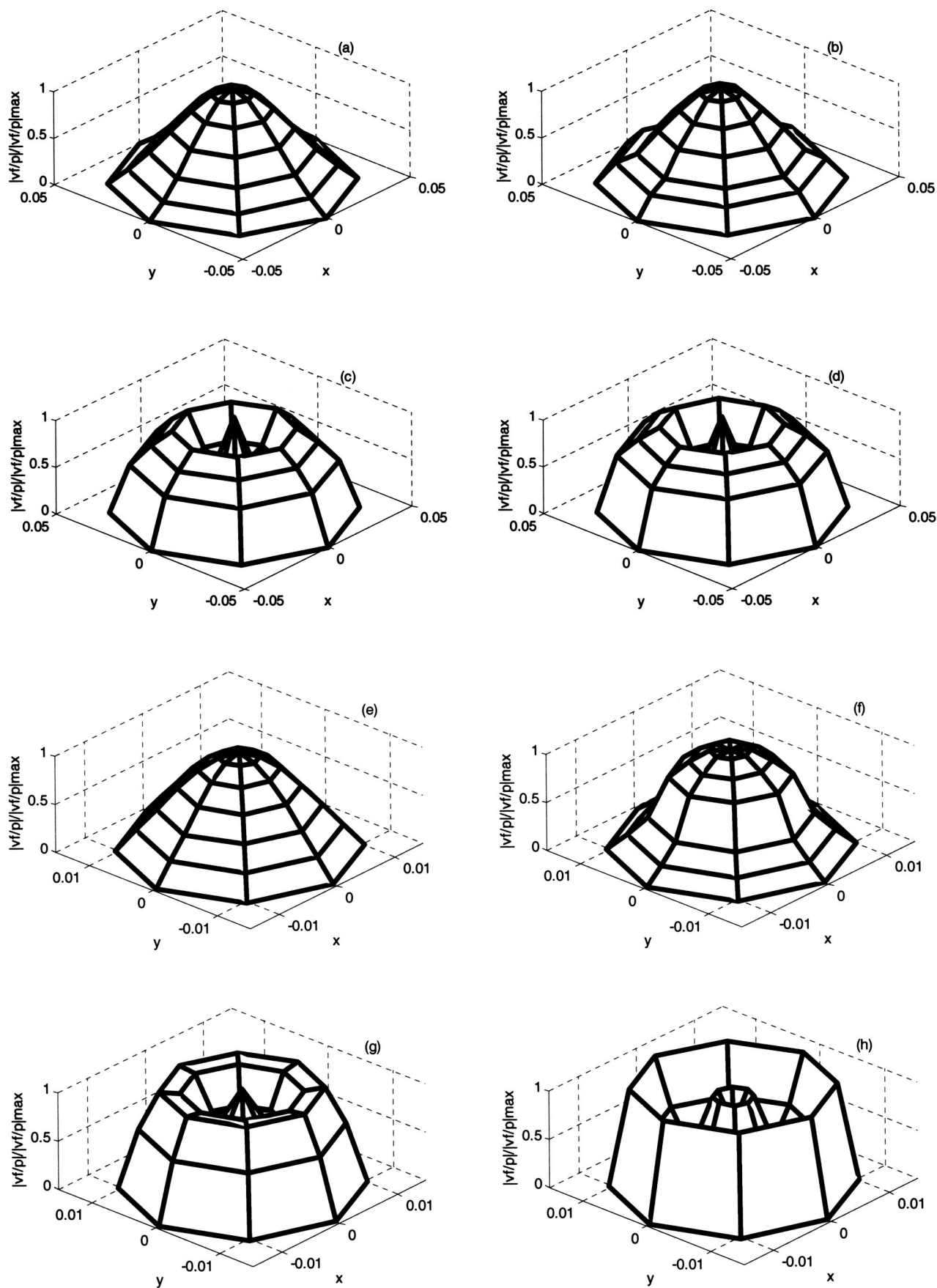


FIG. 11. The first and second mode shapes of the green materials. Constrained in the large impedance tube: (a) FE prediction at 130 Hz, (b) measurement at 130 Hz, (c) FE prediction at 420 Hz, (d) measurement at 420 Hz. Constrained in the small impedance tube: (e) FE prediction at 480 Hz, (f) measurement at 480 Hz, (g) FE prediction at 1120 Hz, (h) measurement at 1120 Hz. Cross-sectional dimensions in meters.

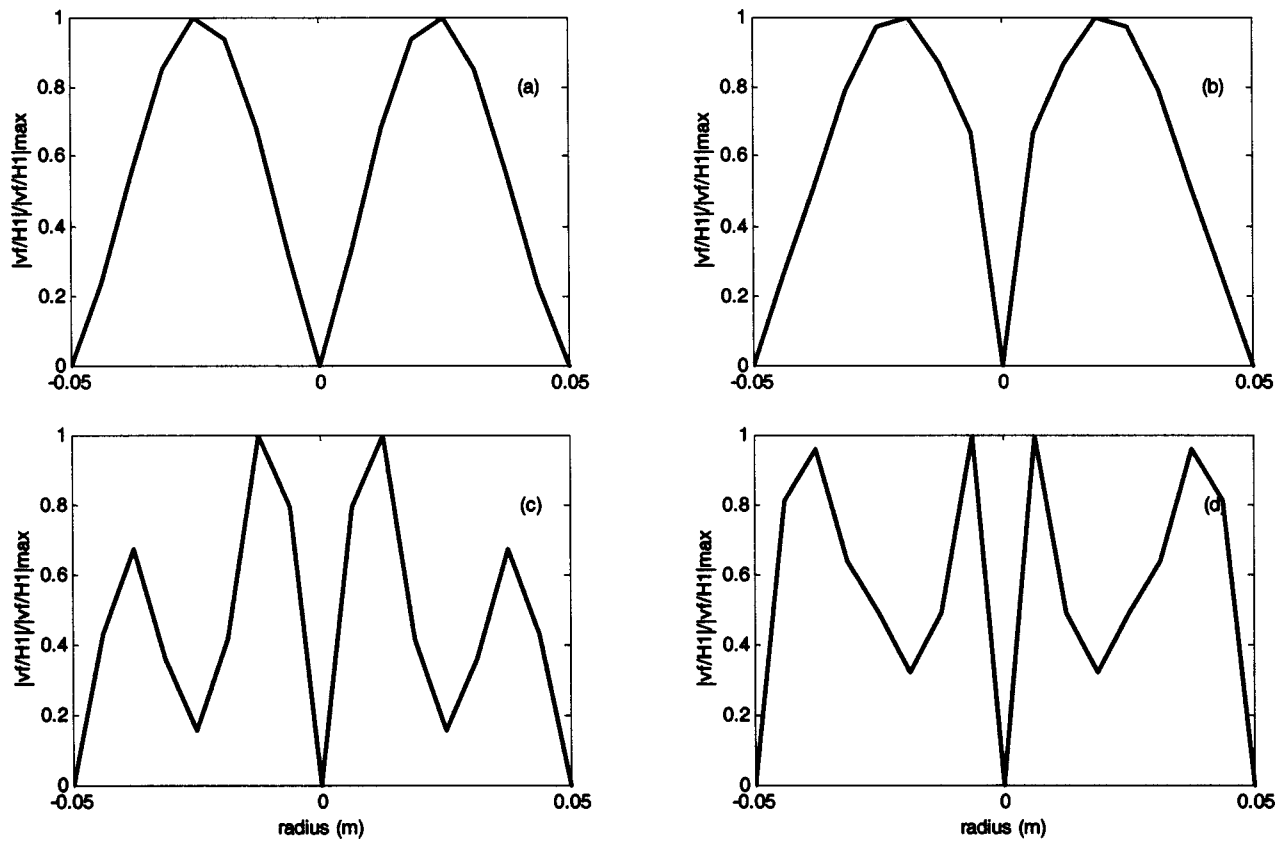


FIG. 12. The first and second mode shapes of the yellow, plane-constrained glass fiber materials in the large impedance tube: (a) FE prediction at 200 Hz, (b) measurement at 200 Hz, (c) FE prediction at 500 Hz, (d) measurement at 500 Hz.

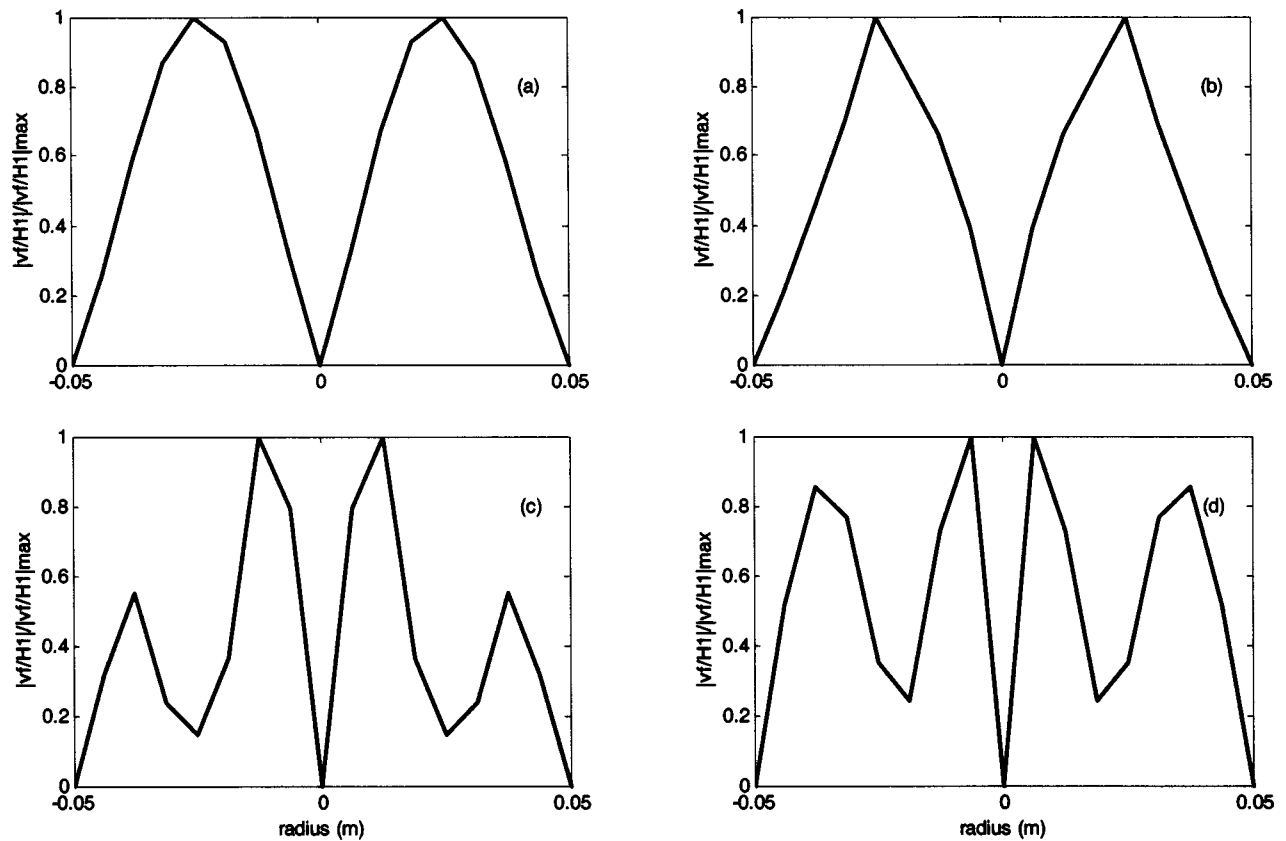


FIG. 13. The first and second mode shapes of the green, plane-constrained glass fiber materials in the large impedance tube: (a) FE prediction at 240 Hz, (b) measurement at 240 Hz, (c) FE prediction at 640 Hz, (d) measurement at 640 Hz.

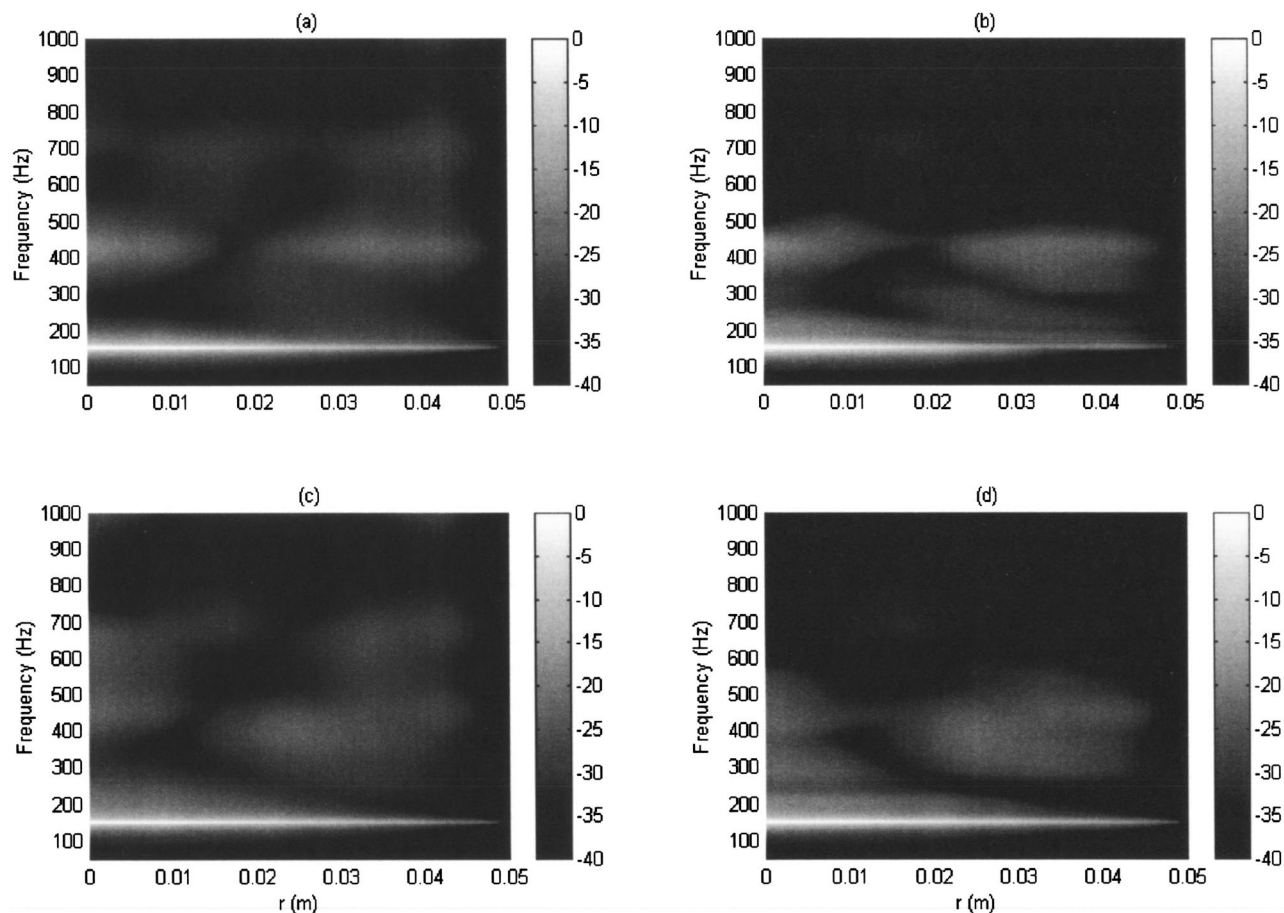


FIG. 14. Normalized frame velocity of glass fiber materials in large tube, edge-constrained, hard termination case: (a) yellow sample (FE prediction), (b) yellow sample (measurement), (c) green sample (FE prediction), (d) green sample (measurement).

ment within the porous material was not constrained at the duct wall, i.e., the fluid phase of the poroelastic medium was allowed to “slip” at the boundary. The parameters that were used to specify the properties of the poroelastic medium were: flow resistivity, tortuosity, porosity, bulk density of the expanded material, bulk Young’s modulus and corresponding loss factor, and Poisson’s ratio. The Young’s modulus and Poisson’s ratio were chosen to give the value of the shear modulus shown in Table I.² The nodes at the left hand end of the models were given a unit axial velocity, and the impedance at the right hand end of the models was set to $\rho_0 c$ to create an anechoic termination.

Models were also created to simulate the laser measurement configuration. Shown in Fig. 5(c) is the axisymmetric FE model used to represent the low frequency laser measurement (i.e., rigid termination) case. The model comprised 18 poroelastic elements and 54 air elements for a total of 72 elements having 91 nodes in the low frequency cases. In the high frequency case, the axisymmetric model comprised 66 poroelastic elements and 84 air elements for a total of 150 elements having 182 nodes: see Fig. 5(d). In these cases, the axial velocities at the nodes at the downstream end of the models were set to zero to represent the rigid termination; otherwise, the boundary conditions were as described in the previous paragraph. In both cases, the pressure “measured” at the left-most node on the tube wall was used to normalize

the solid phase displacement at the rear surface of the sample (as in the experimental procedure). That pressure, along with the pressure “measured” at the next node to the right was used to calculate the reflection coefficient and surface normal impedance of the sample by using the two-microphone procedure described in Sec. III B.

The plane-constraint results presented here were calculated using three-dimensional models (ANSYS Fluid 30 elements were used to create the mesh). In all cases, the plane-constraint was assumed to be rigidly fixed to the tube walls and to have zero thickness. The model comprised 576 poroelastic elements and 1728 air elements for a total of 2304 elements having 2717 nodes: see Fig. 6. The normal fluid and solid phase displacements, and the axial solid phase displacement were set to zero where the poroelastic elements contacted the constraint surface. Otherwise, the boundary conditions were as described previously as were the procedures for normalizing the solid phase displacement of the porous material and for calculating the reflection coefficients and surface normal impedances.

V. RESULTS

A. Surface normal impedances

The surface normal impedances for the two sample types are shown in Fig. 7. The results are typical for a porous

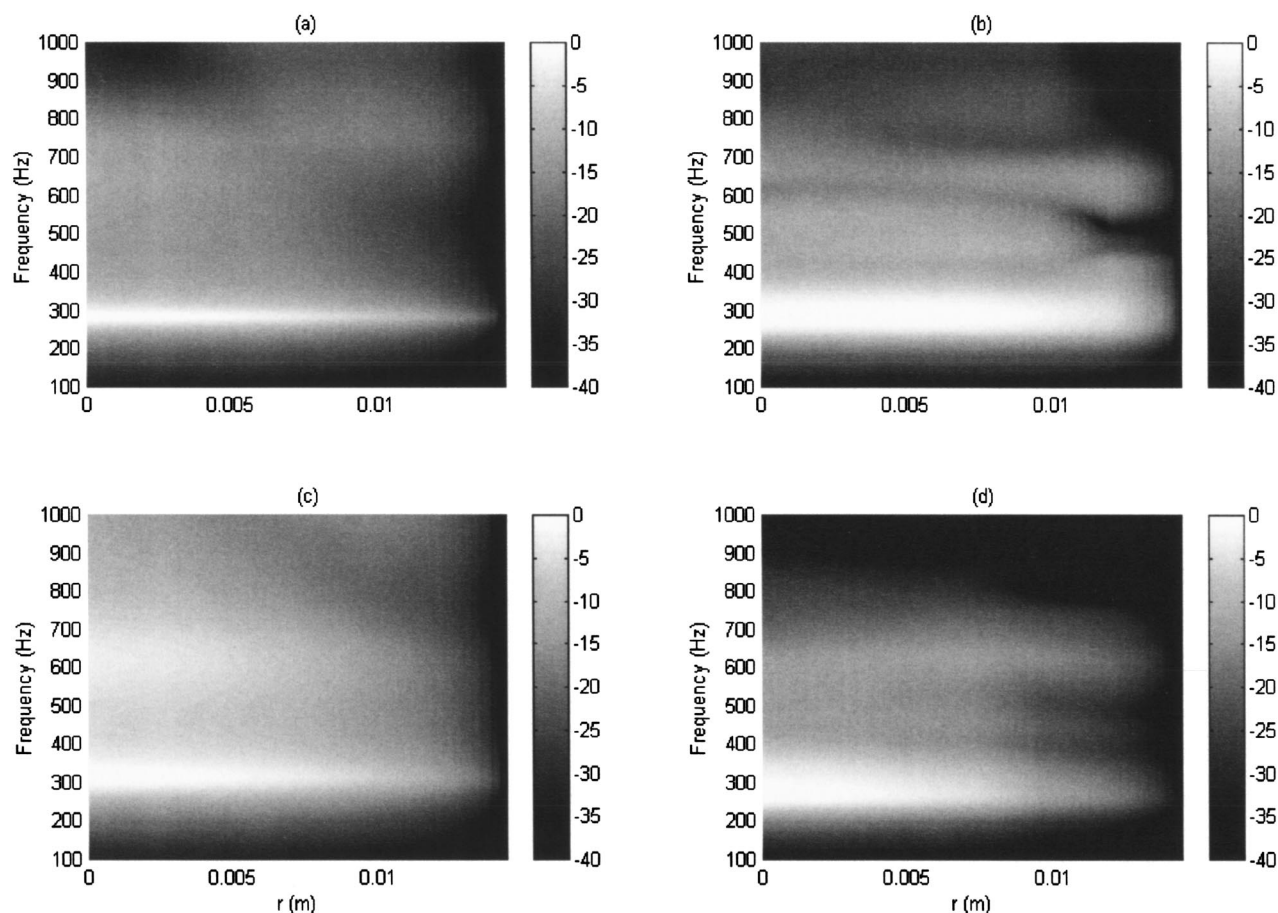


FIG. 15. Normalized frame velocity of glass fiber materials in small tube, edge-constrained, hard termination case: (a) yellow sample (FE prediction), (b) yellow sample (measurement), (c) green sample (FE prediction), (d) green sample (measurement).

layer placed in front of a rigidly terminated air column. That is, the reactive impedance is primarily that of the rigidly terminated air column: stiffness-like at low frequencies and then going through a sequence of depth modes that are progressively more heavily damped as the frequency increases owing to the effect of the fibrous sample [see especially Figs. 7(c) and (d)]. The large and small tube results are essentially similar in the frequency range in which they overlap since the backing airspace depth was approximately the same in both cases. The frequency of the first zero crossing of the reactive impedance is increased by approximately 100 Hz in the small tube case compared to the large tube case since the sample edge-constraint stiffens the sample in the former case.²

It can also be seen in Fig. 7 that the predictions of surface normal impedance made using the FE model in conjunction with the material parameters listed in Table I were in generally good agreement with the measurements for both yellow and green samples. The low frequency results for the small diameter tube in particular show some discrepancies that may be related to the difficulty of maintaining uniform wall contact in the case of 2.9 cm diameter, single layer samples.¹

B. Space-averaged frame velocity

The space-averaged frame velocity results are shown in Fig. 8. It can be seen that there was reasonable agreement

between the FE prediction and the corresponding measurements for both yellow and green samples, particularly with respect to the location of the primary and secondary peaks. However the agreement at higher frequencies in the small tube case was not so close: the FE model generally predicted higher velocity levels than were measured, perhaps due to an underestimate of the frame loss factor. Nonetheless, it was felt that the agreement between measurements and predictions was satisfactory.

C. Identification of sample vibrational modal frequencies

It was desired to visualize the axial motion of the solid phase of the fibrous materials at the frequencies at which they responded primarily in their first and second transverse shearing modes. These frequencies were selected on the basis of FE predictions made for anechoically terminated samples (calculated using the material properties listed in Table I). The anechoic termination condition was chosen for this purpose in order to minimize the influence on the magnitude of the normalized frame velocity of the longitudinal standing waves in the rigidly terminated tube. Further, it was assumed that the frequency ranges over which particular shearing modes of the samples were significant were not strongly affected by the sample termination conditions.

In Fig. 9, the FE-predicted frame velocity results for

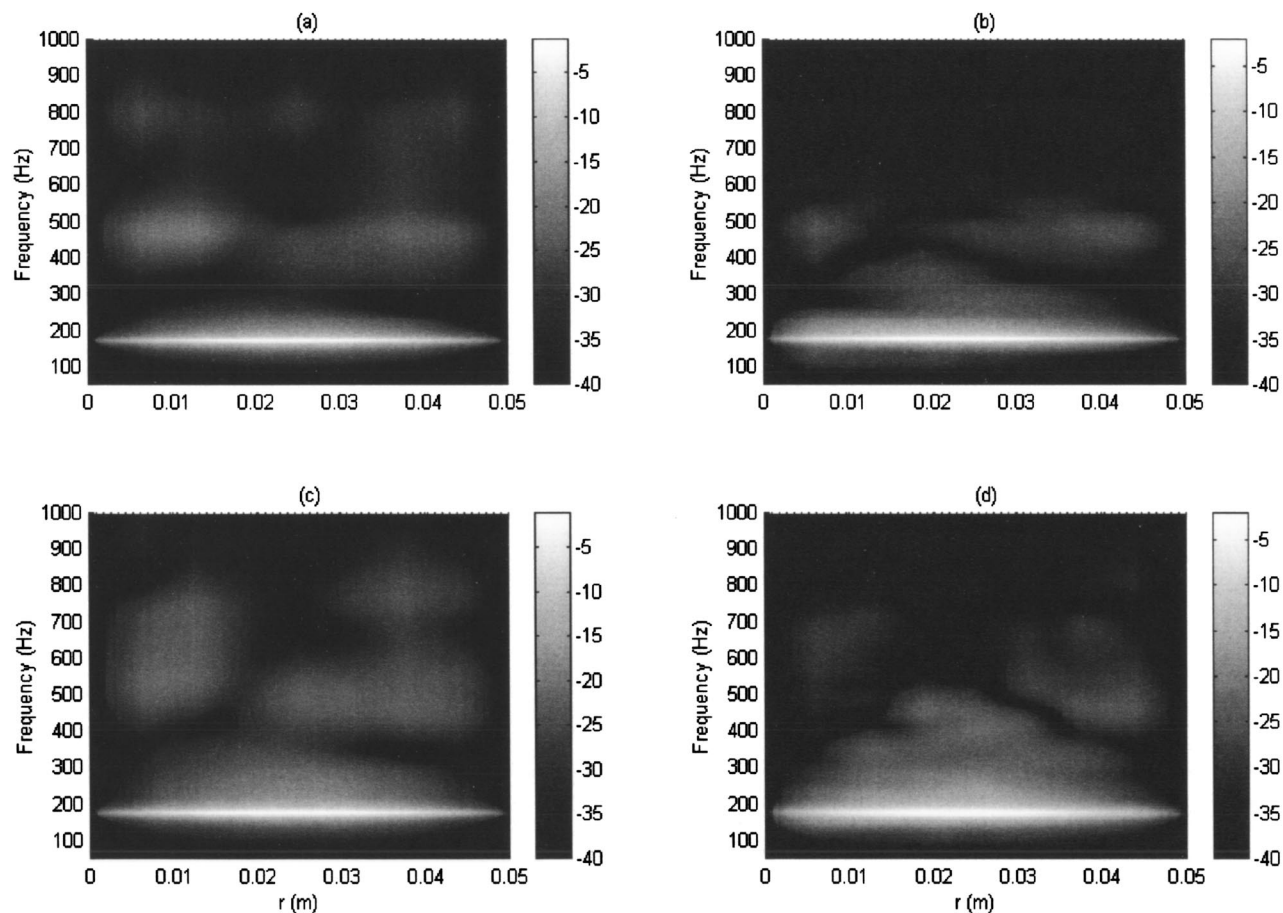


FIG. 16. Normalized frame velocity of glass fiber materials in large tube, plane-constrained, hard termination case: (a) yellow sample (FE prediction), (b) yellow sample (measurement), (c) green sample (FE prediction), (d) green sample (measurement).

both materials in the large and small tube configurations in the edge-constrained cases and in the large tube configurations for the plane-constraint cases are plotted as a function of frequency and radius. The frame velocity amplitudes were normalized by their maximum value within each data set, and the results are presented using a decibel vertical scale. As can be seen in Fig. 9(a), the first transverse mode of the yellow sample in the large tube is clear near 100 Hz and a second mode is visible near 350 Hz. The first mode of the green sample in the large tube is visible near 130 Hz and the second mode is evident near 420 Hz: see Fig. 9(b). In Fig. 9(c), it can be seen that the first mode of the yellow sample in the small tube occurs near 390 Hz while there is a second mode visible near 830 Hz. The first mode of the green sample in the small tube occurs near 480 Hz and a second mode is visible near 1120 Hz: see Fig. 9(d). Note that the various frequencies identified here also correspond approximately with the locations of the minima in the anechoic transmission loss spectra for these various combinations of samples and tube sizes. In the same way, frequencies of 200 Hz and 500 Hz were selected for the yellow material in the plane-constraint case, and 240 Hz and 640 Hz were selected for the green material in the same configuration: see Figs. 9(e) and (f), respectively.

The spatial distribution of the frame velocity when the samples were backed by a finite-depth air space and a rigid

termination were then plotted at the frequencies mentioned above: those results will be discussed in the next sub-section.

D. Comparison of measured and FE-predicted mode shapes

As described in Sec. IV, the spatial variation of the axial velocity of the solid phase of the glass fiber materials was predicted using the code COMET/Safe. The mode shapes in the hard backing configuration were then plotted at the frequencies selected on the basis of the anechoic termination cases. The predicted and measured first and second mode shapes for the edge-constrained materials in the large and small tubes are plotted in Figs. 10 and 11 for the yellow and green materials, respectively, and the plane-constraint results for the two materials are shown in Figs. 12 and 13. Note that although the edge-constraint results are shown over the entire rear surface of the sample, the data was measured over only one diametral line. Recall that data measured at equiradial points were averaged, and those results were further averaged over the set of twenty samples. The latter results were then “rotated” through 360 degrees to give the results shown in Figs. 10 and 11. Further, the velocity at the tube wall location, which could not be measured directly, was set equal to zero for the purpose of plotting. In the same way, the plane-constraint results presented in Figs. 12 and 13 were

“mirrored” with respect to the constraint-plane location, and the velocity both at the constraint plane and the tube wall was set equal to zero for the purpose of plotting.

It can be seen in Figs. 10 and 11 that the measured mode shapes are in generally very good agreement with the FE predictions, particularly in the large tube cases and for the first mode in the small tube cases. The agreement for the second mode in the small tube is not as close; nevertheless, it is clear that the samples are responding in a second-mode-like manner. That is, they show a peak midway between the sample center and the tube wall and there is evidence of a circumferential nodal line between the latter peak and the sample center. In the plane constraint cases (Figs. 12 and 13), there is very good agreement between the measured and predicted results for the first mode. The locations and relative values of the peaks differ somewhat for the second modes, but the measurements and predictions are nonetheless very similar in overall character.

It is clear from the results shown in Figs. 10–13 that the frame motion of the fibrous samples approaches a minimum at the tube wall, thus confirming that the frame is constrained to have zero axial velocity where it contacts the inner tube surface. Also note the circumferential nodal line that is visible in the case of the second modes in both large and small tube, edge-constrained results: see Figs. 10 and 11. These measured results, and those for the plane constraint cases, when combined with the relatively good agreement with the predicted mode shapes, thus confirm that the solid phase of edge-constrained fibrous samples placed in a standing wave tube responds in a shearing motion when driven by plane sound fields. We believe that the present results provide unequivocal evidence that the transmission loss minima discussed in Ref. 2 result from transverse shearing resonances of the solid phase of the constrained samples.

The complete set of predicted and measured normalized frame velocity spectra in the hard backing configuration are plotted as a function of radius in Figs. 14–16: all results are plotted using a decibel vertical scale. In Fig. 14 it can be seen that there is good overall agreement between the measurements and FE predictions in the large tube case, although the high frequency response is generally overestimated in the FE predictions and the higher order mode features occur at lower frequencies in the measurements than in the predictions for the green material, in particular: compare Figs. 14(c) and (d). The agreement in the latter case could have been improved if the shear modulus and loss factor of the green material were adjusted appropriately. However, it was decided here to make predictions using only the set of material parameters identified in Ref. 2 in order to maintain consistency with earlier work. Note that the frequencies at which the normalized velocity reached its peak first mode amplitude in the large tube was higher in the rigid termination case than in the anechoic termination case owing to the quarter-wave effect mentioned in connection with the space-averaged frame velocity results (Fig. 8). In the small tube results presented in Fig. 15, the region of good agreement is largely restricted to the vicinity of the first mode. However,

in the case of the green material, the predicted and measured frequencies of the second mode agree well, but the level of response at those frequencies is substantially lower in the measured than in the predicted results. Again, it may have been possible to improve the quality of the agreement between the measured and predicted results by altering the material parameters of the two media; however, it was felt in this case that the discrepancy related more likely to the difficulty of placing relatively small and fragile samples into the impedance tube so that uniform contact was established around the circumference of the sample. Finally, it can be seen in Fig. 16 that the predicted and measured first mode features are in good agreement in the plane-constraint case, but that higher frequency features are more attenuated in the measured than in the predicted results.

VI. CONCLUSIONS

In this paper, direct measurements of the vibrational velocity of the solid skeleton of fibrous samples placed in a standing wave tube have been presented. It was found that such samples respond in a shearing motion when driven by an incident plane wave field owing both to the viscous coupling of the fluid and solid phases of the fibrous material and to the frictional constraint of the samples where they contact the tube wall. The same effect was observed when a plane constraining surface was inserted axially through a sample. Further, the measured and FE-predicted spatial distribution of the frame velocity were in generally good agreement for the first two shearing modes of both edge- and plane-constrained samples. Thus the present results confirm that glass fiber samples placed in a standing wave tube exhibit shearing modes, as inferred from the work described in Refs. 1 and 2, and that those modes are associated with previously observed transmission loss minima.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Jason D. McIntosh of Motorola, Inc., who originally suggested the use of a laser Doppler velocimeter to measure the vibrations of porous materials placed in a standing wave tube.

¹B.H. Song and J.S. Bolton, “A transfer matrix approach for estimating the characteristic impedance and wave number of limp and rigid porous materials,” *J. Acoust. Soc. Am.* **107**, 1131–1152 (2000).

²B.H. Song and J.S. Bolton, “Effect of circumferential edge constraint on the acoustical properties of glass fiber materials,” *J. Acoust. Soc. Am.* **110**, 2902–2916 (2001).

³D.H. Kruger and J.A. Mann III, “Minimizing the sound power radiated by a cube as a function of the size of constrained layer damping patches,” *J. Acoust. Soc. Am.* **105**, 1714–1724 (1999).

⁴S.A. Mueller and F.A. Moslehy, “An inverse vibration method applied to bond evaluation of space shuttle tiles,” *Trans. ASME* **118**, 454–462 (1996).

⁵P.S. Dubbelday, “Poisson’s ratio of foamed aluminum determined by laser Doppler vibrometry,” *J. Acoust. Soc. Am.* **91**, 1737–1744 (1992).

⁶E. Mariez, S. Sahraoui, and J.F. Allard, “Elastic constants of polyurethane foam’s skeleton for Biot model,” *Proceedings of INTER-NOISE 96*, 1996, pp. 951–954.

⁷K. Khirnykh and A. Cummings, “The measurement of frame motion in flexible porous materials,” *J. Acoust. Soc. Am.* **105**, 755–761 (1999).

⁸S. Temkin, *Elements of Acoustics* (Wiley, New York, 1981).

Experimental validation of two elastodynamic models for the wave field generated by ultrasonic transducers

Maarten C. M. Bakker^{a)}

Aerospace Materials Laboratory, Faculty of Aerospace Engineering, P.O. Box 5058, 2600 GB Delft, The Netherlands

Martin D. Verweij^{b)}

Laboratory of Electromagnetic Research, Faculty of Information Technology and Systems, Centre for Technical Geoscience, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands

(Received 30 August 2002; revised 3 January 2003; accepted 12 January 2003)

Two different three-dimensional elastodynamic models are introduced to simulate the wave field generated in steel by two types of surface mounted ultrasonic transducers. By replacing the actual transducer by an equivalent surface source distribution, the models become amenable to an exact analytical analysis. The first model simulates the action of a contact transducer through a distribution of nonmoving line segment sources. The second model simulates the action of an angle beam transducer through a single moving line segment source. Almost any transducer aperture shape may be modeled, while the source may apply a nonuniform traction. To speed up the numerical space-time domain calculations, the Cagniard-De Hoop method is employed to analytically evaluate the wave field produced by a single nonmoving line segment source. This solution provides the integrand for both single-integral models. The models are experimentally validated for a contact transducer and three different angle beam transducers. The validation involves a comparison of the wave-field patterns, the directivity curves and some time-domain signals from the wave field. It is shown that the models reliably identify the wide variety of waves generated by ultrasonic transducers, such as focused waves, edge waves, Rayleigh waves and head waves. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1557213]

PACS numbers: 43.20.Px, 43.20.Rz [LLT]

I. INTRODUCTION

In the practice of nondestructive inspection of steel components for fatigue or corrosion, two important types of ultrasonic transducers are the angle beam transducer and the contact transducer. The angle beam transducer^{1,2} is an assembly of a plastic wedge and a piezo-electric wedge transducer, as depicted in Fig. 1. It is designed to generate a beam of shear waves that propagate under a specific angle with the steel surface onto which the transducer is mounted. For brevity we will refer to an angle beam transducer as a *probe*. The contact transducer is a piezoelectric transducer that is mounted directly onto the steel surface. It is designed to generate a beam of compression waves that propagate perpendicular to the steel surface. Both types of transducer make proper acoustical contact with the steel through a thin layer of low-viscous liquid couplant. As a consequence, only the normal traction is continuous through the transducer-couplant-steel interface, and we may therefore neglect the shear tractions when prescribing the equivalent surface source.³ The part of the interface where the energy is effectively exchanged between the transducer and the steel is referred to as the aperture Ω , which size determines the size of the finite source domain in our model. In the practice of

inspection of steel components, the frequencies generated by ultrasonic transducers lie typically in the range of 0.1 MHz to 10 MHz.

Quite a few authors have investigated equivalent surface sources that could be employed as models for wave transducers. For example, at a sufficient distance one could consider a moving point source⁴⁻⁶ to model a probe. For a rectangular contact transducer a line source or strip source⁷⁻⁹ could serve as a model. Approximating models are usually accurate only in the far field,^{2,10} although some models may still perform satisfactorily in the intermediate field.¹¹⁻¹³ A finite-element¹⁴ or finite-difference¹⁵ calculation may give satisfactory predictions in both the near field and the far field, but the computational effort for high frequencies in combination with large distances can be a drawback.

As will be shown in this paper, the head waves are prominent features of the wave field for small contact transducers and for probes with steep beam angles, while the Rayleigh waves always dominate at the surface. Models that in essence neglect the mode conversions at the free surface fail to give accurate predictions for these important wave field constituents.^{2,12,16,17}

To accurately describe the full wave field of a transducer at any range, the exact elastodynamic solution is required. At a first glance, this solution could be obtained by employing point source superposition. However, in a three-dimensional problem this would result in a double integral over the aperture. This requires an enormous computational effort. The effort increases with decreasing pulse width, since the spatial

^{a)}Electronic mail: bakkerbcm@dutlbcz.lr.tudelft.nl

^{b)}Electronic mail: m.d.verweij@its.tudelft.nl

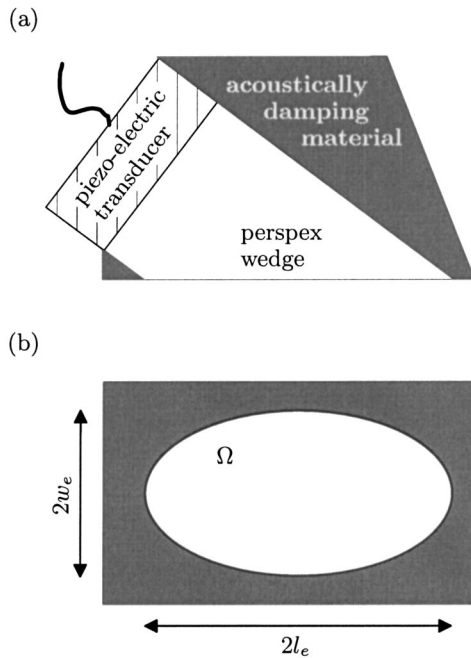


FIG. 1. (a) Side-view of an angle beam transducer (probe). (b) Bottom view of the probe.

discretization of the double integral is linked to the time sampling, which must be relatively fine in view of Shannon's sampling theorem. Although a combination of approximations and smart programming may considerably alleviate the computational effort,¹⁸ the evaluation would still suffer from the approximation drawbacks discussed above. For some source shapes the double integral can be evaluated (partly) analytically using contour integration. Examples are circular sources,^{19–21} annular sources²² and rectangular sources.^{23–25}

In this paper we model the probe by a moving line segment of normal traction, where the motion of the load causes the angled beam of shear waves that is characteristic for this type of transducer. In addition, we model the contact transducer by a distribution of nonmoving line segments of normal traction. Next, we employ a closed form solution for the space–time domain wave field of a single, nonmoving line segment source. This intermediate result is obtained using the Cagniard–De Hoop technique,²⁶ which is also demonstrated in Refs. 6, 8, 9, and 25. For both models, the intermediate result needs to be integrated to account for the entire aperture. This integration is performed numerically. In fact, both models employ the superposition of closed form results for line segment sources, which requires considerably less computational effort than the superposition of closed form results for point sources.

In this paper we focus on the elliptically shaped transducer aperture, which case covers most common commercial transducers. Nevertheless, with our approach almost any aperture shape could be modeled. To model the pulse generated by an ultrasonic transducer, we apply a time convolution with a source signature obtained from a calibration measurement. Due to its convolution structure, the additional integral does not burden the total calculation time. The elastodynamic models are validated against measurements involving

three types of shear wave probe and one compression wave contact transducer.

In Sec. II we describe the model configuration and present the space–time domain wave field solutions. In Sec. III we outline the experimental setup. In Sec. IV we compare and discuss the modeled and measured wave fields, directivity curves, and time domain signals, and in Sec. V we draw some conclusions.

II. MODEL CONFIGURATION AND SPACE–TIME DOMAIN WAVE FIELD

A. Configuration

We consider a homogeneous, isotropic, lossless elastic half-space with an initially traction-free surface. A fixed Cartesian coordinate system is chosen, where a positive vertical z -coordinate indicates the depth of the observation point $\mathbf{x} = x\mathbf{i}_x + y\mathbf{i}_y + z\mathbf{i}_z$ below the surface. At $t=0$ a part of the surface is loaded by a source of normal traction T_z . The elliptical source domain Ω lies in between $x = -l_e$ and $x = l_e$ in one horizontal direction, and in between $y = -w(x)$ and $y = w(x)$ in the other horizontal direction, where $w(x) = w_e \sqrt{1 - (x/l_e)^2}$. The values of l_e and w_e follow from the length $2l_e$ and the width $2w_e$ of the aperture of an actual transducer.

B. Source model for a contact transducer

We describe the contact transducer by the distribution of nonmoving line sources

$$T_z(x, y, t) = \int_{-l_e}^{l_e} H(t) \delta(x - x') g^{ac}(x') \times \{H[y + w(x')] - H[y - w(x')]\} dx'. \quad (1)$$

This distribution fills up the source domain Ω with shifted line sources that are directed parallel to the y axis and have a length $2w(x')$, where x' is the position of a line segment along the x axis. One such line source is shown in Fig. 2(a). The temporal behavior of a line source is determined by the step function $H(t)$, its local strength is given by the traction amplitude $g^{ac}(x')$ with dimension force per area. Suppose that we know the displacement Green's function $G_k(\mathbf{x}, t)$ due to an impulsive line source $\delta(t)\delta(x)[H(y + w) - H(y - w)]$. Closed-form expressions for this space–time domain Green's function are presented in the Appendix. Superposition of the responses to all line sources inside Ω gives us the displacement due to the contact transducer as

$$u_k(\mathbf{x}, t) = H(t) * \int_{-l_e}^{l_e} g^{ac}(x') G_k(x - x', y, z, t) dx', \quad (2)$$

where $G_k(x - x', y, z, t)$ also depends on x' through $w(x')$. The above process may be compared to the conventional point source superposition that would result in a double spatial integral. To change the time behavior of the contact transducer into something else than a step function, a temporal convolution of Eqs. (1) and (2) with an appropriate unit amplitude source signature $g^t(t)$ should be applied.

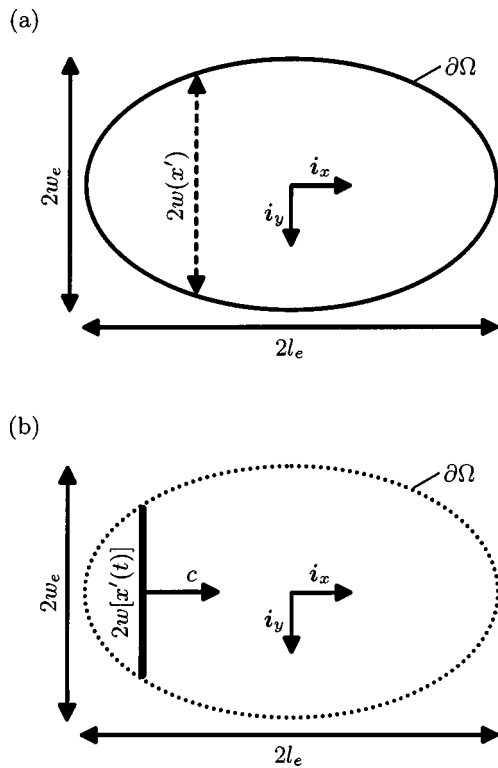


FIG. 2. (a) The contact transducer aperture, which as a source domain of normal traction is build up from line segments of length $2w(x')$. (b) The probe aperture, which as a source domain is traversed by a moving line segment of normal traction of length $2w[x'(t)]$.

C. Source model for a probe

We describe the probe by the moving line source

$$T_z(x, y, t) = H(t) \delta[x - x'(t)] g^{\text{ap}}(x) \times \{H[y + w(x)] - H[y - w(x)]\} \quad (0 \leq t \leq T_c), \quad (3)$$

with $x'(t) = ct - l_e$ being the momentary position of the line source at $t = (x' + l_e)/c$. The line source moves inside the source domain Ω at a constant speed $c > 0$ along the x axis; see Fig. 2(b). It becomes active at $t = 0$ in $x = -l_e$, which is at the back of the probe aperture, and it vanishes at $t = T_c = 2l_e/c$ in $x = l_e$, which is at the front of the probe aperture. The local strength of the line source is given by the traction amplitude $g^{\text{ap}}(x)$ with dimension force per length. Next we write the moving line source as the distribution of delayed and shifted nonmoving line sources

$$T_z(x, y, t) = \int_0^{T_c} \delta(t - \tau) \delta[x - x'(\tau)] g^{\text{ap}}(x) \times \{H[y + w(x)] - H[y - w(x)]\} d\tau. \quad (4)$$

Suppose that we know the displacement Green's function $G_k(\mathbf{x}, t)$ due to a line source $\delta(t) \delta(x) [H(y + w) - H(y - w)]$, as presented in the Appendix. Then superposition of the responses to all delayed and shifted line sources inside Ω renders us the displacement¹¹

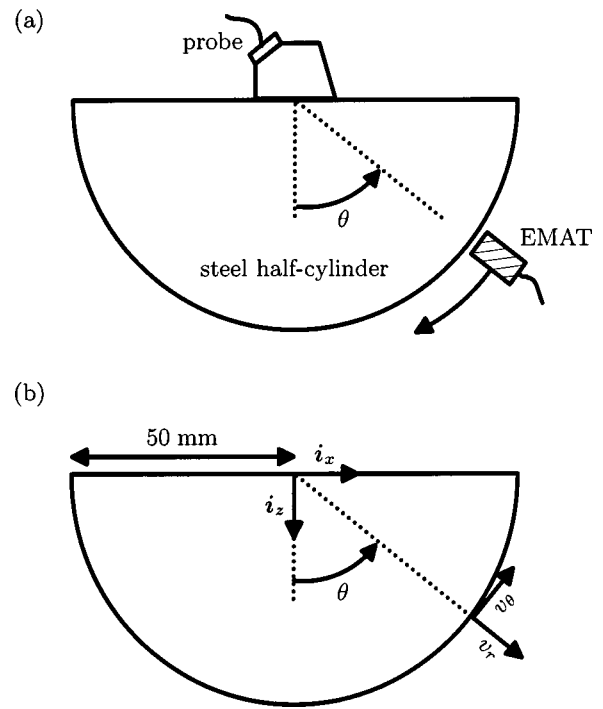


FIG. 3. (a) The experimental setup with a probe centrally attached to the plane surface of the steel half-cylinder. The EMAT is automatically rotated around the curved cylinder surface in steps of 2° . (b) The velocity components v_r and v_θ in the employed cylindrical coordinate system.

$$u_k(\mathbf{x}, t) = \int_0^{\min\{t, T_c\}} G_k[x - x'(\tau), y, z, t - \tau] g^{\text{ap}}[x'(\tau)] d\tau, \quad (5)$$

where $G_k(x - x', y, z, t)$ also depends on x' through $w(x')$. The above process may be compared to the conventional moving point source superposition that would result in a double spatial integral. The steplike time behavior of the probe may be changed into something else by a temporal convolution of Eqs. (3), (4), and (5) with an appropriate unit amplitude source signature $g'(t)$.

D. Particle velocities in the cylindrical coordinate system

In our experimental setup we will measure the particle velocity $\mathbf{v} = \partial_t \mathbf{u}$. Since the interior of the steel half-space is inaccessible for measurements, we will measure the wave field along a circular cylindrical surface, as depicted in Fig. 3(a). To facilitate these measurements, we introduce a cylindrical coordinate system with coordinates (r, θ, ζ) to replace the Cartesian coordinates (x, y, z) . The particle velocity components v_r and v_θ [see Fig. 3(b)] are obtained from the Cartesian ones by using

$$v_r = v_x \sin(\theta) + v_z \cos(\theta), \quad (6)$$

$$v_\theta = v_x \cos(\theta) - v_z \sin(\theta). \quad (7)$$

The component $v_\zeta = v_y$ is perpendicular to the plane of the drawing in Fig. 3(b). This component is zero in the symmetry plane of the transducer in which the validation will take place.

TABLE I. Parameters of the employed transducers.

Transducer type	Diameter of crystal (mm)	Frequency (MHz)	l_e, w_e (mm, mm)	Load speed (m/s)
Contact transducer	12.7	2.25 MHz	6.35, 6.35	
32° probe	12.7	5.00 MHz	7.90, 6.35	6658
45° probe	25.4	2.25 MHz	17.5, 14.0	4732
70° probe	25.4	5.00 MHz	22.5, 14.0	3506

III. EXPERIMENTAL SETUP

Figure 3(a) depicts a side-view of our setup for measuring the wave field from an ultrasonic transducer. The transducer is mounted centrally on the flat surface of a steel half-cylinder, so that it radiates towards the curved cylinder surface. An automated stepping machine rotates an EMAT, being a type of noncontact ultrasonic receiver, along the curved cylinder surface. In this fashion, we have taken automatic wave-field measurements from $\theta = -89^\circ$ to $\theta = 89^\circ$ in steps of 2° . The radius of the cylinder is 50 mm and the bulk wave speeds of the steel are measured as $C^P = 5888$ m/s and $C^S = 3228$ m/s. The frequencies of the measured signals lie roughly in the range from 0.5 MHz to 6 MHz. The signals are sampled in time at 50 ns intervals (20 MHz) and in amplitude at 255 levels (8 bits), and are subsequently stored on a computer disk for further processing.

We have used an electromagnetic acoustic transducer¹⁵ (EMAT) to measure the wave field. The EMAT has the advantages that it has a small aperture and does not have to make contact with the steel, since its signal is obtained through electromagnetic induction. The motion of the steel particles, which are magnetized by a permanent magnet in the EMAT, induces a small emf in the coil of the EMAT. The detected signal is proportional to the velocity of the steel surface. As a consequence of this principle, the receiver is sensitive to lift-off, i.e., the size of the airgap between the EMAT and the steel surface.

Two dedicated types of EMAT have been designed. An R-EMAT is constructed to be most sensitive to the radial velocity component v_r , while a T-EMAT is most sensitive to the transversal velocity component v_θ . The spatial resolution of the EMAT proves to be sufficient to have a negligible influence on the results shown in this paper. The suppression of the cross-talk, i.e., the undesired complementary velocity component, is expected to be much better for the T-EMAT than for the R-EMAT. The signal-to-noise ratio (SNR) is about 35 dB for the T-EMAT and about 25 dB for the R-EMAT. The fact that the two EMATs have quite different characteristics prohibits us from separately validating the components of the wave field. Therefore, we will use the generally better T-EMAT, with an exception for the wave field of the 45° probe where the R-EMAT yields some more details. The measurements have been conducted in the symmetry plane $\zeta = 0$ of the transducer, which leaves us with the cylindrical velocity components v_r and v_θ , as depicted in Fig. 3(b).

The curved cylinder surface will cause reflected waves of all wave types and cause mode converted waves from mainly the head waves, Rayleigh waves and edge waves. These will not show up in our calculations based on a steel

half-space. By comparing the measurements with the model calculations, these waves are easily identified. Moreover, these may even be useful. An example is the mode conversion of the Rayleigh waves at the corners of the plane surface, i.e., where $\theta = \pm 90^\circ$. Here, the magnetic field produced by the EMAT is distorted, which results in a rather poor detection of the wave field. Nevertheless, when a Rayleigh wave reaches a corner, a part of its energy will be converted into a surface wave that can propagate along the curved surface.²⁷ The latter wave shows up clearly in the data as an approximately linear feature, which forms clear evidence of the presence of a Rayleigh wave at the plane surface.

The experimental setup introduces three more or less unavoidable causes of error with regard to the location of a measurement. First, the transducer may be misaligned with respect to the observation plane, which should be the symmetry plane of the transducer. Second, an offset in the y direction may cause the measurements to take place in a plane that is offset with respect to the symmetry plane of the transducer. Third, the machine that rotates the EMAT around the curved cylinder surface may be attached with an angular offset, which causes the measured range to deviate from $-89^\circ \leq \theta \leq 89^\circ$. These errors may affect the travel time, the relative strength and the pulse shape of the wave-field constituents, and will show up most clearly in the directivity pattern and in the time domain signal.

Table I shows the diameter and the center frequency of the piezoelectric crystals used in the ultrasonic transducers. It also shows the effective aperture sizes l_e and w_e as well as the load speed c . The latter is calculated from

$$\arcsin\left(\frac{C^S}{c}\right) = \theta^B, \quad (8)$$

where θ^B is the direction of the main beam obtained from the measurements. These measured values of θ^B may differ a few degrees from the value given by the manufacturer. This is due to a difference in properties of the employed steel.

IV. COMPARISON OF MODELED AND MEASURED WAVE FIELDS

As a first way of comparing the modeled and measured wave fields, we look at the wave-field pattern. This is presented as a greyscale plot showing the magnitude, i.e., the absolute value, of the wave field versus time and observation angle. Each plot is normalized to the largest magnitude that would appear in the corresponding unnormalized plot. Depending on its construction, a practical transducer causes a specific internal time delay before the excitation of the steel

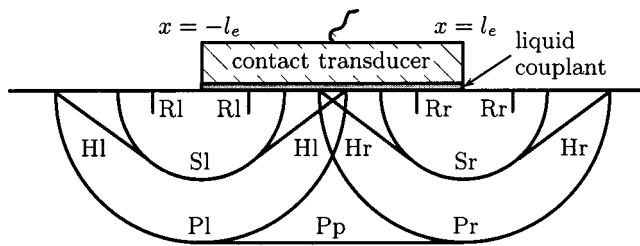


FIG. 4. The theoretical wave fronts in the symmetry plane $y=0$ for a contact transducer that is coupled to a steel surface through a thin layer of liquid couplant. In the symmetry plane, the wave fronts of the edge waves originate from the back edge $x=-l_e$ and the front edge $x=l_e$ of the circular aperture. The labels of the edge waves are explained in Table II. The dominant quasiplanar compression wave travels straight down from the transducer and is labeled as Pp.

surface takes place. To enable the comparison of the modeled and measured wave-field patterns, we synchronize both wave fields by removing this constant time delay from the measured data.

We look at the directivity of the modeled and measured wave fields as another way of comparison. For the directivity measurements we always use the T-EMAT since this yields the best SNR with the smallest crosstalk interference. Consequently, the directivity $D(\theta)$ of the particle velocity component v_θ is defined as

$$D(\theta) = \frac{\int_0^\infty v_\theta^2(\theta, t) dt}{\max \left\{ \int_0^\infty v_\theta^2(\theta, t) dt \right\}}. \quad (9)$$

The directivity is a measure of the energy transmitted in a direction θ , normalized by the maximum of this energy measure over all directions θ .

As a third means of comparison we look at the time-domain signals of the modeled and measured wave fields for a particular value of the observation angle. Each graph is normalized to the largest field value that would appear in the corresponding unnormalized graph.

All our measured results apply to a half-cylinder with a radius $R=50$ mm. We determine the source signature g^t from the measured signal in the main beam of the transducer at hand. In the model calculations we use an adaptive Gauss quadrature rule with a relative accuracy of 10^{-6} to evaluate the single integral in Eqs. (2) and (5). The time sampling rate in the model is 40 MHz. In view of this, the sampling rate of the measured source signature is artificially increased from 20 MHz to 40 MHz by spline interpolation. In the model we apply uniform traction amplitudes g^{ac} and g^{ap} unless noted otherwise.

In Fig. 4 we have depicted the wave fronts of all the waves that in theory show up with a contact transducer. In Table II we have explained the labels of the occurring edge waves. The same types of edge wave also show up in the wave field of a probe. In the symmetry plane $y=0$ the edge waves may be traced back to either the back edge of the aperture ($x=-l_e$) or the front edge of the aperture ($x=l_e$). The labels in Table II are used to indicate the important events in the numerical results.

TABLE II. Overview of the edge waves and their points of origin.

Wave label	Wave type	Point of origin
Pl	Compression wave	$x=-l_e$
Sl	Shear wave	$x=-l_e$
Hl	Head wave	$x=-l_e$
Rl	Rayleigh wave	$x=-l_e$
Pr	Compression wave	$x=l_e$
Sr	Shear wave	$x=l_e$
Hr	Head wave	$x=l_e$
Rr	Rayleigh wave	$x=l_e$

A. Compression wave contact transducer

1. Wave-field pattern

In Fig. 5(a) we show the wave-field pattern as measured by the T-EMAT, and in Fig. 5(b) we show the wave-field pattern as modeled for the particle velocity component v_θ . The complex field pattern for this relatively simple type of transducer has also been observed by Djelouah and Baboux,²¹ who compared their model results with laser interferometry measurements.

Around $9 \mu s$ the wave field is dominated by a focal spot due to constructive interference of a quasiplanar Pp wave and the edge waves Pl and Pr. The focal spot around $16 \mu s$ is less wide than the first one. The cause is that the wavelength is only half as large as for the compression waves. Moreover, the contact transducer does not generate a quasiplanar shear wave, so this second focal spot is due to constructive interference of only the edge waves Sl and Sr. The weaker waves that in Fig. 5(a) arrive directly after the focal spots result from multiple internal reflections inside the contact transducer.

Since the contact transducer is circular symmetrical, the measured wave-field pattern should be symmetrical with respect to $\theta=0^\circ$, just like the modeled wave field pattern. However, the lower right branch of the X-pattern of the waves Pl and Pr is clearly missing in the measurements. This may be caused by a reduced acoustical coupling at the rightmost edge of the transducer. A less clearly observable asymmetry is that the focal spots appear at -2° instead of 0° . This is caused by an angular offset of the machine that rotates the EMAT around the curved cylinder surface.

In the modeled wave field we detect two of each head waves Hl and Hr and two of each Rayleigh waves Rl and Rr. However, in the measured wave field we only observe the first arriving head waves and Rayleigh waves. The attenuation of the measured waves that arrive later is caused by the presence of the transducer on the steel surface, see Fig. 4, which perturbs the propagation of the surface waves. This attenuation does not occur in the model, which allows unperturbed wave propagation along the surface below the source domain. The measured Rayleigh waves can be identified by tracing the linear features, due to waves creeping along the curved cylinder surface, back to the edge of the plane surface at $\theta=\pm 90^\circ$.

2. Directivity

In Fig. 6(a) we present the measured and modeled directivity D of the velocity component v_θ ; see Eq. (9). While the

calculated directivity reaches values as low as -30 dB, the measured curve does not fall below -20 dB due to measurement noise. In the modeled curve there is a dip around 0° since here all the energy of the strong Pp wave is related to the velocity component v_r . The asymmetries in the measurements, as mentioned above, are apparent.

3. Time-domain signal

In Fig. 7(a) we give the measured and modeled time-domain signal of v_θ for an observation angle $\theta = -5^\circ$. At $8.5 \mu\text{s}$ a focused compression wave arrives, which is followed at $15.5 \mu\text{s}$ by a focused shear wave. The figure shows a good correspondence between the arrival times and the relative magnitudes of the measured and the modeled signals. For the small observation angle considered here, the measured and the modeled signal shapes correspond well. However, for large observation angles, the correspondence deteriorates for the weaker edge waves.

We suppose that the difference between the measured and the modeled signal shapes has three different causes.

First, the difference is sensitive to a number of experimental factors that are quite difficult to reproduce or control. The most important of these are the liquid coupling between the transducer and the steel surface and the alignment of the EMAT. Second, our model applies to a half-space and does therefore not account for the mode conversions in the experimental setup at the free, curved surface of the half-cylinder. Since the amplitude of the mode converted waves increases with an increasing angle of incidence, this mostly affects the signal shape of head waves, Rayleigh waves and edge waves. Third, in our source model it has been assumed that g^t is independent of the horizontal coordinates, i.e., it is assumed that the source signature is the same over the whole transducer aperture. In practice, the effective mechanical stiffness of the transducer will be lower near the edges of the aperture. Moreover, phase variations may occur over the piezoelectric crystal. Both effects locally influence the source signature, and manifest themselves mainly in the resulting signal shape of the edge waves.

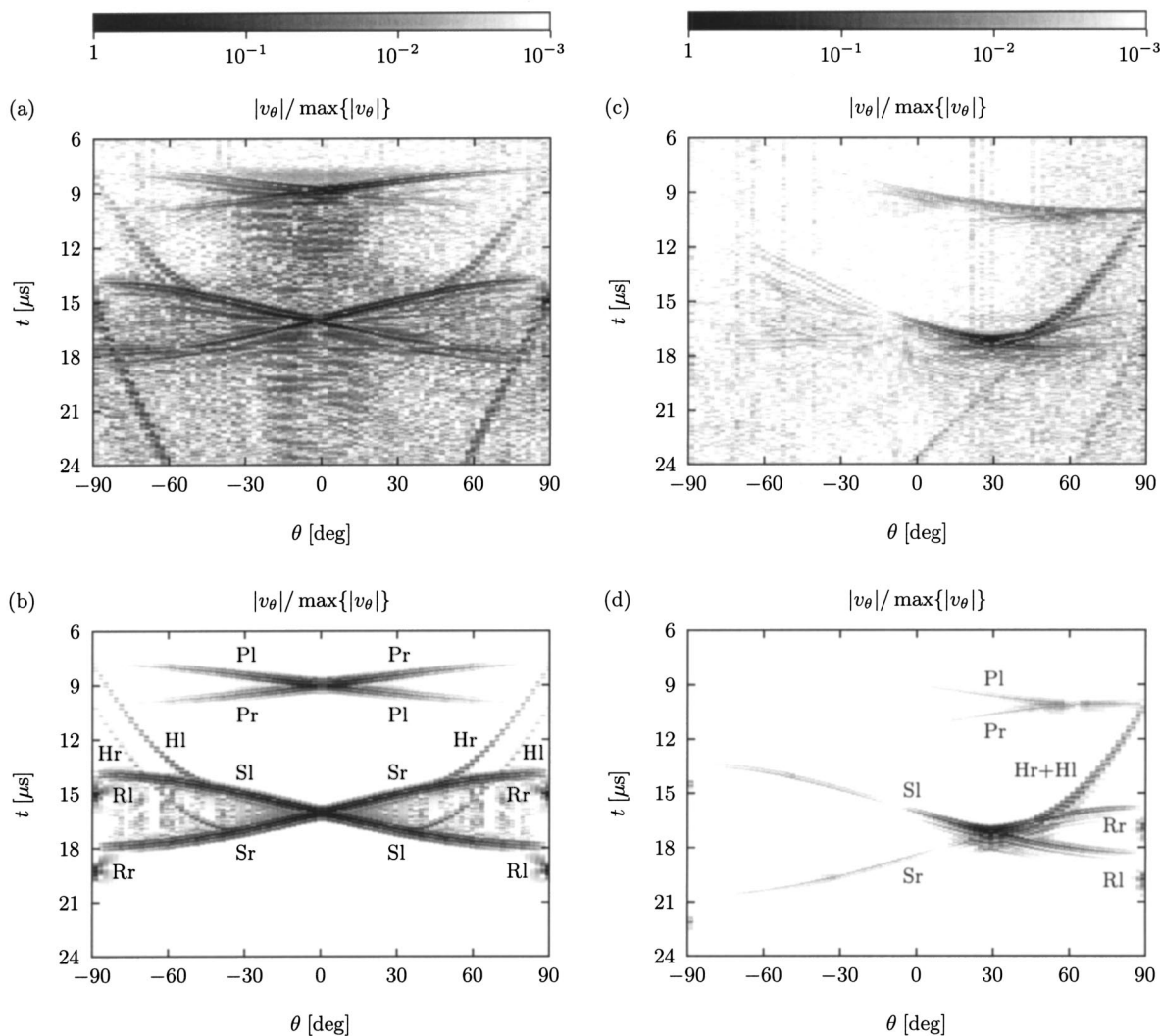


FIG. 5. Wave-field patterns in the half-plane $y=0$, $z \geq 0$, showing the normalized versions of (a) $|v_\theta|$ measured for the contact transducer; (b) $|v_\theta|$ modeled for the contact transducer; (c) $|v_\theta|$ measured for the 32° probe; (d) $|v_\theta|$ modeled for the 32° probe; (e) $|v_r|$ measured for the 45° probe; (f), (g) $|v_r|$ modeled for the 45° probe; (h) $|v_\theta|$ measured for the 70° probe; (i), (j) $|v_\theta|$ modeled for the 70° probe. For (g) and (j) the nonuniform traction amplitudes from Fig. 8 have been used.

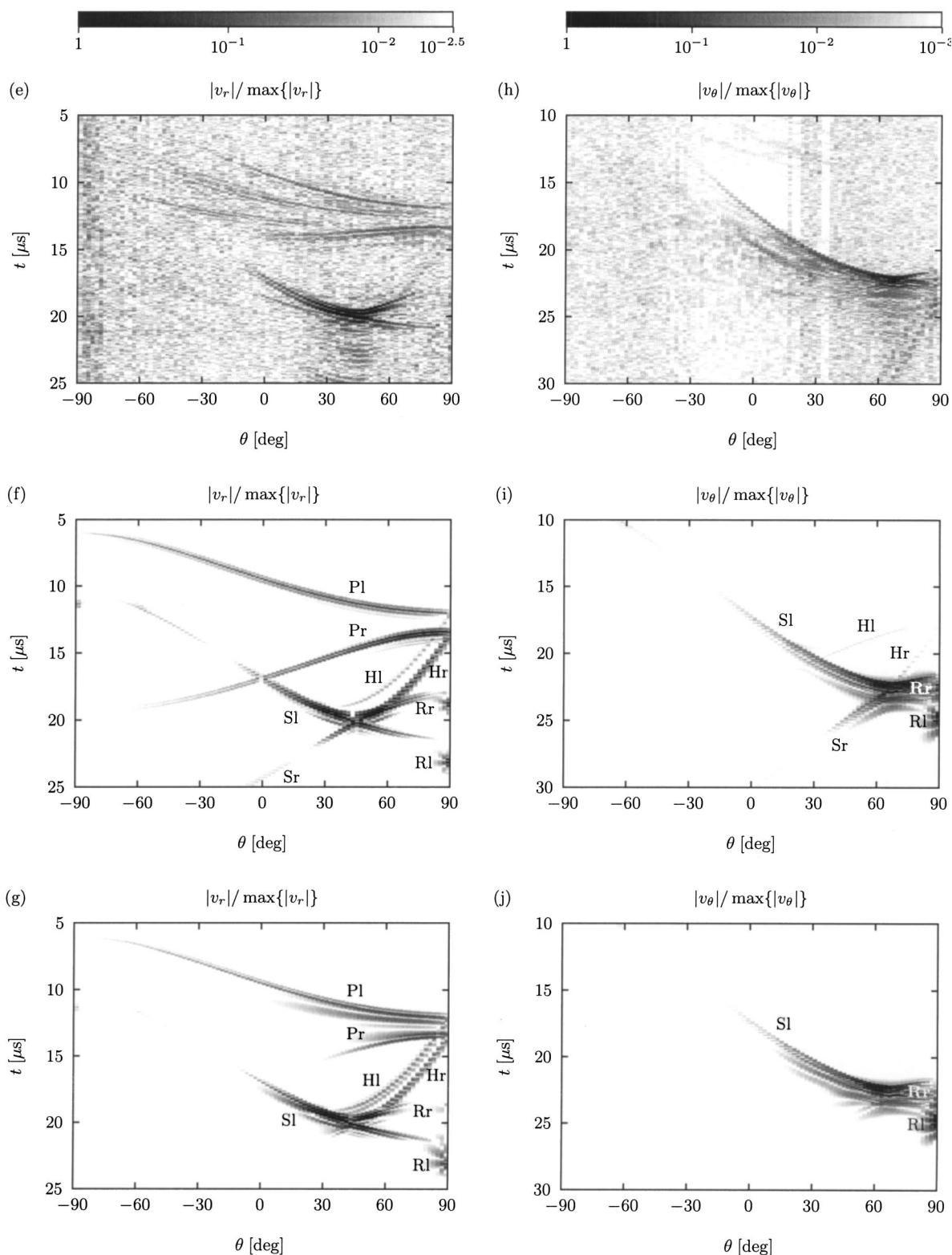


FIG. 5. (Continued.)

B. 32° shear wave probe

1. Wave-field pattern

We show the measured and modeled wave-field patterns for the T-EMAT and v_θ in Figs. 5(c) and (d), respectively. Around 18 μs the wave field is dominated by a focal spot due to constructive interference of a quasiplanar shear wave

Sp, which travels under 29°, and the edge waves Sl and Sr. There is also a significant focal spot around 10 μs due to constructive interference of a quasiplanar Pp wave, which travels under 63°, and the edge waves Pl and Pr. Upon considering that the relative contribution of the compression waves is even larger in v_r , we think that in the present case the name shear wave probe should not be taken too literally.

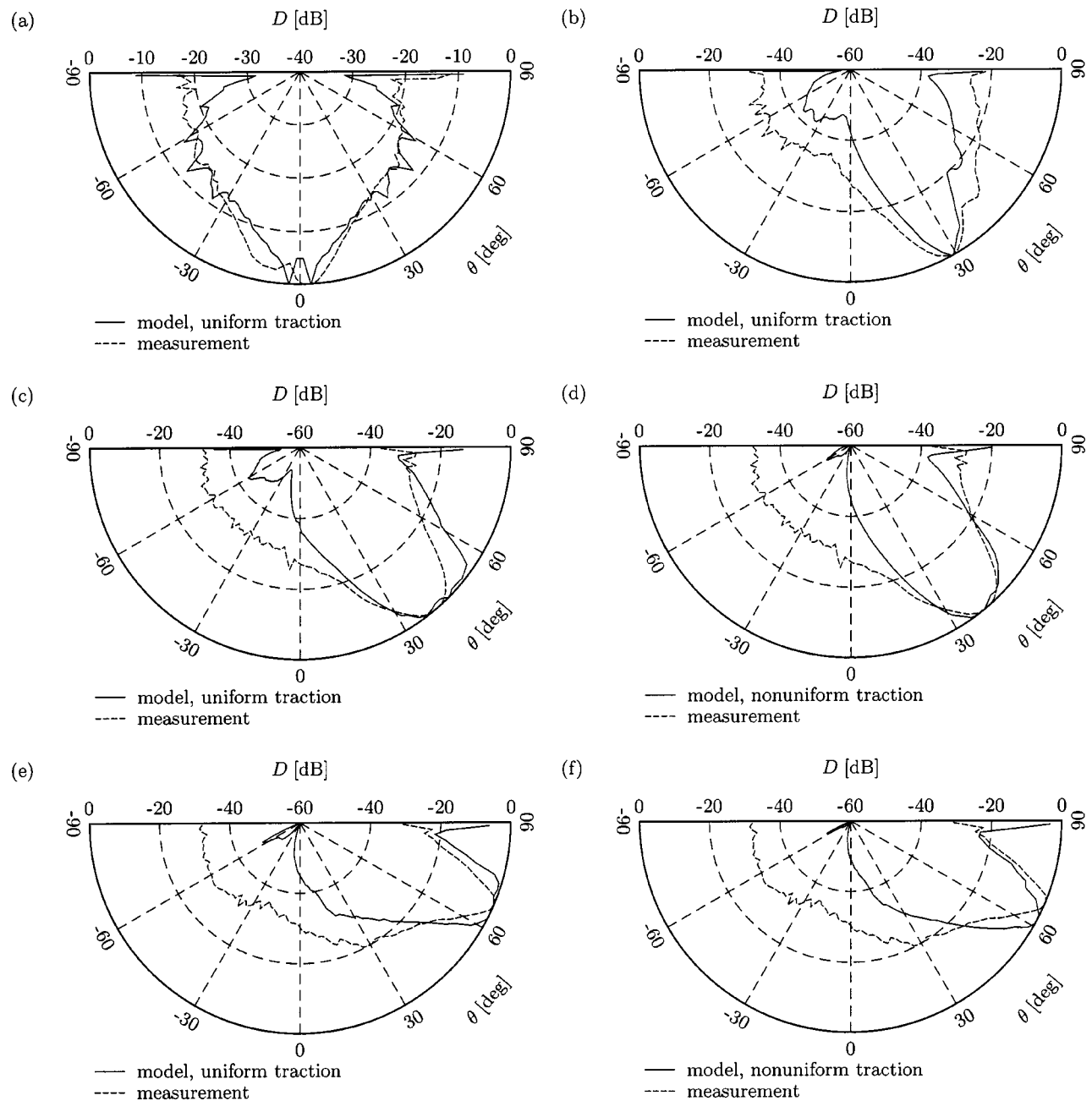


FIG. 6. Directivities in the half-plane $y=0$, $z \geq 0$, obtained from v_θ for (a) the contact transducer; (b) the 32° probe; (c), (d) the 45° probe; (e), (f) the 70° probe. For the modeled directivities in (d) and (f) the nonuniform traction amplitudes from Fig. 8 have been used.

We see that the last arriving Rayleigh wave Rl does not show up in the measurements. This is due to the presence of the probe on the steel surface. From the model results we conclude that the two observable head waves Hl and Hr coincide. This is as we expect, since the main beam direction $\theta^B = 29^\circ$ is close to the critical angle $\theta^C = \arcsin(C^S/C^P) = 33^\circ$ under which the head waves propagate in the steel, and c and C^P are of similar magnitude so the headwaves are nearly focused as well. Whether Hl is actually present in the measurements cannot be established from Fig. 5(c), but it is quite likely that it is also attenuated due to the presence of the probe on the steel interface. The linear feature occurring in the measurements after $18 \mu s$, between 0° and 45° , is due to a reflection inside the small plastic wedge of the probe.

The measured edge waves Pr, Sr, and Rr are relatively

weak. In between the measured edge waves Sl and Sr we observe some faint waves that do not appear in the modeled wave-field pattern of Fig. 5(d). We suppose that both differences are due to the assumption of a uniform traction amplitude g^{ap} in the model, as we will explain with the 45° probe.

2. Directivity

We present the measured and modeled directivity D of v_θ in Fig. 6(b). The calculated directivity reaches values as low as -60 dB, while the measured directivity does not fall below -35 dB due to measurement noise. In general, the SNR is better for the current probe than for the contact transducer since the first shows a more pronounced directionality than the latter. The measurements yield a wider main lobe

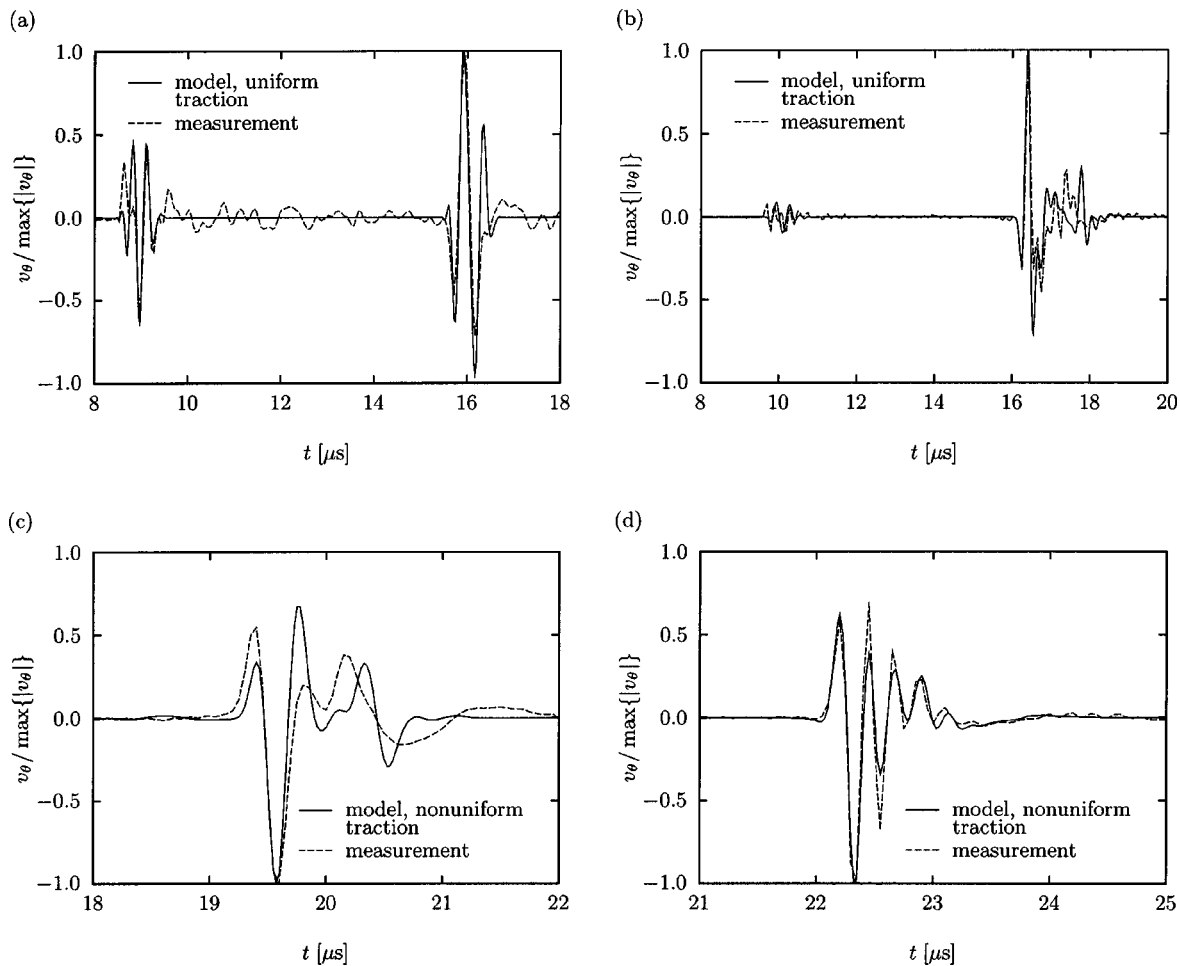


FIG. 7. Time-domain signals in the half-plane $y=0$, $z \geq 0$, showing the normalized versions of v_θ for (a) the contact transducer and $\theta = -5^\circ$; (b) the 32° probe and $\theta = 49^\circ$; (c) the 45° probe and $\theta = 49^\circ$; (d) the 70° probe and $\theta = 67^\circ$.

than the model. We suppose that this is due to the fact that the measured quasiplanar Sp wave is not as dominant with respect to the edge waves as we have calculated. This may happen when a misalignment causes the measurements to be taken outside the plane of symmetry of the probe.

3. Time-domain signal

In Fig. 7(b) we give the measured and modeled time-domain signal of v_θ for an observation angle $\theta = 49^\circ$. The edge waves Pl and Pr arrive first at $9.6 \mu\text{s}$. At $16 \mu\text{s}$ the strong head waves Hl and Hr arrive, which are followed within $2 \mu\text{s}$ by the edge waves Sr and Sl. As for the contact transducer, the correspondence between the arrival times and the relative magnitudes of the measured and the modeled signals is quite good. However, the signal shape is predicted satisfactorily only for the strong head wave. We suppose that, apart from the misalignment of the probe, these differences have the same causes as mentioned with the contact transducer.

C. 45° shear wave probe

1. Wave-field pattern

In Fig. 5(e) we show the wave field pattern as measured by the R-EMAT, and in Fig. 5(f) we show the wave field

pattern as modeled for the particle velocity component v_r . Only for these plots we have switched to the radial velocity component because the resulting measurements show also the edge P waves, which are indistinctive in the T-EMAT measurements. Around $20 \mu\text{s}$ the wave field is dominated by a focal spot due to constructive interference of the quasiplanar wave Sp, which travels under 43° , and the edge waves Sl and Sr. There is no focal spot for the compression waves. In comparison to the modeled wave field, the shear waves in the measurements are relatively stronger than the compression waves. We attribute this to the strong crosstalk of the R-EMAT, which favors the shear waves. Just as for the 32° probe, the presence of the probe on the steel interface severely attenuates the edge waves Hl and Rl.

The edge waves Pr, Sr, Hr, and Rr from the right-hand side of the probe are practically absent in our measurements. Second, as opposed to the modeled wave field, in the measured wave field we observe some waves to appear in between the edge waves Pl and Pr and in between the edge waves Sl and Sr. As we will now elucidate, we have strong indications that these differences may be resolved by using a nonuniform traction amplitude g^{ap} in our source model. In Fig. 8 we have depicted the normalized applied traction, measured with a small piezoelectric contact transducer along the length axis of the probe aperture. The curve shows that

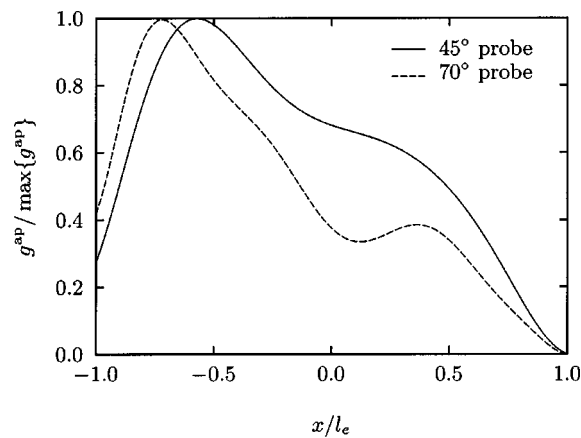


FIG. 8. Normalized measured traction amplitude for the 45° probe and the 70° probe.

the applied traction falls off at the edges, and is especially low at the right edge. This explains the attenuation of the corresponding edge waves.²⁸ On the other hand, the variation of the applied traction over the interval $[-l_e, l_e]$ explains the occurrence of the additional waves showing up in between the edge waves. To support our idea, we have re-evaluated the modeled wave field for the nonuniform g^{ap} shown in Fig. 8. The resulting wave field, shown in Fig. 5(g), indeed has a much better correspondence with the measured wave field than Fig. 5(f). From this we conclude that traction variations in the y direction, which cannot be facilitated by our model, have justifiably been neglected.

2. Directivity

In Fig. 6(c) we present the measured and modeled directivity D of the velocity component v_θ . While the calculated directivity reaches values as low as -60 dB, the measured curve does not fall below -30 dB due to measurement noise. The modeled main lobe is somewhat wider than the measured main lobe, which has its maximum at 43° . We think that this deviation is mainly caused by the use of a uniform traction amplitude g^{ap} in the model. This produces relatively strong edge waves, which widen the main lobe. In Fig. 6(d) we present the measured and the modeled directivity, where the latter has now been obtained with the nonuniform g^{ap} shown in Fig. 8. The width of the main lobe of the measured and the modeled directivity are now in good agreement. The enhanced directionality has increased the maximum signal strength in the modeled curve by 8 dB.

3. Time-domain signal

In Fig. 7(c) we give the measured and modeled time-domain signal of v_θ for an observation angle $\theta = 49^\circ$. We have used the nonuniform traction amplitude g^{ap} of Fig. 8 in the evaluation of the modeled signal. At $18.3 \mu\text{s}$ the head wave H1 arrives, which is followed at $19.2 \mu\text{s}$ and $20 \mu\text{s}$ by the edge waves Sr and Sl, respectively. As for the preceding transducers, the correspondence between the arrival times and the relative magnitudes of the measured and the modeled signals is quite good, while the signal shape of only the strongest edge wave is predicted satisfactorily.

D. 70° shear wave probe

1. Wave-field pattern

We show the measured and modeled wave-field patterns for the T-EMAT and v_θ in Figs. 5(h) and (i), respectively. Around $23 \mu\text{s}$ the wave field is dominated by a focal spot due to constructive interference of a quasiplanar wave Sp, which travels under 67° , and the edge waves Sl and Sr. There is no focal spot for the compression waves, and their magnitudes are more than 30 dB below the largest magnitude occurring in the shear wave focal spot. We observe similar phenomena as for the two preceding probes. The presence of the probe on the steel interface severely attenuates the edge waves Hl and Rl. The nonuniform normal traction amplitude g^{ap} of the actual probe causes a weakening of the measured edge waves Sr, Hr, and Rr as compared to their model counterparts obtained with a uniform g^{ap} . Moreover, the nonuniformity of g^{ap} of the actual probe causes some faint waves to appear in between the edge waves Sl and Sr. To prove the last two statements, we have re-evaluated the modeled wave field for the nonuniform g^{ap} shown in Fig. 8. The resulting wave field is presented in Fig. 5(j) and confirms the improved correspondence with the measurements.

2. Directivity

We present the measured and modeled directivity D of v_θ in Fig. 6(e). The calculated directivity reaches values as low as -60 dB, while the measured directivity does not fall below -30 dB due to measurement noise. As for the 45° probe, the modeled curve shows a wider main lobe than the measured curve because a uniform traction amplitude g^{ap} is used in the model. To prove this, we show the measured and the modeled directivity in Fig. 6(f), where the latter has now been obtained with the nonuniform g^{ap} in Fig. 8. The width of the main lobe of the measured and the modeled directivity are now in good agreement. The improved correspondence between the measured and the modeled main lobes is clearly visible. The enhanced directionality, which with a uniform g^{ap} is already better than for the 45° probe, has increased the maximum signal strength in the modeled curve by 4 dB. We suppose that the modeled edge waves from the left-hand side of the source domain are still too strong, since the angle $\theta^{\text{B}} = 63^\circ$ of the modeled main beam is 4° too low as compared with the measured main beam.

3. Time-domain signal

In Fig. 7(d) we give the measured and modeled time-domain signal of v_θ for an observation angle $\theta = 67^\circ$. We have used the nonuniform traction amplitude g^{ap} of Fig. 8 in the evaluation of the modeled signal. At $21.9 \mu\text{s}$ a focused shear wave arrives, which is the dominating feature of the signal. The measured signal looks a bit triangular due to the moderate sampling frequency in relation to the frequencies in the focused wave. As for the preceding transducers, the correspondence between the arrival times and the relative magnitudes of the measured and the modeled signals is quite good, while the signal shape of only the strongest edge wave is predicted satisfactorily.

V. CONCLUSIONS

We have presented source models for the compression wave contact transducer and the shear wave angle beam transducer (probe). Both models are based on an equivalent line source distribution of normal traction over the source aperture. The elastodynamic space-time domain Green's function for a uniform line source on a half-space has been obtained in closed form. In view of this, these models have enabled us to efficiently and accurately simulate the three-dimensional, time-domain elastodynamic wave field that is generated in a steel half-space by a given ultrasonic transducer.

The model calculations have been verified against measurements for one type of contact transducer and three types of probe. Validations based on the wave-field pattern have shown that the source models yield accurate simulations as far as the arrival times and the relative magnitudes of the wave fields are concerned. With the aid of the modeled wave-field patterns, the focal spots, edge waves, Rayleigh waves and head waves in the measured wave-field patterns could be clearly recognized. Having identified all the waves predicted by the half-space model, it has become tractable to identify the nature and the origin of the disturbances caused by the half-cylinder used in the experimental setup. Validations based on the directivity and the time-domain signals have also shown that the source models perform satisfactorily. However, the shape of the signals can be difficult to validate due to experimental conditions.

The measurements show that in models for the probes the nonuniformity of the traction amplitude cannot be ignored. This nonuniformity results in the observed attenuation of the edge waves and, consequently, in an enhanced directivity.

A difference between the source model and an actual transducer is the influence imposed by the transducer on the surface related waves, such as head waves and Rayleigh waves. Our source models allow these waves to propagate undisturbedly along the source domain. In the experimental situation, however, the transducer-steel interface does not support an unperturbed propagation of these waves.

The models introduced in this paper may easily be extended to sources of shear traction, since a closed-form expression for the Green's function may also be obtained for a source of shear traction.

ACKNOWLEDGMENTS

We thank René Breeuwer from TNO TPD for his technical support and advice. Further we are grateful to TNO TPD for the use of their experimental equipment, and to the Dutch Technology Foundation (STW) for financially supporting this research.

APPENDIX: CLOSED FORM SOLUTION FOR A LINE SEGMENT SOURCE

Here we regard the situation where the surface of the elastic half-space is loaded by an impulsive line source of normal traction

$$T_z(x, y, t) = g^a \delta(t) \delta(x) [H(y + w) - H(y - w)], \quad (A1)$$

where g^a is the traction amplitude giving the strength of the line source. This line source generates a wave field $G_k(x, y, z, t)$ with the following symmetries:

$$G_x(x, y, z, t) = G_x(x, -y, z, t) = -G_x(-x, y, z, t), \quad (A2)$$

$$G_y(x, y, z, t) = G_y(-x, y, z, t) = -G_y(x, -y, z, t), \quad (A3)$$

$$G_z(x, y, z, t) = G_z(-x, y, z, t) = G_z(x, -y, z, t). \quad (A4)$$

In view of these symmetries, we will only consider observation points with $x \geq 0$ and $y \geq 0$. The following results have been obtained by Bakker²⁵ with the aid of the Cagniard-De Hoop method.²⁶

The exact wave field due to the line source may be written as

$$G_k = G_k^{P,+1} + G_k^{S,+1} + G_k^{P,-1} + G_k^{S,-1}, \quad (A5)$$

where the superscript P indicates a compression wave term and the superscript S indicates a shear wave term. In turn, each of these terms separates into a number of contributions

$$G_k^{P,m} = G_k^{Pc,m} + G_k^{Pwc,m}, \quad (A6)$$

$$G_k^{S,m} = G_k^{Sc,m} + G_k^{Sh,m} + G_k^{Swc,m} + G_k^{Sw,h,m}, \quad (A7)$$

with $m = \pm 1$. The two contributions with superscripts Pc and Sc are associated with body waves, which expand with hemispherical wave fronts. These contributions are of the form

$$G_k^{Pc,m} = \frac{g^a m}{2\pi^2 \mu} H(t - T^{Pc}) \int_0^{Q^{Pc}} \text{Im} \left[\frac{\mathcal{E}_k^P \mathcal{J}^{Pc}}{(p - p^w)(p + p^w) \mathcal{R}} \right] dq, \quad (A8)$$

$$G_k^{Sc,m} = \frac{g^a m}{2\pi^2 \mu} H(t - T^{Sc}) \int_0^{Q^{Sc}} \text{Im} \left[\frac{\mathcal{E}_k^S \mathcal{J}^{Sc}}{(p - p^w)(p + p^w) \mathcal{R}} \right] dq. \quad (A9)$$

The single contribution with the superscript Sh is associated with the head wave connecting the Pc and Sc wave fronts. This contribution is of the form

$$\begin{aligned} G_k^{Sh,m} = & \frac{g^a m}{2\pi^2 \mu} H(t - T^{Sh}) H(T^{Sc} - t) S^h \\ & \times \int_0^{Q^{Sh}} \text{Im} \left[\frac{\mathcal{E}_k^S \mathcal{J}^{Sh}}{(p - p^w)(p + p^w) \mathcal{R}} \right] dq \\ & + \frac{g^a m}{2\pi^2 \mu} H(t - T^{Sc}) H(T^{Shr} - t) S^h \\ & \times \int_{Q^{Sc}}^{Q^{Sh}} \text{Im} \left[\frac{\mathcal{E}_k^S \mathcal{J}^{Sh}}{(p - p^w)(p + p^w) \mathcal{R}} \right] dq. \end{aligned} \quad (A10)$$

The two contributions with superscripts Pw and Sw are associated with waves that expand with cylindrical wave fronts. These contributions are of the form

$$G_k^{Pwc,m} = -\frac{g^a m}{\pi \mu} H(t - T^{Pwc}) S^w \text{Im} \left[\frac{\mathcal{M}_k^P \mathcal{J}^{Pwc}}{\mathcal{R}^w} \right], \quad (A11)$$

$$G_k^{Swc,m} = -\frac{g^a m}{\pi \mu} H(t - T^{Swc}) S^w \text{Im} \left[\frac{\mathcal{M}_k^S \mathcal{J}^{Swc}}{\mathcal{R}^w} \right]. \quad (A12)$$

The single contribution with the superscript Swh is associated with the head wave connecting the Pw and Sw wave fronts. This contribution is of the form

$$G_k^{Swh,m} = -\frac{g^a m}{\pi \mu} H(t - T^{Swh}) H(T^{Swc} - t) S^w S^{wh} \times \text{Im} \left[\frac{\mathcal{M}_k^S \mathcal{J}^{Swh}}{\mathcal{R}^w} \right]. \quad (\text{A13})$$

In Eqs. (A8)–(A13), μ is one of the Lamé parameters of the elastic half-space. Further, the following arrival times show up:

$$T^{Pc} = \frac{R}{C^P}, \quad (\text{A14})$$

$$T^{Sc} = \frac{R}{C^S}, \quad (\text{A15})$$

$$T^{Sh} = za^{sp} + \frac{r}{C^P}, \quad (\text{A16})$$

$$T^{Pwc} = \frac{R^w}{C^P}, \quad (\text{A17})$$

$$T^{Swc} = \frac{R^w}{C^S}, \quad (\text{A18})$$

$$T^{Swh} = za^{sp} + \frac{x}{C^P}, \quad (\text{A19})$$

while in Eq. (A10) there appears the time

$$T^{Shr} = \frac{R^2 a^{asp}}{z}. \quad (\text{A20})$$

Together with the time parameters, there has been introduced the compression wave speed C^P , the shear wave speed C^S , the horizontal distance $r = \sqrt{x^2 + \Delta^2}$ with $\Delta = y + mw$, and the distances $R = \sqrt{r^2 + z^2}$ and $R^w = \sqrt{x^2 + z^2}$. Further, the parameter $a^{sp} = \sqrt{(C^S)^{-2} - (C^P)^{-2}}$ has been employed. In Eqs. (A10)–(A13), the following switches may turn the terms on or off:

$$S^{Sh} = 1 \quad \text{if} \quad \left(\frac{r}{R} \right)^2 - \left(\frac{C^S}{C^P} \right)^2 > 0 \quad \text{else} \quad 0, \quad (\text{A21})$$

$$S^w = 1 \quad \text{if} \quad \Delta > 0 \quad \text{else} \quad 0, \quad (\text{A22})$$

$$S^{wh} = 1 \quad \text{if} \quad \left(\frac{x}{R^w} \right)^2 - \left(\frac{C^S}{C^P} \right)^2 > 0 \quad \text{else} \quad 0. \quad (\text{A23})$$

The integration limits in Eqs. (A8)–(A10) are

$$Q^{Pc} = \sqrt{\left(\frac{t}{R} \right)^2 - \left(\frac{1}{C^P} \right)^2}, \quad (\text{A24})$$

$$Q^{Sc} = \sqrt{\left(\frac{t}{R} \right)^2 - \left(\frac{1}{C^S} \right)^2}, \quad (\text{A25})$$

$$Q^{Sh} = \sqrt{\left(\frac{t - za^{sp}}{r} \right)^2 - \left(\frac{1}{C^P} \right)^2}. \quad (\text{A26})$$

The factors in the numerators of Eqs. (A8)–(A13) are

$$\mathcal{E}_x^P = -\frac{2x}{\Delta} (p^2 - q^2) \left[\left(\frac{1}{C^S} \right)^2 - 2p^2 + 2q^2 \right], \quad (\text{A27})$$

$$\mathcal{E}_y^P = -2(p - p^w)(p + p^w) \left[\left(\frac{1}{C^S} \right)^2 - 2p^2 + 2q^2 \right], \quad (\text{A28})$$

$$\mathcal{E}_z^P = -\frac{2r}{\Delta} \eta p \left[\left(\frac{1}{C^S} \right)^2 - 2p^2 + 2q^2 \right], \quad (\text{A29})$$

$$\mathcal{E}_x^S = \frac{4x}{\Delta} \eta \zeta (p^2 - q^2), \quad (\text{A30})$$

$$\mathcal{E}_y^S = 4 \eta \zeta (p - p^w)(p + p^w), \quad (\text{A31})$$

$$\mathcal{E}_z^S = -\frac{4r}{\Delta} \eta p (p^2 - q^2), \quad (\text{A32})$$

$$\mathcal{M}_x^P = -\bar{q} \left[\left(\frac{1}{C^S} \right)^2 - 2\bar{q}^2 \right], \quad (\text{A33})$$

$$\mathcal{M}_y^P = 0, \quad (\text{A34})$$

$$\mathcal{M}_z^P = -\bar{\eta} \left[\left(\frac{1}{C^S} \right)^2 - 2\bar{q}^2 \right], \quad (\text{A35})$$

$$\mathcal{M}_x^S = 2 \bar{\eta} \bar{\zeta} \bar{q}, \quad (\text{A36})$$

$$\mathcal{M}_y^S = 0, \quad (\text{A37})$$

$$\mathcal{M}_z^S = -2 \bar{\eta} \bar{q}^2, \quad (\text{A38})$$

$$\mathcal{J}^{Pc} = \frac{i \eta}{\sqrt{t^2 - (R/C^P)^2 - R^2 q^2}}, \quad (\text{A39})$$

$$\mathcal{J}^{Sc} = \frac{i \zeta}{\sqrt{t^2 - (R/C^S)^2 - R^2 q^2}}, \quad (\text{A40})$$

$$\mathcal{J}^{Sh} = \frac{\zeta}{\sqrt{(R/C^S)^2 + R^2 q^2 - t^2}}, \quad (\text{A41})$$

$$\mathcal{J}^{Pwc} = \frac{i \eta}{\sqrt{t^2 - (T^{Pwc})^2}}, \quad (\text{A42})$$

$$\mathcal{J}^{Swc} = \frac{i \zeta}{\sqrt{t^2 - (T^{Swc})^2}}, \quad (\text{A43})$$

$$\mathcal{J}^{Swh} = \frac{\zeta}{\sqrt{(T^{Swc})^2 - t^2}}. \quad (\text{A44})$$

In the denominators of Eqs. (A8)–(A13) there appear the Rayleigh factors \mathcal{R} ,

$$\mathcal{R} = \left[\left(\frac{1}{C^S} \right)^2 - 2p^2 + 2q^2 \right]^2 + 4 \eta \zeta (p^2 - q^2), \quad (\text{A45})$$

$$\mathcal{R}^w = \left[\left(\frac{1}{C^S} \right)^2 - 2\bar{q}^2 \right]^2 + 4 \bar{\eta} \bar{\zeta} \bar{q}^2. \quad (\text{A46})$$

Moreover, in the denominators of Eqs. (A8)–(A10) the parameter $p^w = iqx/\Delta$ shows up. In the factors \mathcal{E} , \mathcal{M} , \mathcal{J} , and \mathcal{R} the propagation coefficients

$$\eta = \sqrt{\left(\frac{1}{C^p}\right)^2 + q^2 - p^2}, \quad (\text{A47})$$

$$\zeta = \sqrt{\left(\frac{1}{C^s}\right)^2 + q^2 - p^2}, \quad (\text{A48})$$

$$\bar{\eta} = \sqrt{\left(\frac{1}{C^p}\right)^2 - \bar{q}^2}, \quad (\text{A49})$$

$$\bar{\zeta} = \sqrt{\left(\frac{1}{C^s}\right)^2 - \bar{q}^2}, \quad (\text{A50})$$

are employed. The following parameters p and \bar{q} apply in the indicated cases

$$p = \frac{rt}{R^2} + \frac{iz}{R^2} \sqrt{t^2 - \left(\frac{R}{C^p}\right)^2 - R^2 q^2} \quad (\text{Pc term}), \quad (\text{A51})$$

$$p = \frac{rt}{R^2} + \frac{iz}{R^2} \sqrt{t^2 - \left(\frac{R}{C^s}\right)^2 - R^2 q^2} \quad (\text{Sc term}), \quad (\text{A52})$$

$$p = \frac{rt}{R^2} + \frac{iz}{R^2} \sqrt{\left(\frac{R}{C^s}\right)^2 + R^2 q^2 - t^2} \quad (\text{Sh term}), \quad (\text{A53})$$

$$\bar{q} = \frac{xt}{(R^w)^2} + \frac{iz}{(R^w)^2} \sqrt{t^2 - (T^{\text{Pwc}})^2} \quad (\text{Pwc term}), \quad (\text{A54})$$

$$\bar{q} = \frac{xt}{(R^w)^2} + \frac{iz}{(R^w)^2} \sqrt{t^2 - (T^{\text{Swc}})^2} \quad (\text{Swc term}), \quad (\text{A55})$$

$$\bar{q} = \frac{xt}{(R^w)^2} + \frac{iz}{(R^w)^2} \sqrt{(T^{\text{Swc}})^2 - t^2} \quad (\text{Swh term}). \quad (\text{A56})$$

The integrals in Eqs. (A8)–(A10) contain an integrable singularity and have limits that depend on time. These aspects complicate the numerical evaluation of the integrals. Both numerical obstacles are removed by the respective substitutions

$$q = Q^{\text{Pc}} \sin(\theta), \quad (\text{A57})$$

$$q = Q^{\text{Sc}} \sin(\theta), \quad (\text{A58})$$

$$q = \begin{cases} Q^{\text{Sh}} \sin(\theta), & t < T^{\text{Sc}}, \\ Q^{\text{Sc}} \cosh(\theta), & t \geq T^{\text{Sc}}. \end{cases} \quad (\text{A59})$$

¹A. J. Krautkrämer and H. Krautkrämer, *Ultrasonic Testing of Materials* (Springer, Berlin, 1990), pp. 145–150, 196–202.

²H. Wustenberg, Ph.D. thesis, Technical University of Berlin, 1972, pp. 38–48.

³T. P. Lerch, L. W. Schmerr, and A. Sedov, “The wavefield of an ultrasonic angle beam shear wave transducer: an elastodynamic approach,” in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1997), Vol. 16, pp. 855–892.

⁴D. L. Lansing, “The displacements in an elastic halfspace due to a moving concentrated normal load,” NASA technical report 238, 1966.

⁵D. C. Gakenheimer and J. Miklowitz, “Transient excitation of an elastic half space by a point load traveling on the surface,” *J. Appl. Mech.* **36**, 505–515 (1969).

⁶M. C. M. Bakker, M. D. Verweij, B. J. Kooij, and H. A. Dieterman, “The traveling point load revisited,” *Wave Motion* **29**, 119–135 (1999).

⁷K. J. Langenberg, U. Aulenbacher, G. Bollig, P. Fellingner, H. Morbitzer, G. Weinfurter, P. Zanger, and V. Schmitz, “Numerical modelling of ultrasonic scattering,” in *Mathematical Modelling in Non-destructive Testing*, edited by M. Blakemore and G. A. Georgiou (Clarendon, Oxford, 1988), pp. 132–138.

⁸U. Aulenbacher and K. J. Langenberg, “Transient radiation pattern of ultrasonic transducers on elastic half-spaces,” in *New Procedures in Non-destructive Testing*, edited by P. Holler (Springer, Berlin, 1983), pp. 81–92.

⁹H. J. Stam, “The two-dimensional elastodynamic distributed surface load problem,” *Geophysics* **55**, 1047–1056 (1990).

¹⁰L. W. Schmerr, Jr. and A. Sedov, “Elastodynamic model for compressional and shear wave transducers,” *J. Acoust. Soc. Am.* **86**, 1988–1999 (1989).

¹¹M. C. M. Bakker and M. D. Verweij, “The approximate far field and directivity of elastic wave transducers,” *J. Acoust. Soc. Am.* **111**, 1177–1188 (2002).

¹²D. Gridin and L. J. Fradkin, “The high-frequency asymptotic description of pulses radiated by a circular normal transducer into an elastic half-space,” *J. Acoust. Soc. Am.* **104**, 3190–3198 (1998).

¹³D. Gridin, “The radiating near field of a circular normal transducer of arbitrary apodization on an elastic half-space,” *J. Acoust. Soc. Am.* **106**, 1237–1246 (1999).

¹⁴T. Xue, W. Lord, and S. Upda, “Numerical analysis of the radiated fields of ultrasonic transducers,” *J. Nondestruct. Eval.* **14**, 137–147 (1995).

¹⁵A. Ilan and J. P. Weight, “The propagation of short pulses of ultrasound from a circular source coupled to an isotropic solid,” *J. Acoust. Soc. Am.* **88**, 1142–1151 (1990).

¹⁶A. Lhémy, “A model for the transient ultrasonic field radiated by an arbitrary loading in a solid,” *J. Acoust. Soc. Am.* **96**, 3776–3786 (1994).

¹⁷D. Gridin, “On the radiation of ultrasound into an isotropic elastic half-space via wavefront expansions of the impulse response,” *J. Acoust. Soc. Am.* **105**, 2565–2573 (1999).

¹⁸A. McNab, A. Cochran, and M. A. Campbell, “The calculation of acoustic fields in solids for transient normal surface force sources of arbitrary geometry and apodization,” *J. Acoust. Soc. Am.* **87**, 1455–1465 (1990).

¹⁹F. G. Laturelle, “The stresses produced in an elastic half-space by a normal step loading over a circular area, analytical and numerical results,” *Wave Motion* **12**, 107–127 (1990).

²⁰L. F. Bresse and D. A. Hutchins, “Transient generation of elastic waves in solids by a disk-shaped normal force source,” *J. Acoust. Soc. Am.* **86**, 810–817 (1989).

²¹H. Djelouah and J. C. Baboux, “Transient ultrasonic field radiated by a circular transducer in a solid medium,” *J. Acoust. Soc. Am.* **92**, 2932–2941 (1992).

²²J. C. Baboux and R. Kažys, “Analysis of the transient fields radiated in solids by circular and annular sources,” *J. Acoust. Soc. Am.* **92**, 2942–2951 (1992).

²³F. R. Norwood, “Exact transient response of an elastic half space loaded over a rectangular region of its surface,” *J. Appl. Mech.* **36**, 516–522 (1969).

²⁴F. Guan and M. Novak, “Transient response of an elastic homogeneous half-space to suddenly applied rectangular loading,” *J. Appl. Mech.* **61**, 256–263 (1994).

²⁵M. C. M. Bakker, Ph.D. thesis, Delft University of Technology, 2000, pp. 95–125.

²⁶A. T. de Hoop, “A modification of Cagniard’s method for solving seismic pulse problems,” *Appl. Sci. Res., Sect. B* **8**, 349–356 (1960).

²⁷I. A. Viktorov, *Rayleigh and Lamb Waves, Physical Theory and Applications* (Plenum, New York, 1967), pp. 29–46.

²⁸R. Stacey and J. P. Weight, “Pulse-echo scattering in solids with nonuniform transducers,” *IEE Proc.: Sci., Meas. Technol.* **141**, 363–368 (1994).

Modeling of finite amplitude acoustic waves in closed cavities using the Galerkin method

Robert R. Erickson and Ben T. Zinn

School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0150

(Received 4 March 2002; revised 27 December 2002; accepted 17 January 2003)

Nonlinear resonant gas oscillations in closed ducts are investigated by solving a previously derived, quasi-one-dimensional, nonlinear wave equation that accounts for forcing, gas dynamic nonlinearities, and viscous dissipation. This equation is solved with the approximate Galerkin method to determine the dependence of driven oscillations upon the duct shape, forcing frequency, and forcing amplitude. Initially, the applicability of the developed Galerkin solution approach was studied by investigating oscillations in a straight duct, closed at both ends and periodically oscillated at a single frequency. It is shown that the Galerkin method predictions of shock wave-like oscillations in such ducts are in excellent agreement with results obtained with other numerical solution techniques. Next, this study investigated the forced response of a class of horn-shaped ducts, and it is shown that for a given forcing amplitude, there exists a nonmonotonic increase in compression ratio as the duct's flare constant is increased. Finally, it is shown that oscillations driven in ducts whose shapes were chosen to provide shifting of the second and third natural acoustic mode frequencies exhibit significant waveform distortion and non-negligible increases in compression ratio when compared with oscillations driven in straight ducts. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1559592]

PACS numbers: 43.25.Cb, 43.25.Gf, 43.25.Ts [MFH]

I. INTRODUCTION

The generation of finite amplitude pressure oscillations within ducts has many practical applications in fields ranging from acoustic compression¹ to thermoacoustics² to propulsion.³ Since the performance of these applications generally improves as the amplitude of the excited oscillations increases, there has been an ongoing search for new approaches to increase the amplitude of driven oscillations. Significantly, some of these studies^{1,4,5} have shown that the amplitude of driven oscillations can be increased using ducts with axially varying cross sectional areas (i.e., "shaped ducts") as opposed to constant area ducts (i.e., "straight ducts"). This finding suggests that the performance of wave engines,³ which are propelled by large amplitude pressure oscillations, could be improved by employing shaped ducts. For a preliminary investigation of the use of shaped ducts in wave engines (and to find an optimal configuration), a solution approach is needed that can be used in a parametric study with low expenditures of time and computational resources. In the search for an appropriate solution technique, an application of the approximate Galerkin method, which is a specific example of the method of weighted residuals,⁶ was chosen to analyze finite amplitude oscillations in shaped and straight ducts. This paper describes the results of this study.

To date, many theoretical, numerical, and experimental studies have investigated the driving of finite amplitude oscillations in straight ducts.^{1,7-10} Most of these studies examined straight ducts with a rigid wall at one end and a movable piston at the opposite end. Oscillating the piston at or near the natural acoustic mode frequency of the duct resulted in the generation of finite amplitude oscillations within the duct. These early studies have generally shown that forcing at the frequency of maximum response produces steep-

fronted, "saw-tooth"-shaped waves within the duct. More recently, the response of shaped ducts forced at and near resonance was experimentally investigated by Lawrenson *et al.*¹ These researchers measured the forced response of several different shaped ducts that were closed at both ends. In these experiments, a linear motor periodically oscillated the entire duct along its axis. This mode of forcing eliminates seal losses and temperature gradients (due to heat generation by friction) at the oscillating piston.¹ Furthermore, oscillating the entire duct effectively produces forcing at both ends of the duct, in contrast to piston forcing at only one end. This study concluded that for a given power input, pressure oscillations driven in shaped ducts can have significantly larger amplitude at certain spatial locations than in straight ducts. Furthermore, the waveforms of oscillations driven in the shaped ducts were generally more continuous than the sawtooth waveforms excited in straight ducts.

Finite amplitude oscillations in straight and shaped ducts were also theoretically studied by Il'inskiy *et al.*⁴ They investigated this problem by deriving and solving a nonlinear wave equation expressed in terms of velocity potential. This model equation accounted for all the dominant physical processes of the problem; i.e., linear and nonlinear wave processes, forcing, and dissipation. The solution technique for this nonlinear wave equation assumed periodic time dependence and used a truncated Fourier series expansion to reduce the equation to a two-point boundary value problem. This technique obtained approximate solutions for limit cycle behavior of the driven oscillations, and the numerical calculations agreed well with experimental results.¹

Furthermore, several other approaches using finite difference techniques have examined large amplitude, one-dimensional oscillations in ducts.^{5,11} Vanhille and Campos-

Pozuelo¹¹ derived a fully nonlinear wave equation for the velocity in a straight duct forced by an oscillating piston. This nonlinear wave equation was formulated from the conservation laws and the Tait–Kirkwood equation of state. This equation was solved with a multi-time step, implicit, six-point scheme starting from an initial quiescent flow condition using the Gauss method. The results of this study compare favorably with experimental results from previous studies.¹⁰ Additionally, a numerical investigation by Chun and Kim⁵ of straight and shaped ducts examined the one-dimensional conservation equations of mass, momentum, and energy directly, therefore solving a set of partial differential equations in both space and time. The numerical approach used fourth-order spatial and temporal differentiation schemes and is applicable to all duct shapes within the constraints of the one-dimensional assumption. Results presented in this study show qualitative and quantitative agreement with previously published results.

The decision to apply the Galerkin method in this study was primarily based on the successful application of this method in the investigation of a variety of nonlinear problems.⁶ These studies have shown that the Galerkin method can provide approximate solutions to complex nonlinear equations with relatively low computational cost. This is accomplished by transforming the original nonlinear partial differential equations that describe the problem into systems of ordinary differential equations. The transformation of a partial differential equation into a system of ordinary differential equations has two distinct advantages: it simplifies the numerical treatment and it provides a useful framework for understanding the physics of the problem. Furthermore, the Galerkin method assumes no periodicity in time, thus yielding the complete transient behavior of the problem starting from some initial disturbance and terminating when the solution exhibits limit cycle behavior, if such a behavior exists. For example, there may exist some duct shape configurations and forcing that leads to irregular, chaotic oscillations.

The study described in this paper consists of several tasks. The first task investigates whether the approximate Galerkin method applied to a one-dimensional problem can accurately determine the characteristics of forced, finite amplitude, oscillations in closed ducts. This is accomplished by comparing the Galerkin method solutions with previously published results. Upon demonstration of the applicability of the Galerkin method, the method is then used to investigate the driving of finite amplitude oscillations in several different duct geometries, including straight ducts, horn-shaped ducts, and a specific class of sinusoidally shaped ducts. These ducts, as described by Hamilton *et al.*,¹² theoretically allow for “control” of the magnitude and spacing of individual natural acoustic frequencies of the duct. For this class of duct shapes, the Galerkin technique was used to examine the impact of individual natural acoustic frequency shifts on driven oscillation properties including waveform, amplitude, and overall compression ratio (P_{\max}/P_{\min}).

II. MODEL DEVELOPMENT

In this study, the nonlinear wave equation derived by Il'inskii *et al.*⁴ was used to model oscillations driven in several different duct shapes. This wave equation describes quasi-one-dimensional, finite amplitude oscillations of a calorically perfect gas while accounting for linear and second-order nonlinear processes, duct forcing, and dissipation. Third-order nonlinearities have been neglected in this study to simplify the model, and it will be shown that this has little effect upon the results of interest. Manipulating the mass, momentum, and state equations, Il'inskii derived the following normalized nonlinear wave equation:

$$\begin{aligned} \frac{\partial^2 \Phi}{\partial T^2} - \frac{1}{\Omega^2 \pi^2} \left\{ \frac{1}{S} \frac{dS}{dX} \right\} \frac{\partial \Phi}{\partial X} - \frac{1}{\Omega^2 \pi^2} \frac{\partial^2 \Phi}{\partial X^2} \\ = - \frac{2}{\Omega} \frac{\partial^2 \Phi}{\partial X \partial T} \frac{\partial \Phi}{\partial X} - \frac{(\gamma-1)}{\Omega} \frac{\partial \Phi}{\partial T} \frac{\partial^2 \Phi}{\partial X^2} \\ - \frac{(\gamma-1)}{\Omega} \left\{ \frac{1}{S} \frac{dS}{dX} \right\} \frac{\partial \Phi}{\partial T} \frac{\partial \Phi}{\partial X} - \frac{\Gamma(X,T)}{\Omega^2} + \frac{\xi(X,T)}{\Omega^2}, \quad (1) \end{aligned}$$

where the normalized velocity potential Φ is defined by

$$\frac{\partial \Phi}{\partial x} = \frac{u}{\ell^2 \omega_0}, \quad (2)$$

and X , T , γ , and S are the nondimensional axial distance, time, ratio of specific heats, and cross-sectional area of the duct, respectively. In addition, u , ℓ , and ω_0 are the dimensional axial velocity, duct length, and fundamental natural frequency of a straight duct, respectively. Finally, Eq. (1) also includes the following normalized frequency parameter Ω :

$$\Omega = \frac{\omega}{\omega_0}, \quad (3)$$

where ω is the forcing frequency.

The left side of Eq. (1) is the linear wave operator while the right side describes nonlinear effects (i.e., the first three terms), external forcing (terms involving Γ), and dissipation due to viscous processes (terms involving ξ). The forcing term Γ is described by the following expression:

$$\begin{aligned} \Gamma(X,T) = F \frac{\partial \Phi}{\partial X} + \Omega \frac{dF}{dT} X + (\gamma-1) \left\{ \frac{1}{S} \frac{dS}{dX} \right\} \frac{\partial \Phi}{\partial X} FX \\ + (\gamma-1) F \frac{\partial^2 \Phi}{\partial X^2} X, \quad (4) \end{aligned}$$

where F describes the normalized time-dependent forcing function that accelerates the entire duct in the axial direction. It is important to note that Eq. (1) describes the driven duct oscillations in a coordinate system that is attached to the duct. As will be seen later, the use of this frame of reference simplifies the application of the boundary conditions in the Galerkin method. Finally, the nondimensional, viscous dissipation ξ in Eq. (1) has the form

$$\xi(X,T) = \frac{D\Omega}{\pi^3 S} \frac{\partial^2}{\partial T \partial X} \left(S \frac{\partial \Phi}{\partial X} \right), \quad (5)$$

where D is a dimensionless dissipation parameter whose magnitude is determined by experimental results.¹

A. Application of the Galerkin method

The application of this solution technique begins with the assumption that the velocity potential can be expressed by the following infinite series expansion:

$$\Phi(X, T) = \sum_{n=1}^{\infty} \eta_n(T) \Psi_n(X), \quad (6)$$

where $\Psi_n(X)$, referred to as trial functions, describe the spatial dependence of the natural acoustic mode shapes of a given duct geometry (to be discussed in detail later), and $\eta_n(T)$ are the unknown, time-dependent amplitudes of these acoustic modes. Since in practice Eq. (6) is truncated to a finite number of modes, the resulting solution for Φ is approximate and does not satisfy Eq. (1) exactly. Specifically, when the truncated series for Φ is substituted into Eq. (1), the resulting expression, called the residual $[R(\Phi)]$, is non-zero. The objective of the Galerkin method is to derive a set of equations for the time-dependent amplitudes, $\eta_n(T)$, that when solved will minimize the residual, $R(\Phi)$.

To derive the system of equations for the amplitude functions $\eta_n(T)$, the truncated series in Eq. (6) is first substituted into Eq. (1). The next step in the solution approach takes advantage of the property of orthogonality, which states that a continuous function is zero if it is orthogonal (i.e., inner product equal to zero) to every member of a complete set.⁶ This orthogonality property leads to the formation of the following expression:

$$\int_0^1 R(\Phi(X, T)) \cdot W_j(X) dX = 0, \quad (7)$$

where the weighting functions, $W_j(X)$, must consist of a complete set of functions. In this study, the weighting function is determined based on the duct shape and the trial functions, and will be discussed in detail later. Furthermore, the use of natural acoustic mode shapes as trial functions also satisfies the boundary conditions exactly, thus eliminating any residuals at the boundary. Thus, once the set $W_j(X)$ is chosen, performing the integration over the duct length in Eq. (7) yields the following system of time-dependent coupled ordinary differential equations:

$$\begin{aligned} \sum_n G_{1n} \ddot{\eta}_1 - \sum_n H_{1n} \eta_n + \sum_n \sum_m (Z_{1nm} \dot{\eta}_n \eta_m) \\ + \sum_n E_{1n}(T) - L_{1n}(T) = 0, \\ \sum_n G_{2n} \ddot{\eta}_2 - \sum_n H_{2n} \eta_n + \sum_n \sum_m (Z_{2nm} \dot{\eta}_n \eta_m) \\ + \sum_n E_{2n}(T) - L_{2n}(T) = 0, \\ \vdots \end{aligned} \quad (8)$$

$$\begin{aligned} \sum_n G_{jn} \ddot{\eta}_j - \sum_n H_{jn} \eta_n + \sum_n \sum_m (Z_{jnm} \dot{\eta}_n \eta_m) \\ + \sum_n E_{jn}(T) - L_{jn}(T) = 0, \end{aligned}$$

where the subscripts j , n , and m all go from 1 to the specified number of included modes. G_{jn} and H_{jn} describe linear processes, Z_{jnm} describes nonlinear coupling, L_{jn} describes viscous damping, and E_{jn} describes forcing. These quantities are described by the following expressions:

$$G_{jn} = \int_0^1 W_j(X) \Psi_n dX, \quad (9)$$

$$H_{jn} = \int_0^L W_j(X) \left[\frac{d^2 \Psi_n}{dX^2} + \left\{ \frac{1}{S} \frac{dS}{dX} \right\} \frac{d\Psi_n}{dX} \right] dX, \quad (10)$$

$$Z_{jnm} = \frac{A_{jnm} + B_{jnm} + C_{jnm}}{\Omega}, \quad (11)$$

$$A_{jnm} = (\gamma - 1) \int_0^1 W_j(X) \left\{ \frac{1}{S(X)} \frac{dS(X)}{dX} \right\} \Psi_n \frac{d\Psi_m}{dX} dX, \quad (12)$$

$$B_{jnm} = (\gamma - 1) \int_0^1 W_j(X) \Psi_n \frac{d^2 \Psi_m}{dX^2} dX, \quad (13)$$

$$C_{jnm} = 2 \int_0^1 W_j(X) \frac{d\Psi_n}{dX} \frac{d\Psi_m}{dX} dX, \quad (14)$$

$$\begin{aligned} E_{jn}(T) = & \left[\int_0^1 W_j(X) \frac{d\Psi_n}{dX} dX + (\gamma - 1) \int_0^1 W_j(X) \right. \\ & \times \left\{ \frac{1}{S} \frac{dS}{dX} \right\} \frac{d\Psi_n}{dX} dX \\ & \left. + (\gamma - 1) \int_0^1 W_j(X) \frac{d^2 \Psi_n}{dX^2} dX \right] F \eta_n \\ & + \left[\Omega \int_0^1 W_j(X) dX \right] \frac{dF}{dT}, \end{aligned} \quad (15)$$

and

$$\begin{aligned} L_{jn}(T) = & \left(\frac{D\Omega}{\pi^3} \int_0^L W_j(X) \left[\frac{d^2 \Psi_n}{dX^2} \right. \right. \\ & \left. \left. + \left\{ \frac{1}{S} \frac{dS}{dX} \right\} \frac{d\Psi_n}{dX} \right] dX \right) \dot{\eta}_n. \end{aligned} \quad (16)$$

Specification of initial conditions, trial functions $\Psi_n(X)$, and weighting functions $W_n(X)$ provides all the information needed to solve Eqs. (8), which are numerically integrated to determine the amplitude functions $\eta_j(T)$ from time zero to any desired time. It follows from Eq. (8) that the number of modes, N , included in the approximate solution directly determines the number of coupled differential equations which must be solved. Additionally, the number of included modes also determines the accuracy of the solution, as will be demonstrated later. Finally, once Φ is determined, the conservation equations are used to determine the remaining un-

knowns.⁴ For example, the momentum equation is used to obtain the following expression for the nondimensional pressure at any location:

$$\frac{P}{P_0} = \left[1 - (\gamma - 1) \pi^2 \left(\Omega \frac{\partial \Phi}{\partial T} + \frac{1}{2} \left(\frac{\partial \Phi}{\partial X} \right)^2 + FX - \frac{D}{\pi^3} \frac{\partial^2 \Phi}{\partial X^2} \right) \right]^{\gamma/(\gamma-1)} \quad (17)$$

B. Trial functions

The choice of trial functions $\Psi_n(x)$ used in the Galerkin method are often solutions of simpler, yet related problems. For this study, $\Psi_n(x)$ are chosen to be the natural acoustic modes of the same geometrically similar duct. These functions can be determined analytically for certain duct shapes by solving the following linear wave equation, corresponding to the left-hand side of Eq. (1):

$$\Omega^2 \frac{\partial^2 \Phi}{\partial T^2} - \frac{1}{\pi^2} \frac{1}{A} \frac{dA}{dX} \frac{\partial \Phi}{\partial X} - \frac{1}{\pi^2} \frac{\partial^2 \Phi}{\partial X^2} = 0, \quad (18)$$

where A describes the axial dependence of the cross-sectional area of a duct that may equal or differ from that of the investigated duct. Note that Eq. (18) is the Webster horn equation¹³ expressed in terms of velocity potential.

To illustrate the determination of trial functions used in this study, consider a duct with exponentially varying axial cross-sectional area

$$A = A_0 \exp(\alpha X), \quad (19)$$

where A satisfies the relationship

$$\frac{1}{A} \frac{dA}{dX} = \alpha, \quad (20)$$

and α is referred to as the flare constant. For this duct shape, Eq. (18) becomes

$$(\pi \Omega)^2 \frac{\partial^2 \Phi}{\partial T^2} - \alpha \frac{\partial \Phi}{\partial X} - \frac{\partial^2 \Phi}{\partial X^2} = 0. \quad (21)$$

Given that the frame of reference of this study is attached to the duct, Φ must satisfy the following zero velocity boundary conditions:

$$\begin{aligned} \frac{d\Phi(0,T)}{dX} &= 0, \\ \frac{d\Phi(1,T)}{dX} &= 0. \end{aligned} \quad (22)$$

Assuming $\Phi(X,T) = \tau(T)\Psi(X)$ and using the method of separation of variables to solve Eq. (21) subject to the boundary conditions described by Eq. (22), one obtains an equation of the classical Sturm Liouville form¹⁴ with the following sets of eigenfunctions and eigenvalues:

$$\begin{aligned} \Psi_n(X) &= \exp\left(\frac{-\alpha X}{2}\right) \left(\cos(n\pi X) + \frac{\alpha}{2n\pi} \sin(n\pi X) \right) \\ n &= 1, 2, 3, \dots, \end{aligned} \quad (23)$$

TABLE I. Calculation of natural acoustic mode frequencies for straight and horn shaped ducts based on Eq. (24) assuming a sound speed (c_0) of 330 m/s.

Natural acoustic mode	Straight duct natural frequency (Hz) $\ell = 0.165$ m	Horn-shaped duct natural frequency (Hz) $\alpha = 2.5$ $\ell = 0.178$ m	Horn-shaped duct natural frequency (Hz) $\alpha = 5.75$ $\ell = 0.224$ m
1	1000	1000	1000
2	2000	1896	1623
3	3000	2813	2314
4	4000	3737	3027

$$k_n^2 = \left(\frac{\omega_n}{c_0} \right)^2 = \left(\frac{n\pi}{\ell} \right)^2 + \frac{1}{4} \left(\frac{\alpha}{\ell} \right)^2 \quad \text{for } n = 1, 2, 3, \dots, \quad (24)$$

where $\Psi_n(X)$ forms a complete set and k_n are the natural acoustic wave numbers of the duct.

An examination of the natural acoustic mode frequencies determined by Eq. (24) illustrates an important difference between straight and shaped ducts. Consider, for example, the natural acoustic mode frequencies of three ducts, one straight and two horn shaped, having the same fundamental natural acoustic mode frequency. Table I presents the first four natural acoustic mode frequencies, calculated using Eq. (24) for these three ducts. This calculation of natural acoustic mode frequencies provides a quantitative measure of the well-documented^{1,4} shifting of natural acoustic mode frequencies important to this problem. Furthermore, Table I shows that as flare constant increases, the natural acoustic mode frequencies decrease further away from their respective straight duct values.

Upon determination of the trial functions, the final terms needed for a complete description of the time-dependent amplitude Eqs. (8) are the weighting functions, $W_j(X)$. The choice of weighting functions must satisfy the orthogonality condition of the Sturm Liouville equation¹⁴ governing the natural acoustic mode shapes expressed in Eq. (23), and is defined in this study as

$$W_n(X) = A(X)\Psi_n(X). \quad (25)$$

This form of weighting functions is chosen in order to simplify the first terms of Eqs. (8) by reducing the summation to a single term where $n=m$. This simplification transforms Eqs. (8) into a system of nonlinear oscillators where the energy transfer between modes is modeled as the coupling between elements oscillating at the natural frequencies of the duct.

While the discussion above was limited to ducts with exponentially varying cross-sectional areas, similar analytical solutions can be found for other classes of duct shapes. For example, it can be shown that ducts with linearly varying cross-sectional area have natural acoustic mode shapes that are described by Bessel functions. In addition, as long as the chosen set of trial functions forms a complete set and satisfies the boundary conditions, this method can be applied to model any arbitrary duct shape (within the applicability of the model equation). As a result, this method can be applied in the event that the actual duct shape and the shape assumed to determine the trial functions are not the same. However, it is important to note that the more the actual duct shape dif-

fers from the duct shape assumed to get the trial functions, the more unrelated the actual physics of the system is from the resulting nonlinear oscillator equations obtained from the Galerkin procedure.

III. RESULTS

First, to check the accuracy of the developed Galerkin method solution approach, it was applied to predict the characteristics of nonlinear oscillations driven in straight ducts and the results were compared with those obtained in related studies.^{4,5} Next, results are presented for driven oscillations in horn-shaped ducts, identical to those used in the model development. Results from a parametric study of the influence of the flare constant on compression ratio are presented as an investigation into the optimal duct shaping for this class of duct shape. Finally, driven oscillations in several ducts, whose shapes were chosen following a procedure described in a recent study by Hamilton *et al.*¹² were investigated for two reasons; first, to display the ability of the developed Galerkin solution approach to model-driven oscillations in complex duct shapes, and second, to determine the impact of shifting the frequencies of specific natural acoustic modes upon the resulting duct oscillations.

Excluding the examination of numerical aspects of the developed Galerkin solution approach, all the reported results were obtained using approximate series solutions consisting of 20 modes. These solutions were rapidly computed on a desktop computer (<10 min for over 100 oscillations on a 400-MHz Pentium® II), thereby allowing quick turnaround parametric studies. To obtain the coefficient in the time-dependent amplitude equations, the integrals in Eqs. (9)–(16) were numerically solved by dividing the spatial domain into 1000 segments, leading to a resolution of a minimum of 50 points per wavelength for the highest frequency mode. In addition, a nondimensional time step of 0.01π was used, resulting in 10 time steps per cycle for the shortest period mode. Unless otherwise noted, the dissipation constant (D) and the ratio of specific heats (γ) were chosen as 0.01 and 1.2, respectively. Finally, in all of the studies, the entire duct was oscillated by the following sinusoidal forcing:

$$F = F_0 \cos(T). \quad (26)$$

A. Driven oscillations in straight ducts

Figure 1 compares the predictions of the Galerkin method for forced oscillations in straight ducts with two previously published solutions of the same problem.^{4,5} The results describe the pressure oscillations at one end of a straight duct forced at its fundamental natural acoustic frequency. For one of the comparisons, the Galerkin method and Ilinskii *et al.*⁴ numerical approaches were used to solve a nearly identical model equation [i.e., Eq. (1)]. The one difference between the two approaches, other than the differing solution techniques, is the exclusion of third-order nonlinearities from the equation solved by the Galerkin method. The third-order nonlinearities are apparently not important in the investigated problem as the Galerkin method solution is in excellent agreement with the Ilinskii *et al.* results. The second comparison in Fig. 1 corresponds to a study by Chun

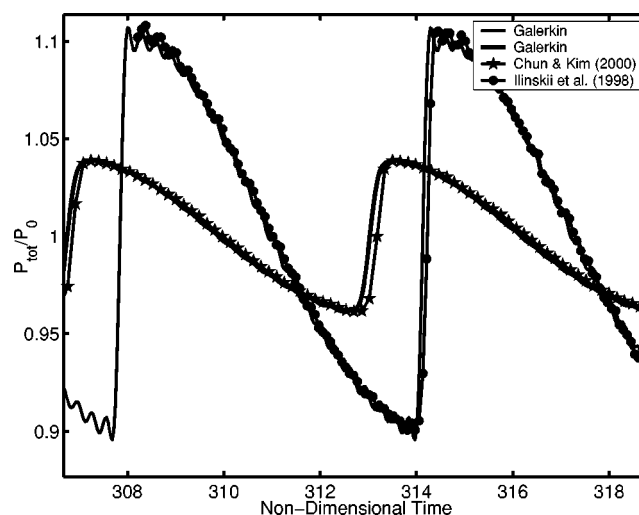


FIG. 1. Comparison of driven pressure oscillations within straight ducts using the Galerkin method with two other published results. Results from Galerkin method for comparison with Ilinskii *et al.* (Ref. 4) data correspond to larger forcing amplitude than that of the Chun and Kim (Ref. 5) study.

and Kim,⁵ who examined a sinusoidally driven 0.2 m long duct filled with 27 °C R-12 refrigerant (mol. wt=120.09, $\gamma=1.129$). To apply the Galerkin method solution to this problem, the dimensional forcing amplitude (a_0) of 100 m/s² reported by Chun and Kim was described by the normalized parameter $F_0 = a_0 / L\omega_0^2$, resulting in $F_0 = 8.939E-5$. It is noteworthy that in spite of the fact that Chun and Kim solved a different set of model equations and used a different numerical solution technique, their solution is in excellent agreement with the Galerkin method solution. Furthermore, the similar pressure amplitudes predicted in the two studies are attributed to the use of the same value for the viscosity coefficient, which was based on the value chosen by Ilinskii *et al.*⁴

The initial time evolution of the pressure at one end of a straight duct driven at its fundamental natural acoustic frequency ($\Omega = 1.0, F_0 = 5E-4$) is presented in Fig. 2. It shows that the pressure amplitude increases with time until nonlinear processes and dissipation limit its growth and a limit cycle is established. The latter occurs when the time average per cycle of the energy supplied by the forcing is balanced by that lost by dissipation. In this example, the initial transient behavior subsides after approximately 20 periods.

Figure 3 shows the time dependence of limit cycle oscillations forced in a straight duct and their FFT. Note that the power of each mode in the FFT is normalized with respect to the power in the fundamental mode. Therefore, the normalized power of the fundamental mode equals 1 and it is not fully represented, in order to better describe the power content of the higher harmonics. Figure 3 shows that the power contained in the first harmonic is larger than 16 percent of the power of the forcing frequency, with the relative energy content of higher harmonics decreasing as their frequency increases. It should also be noted that the power content of higher harmonics provides a measure of the efficiency of energy transfer between the driven mode and these harmonics by nonlinear processes.

Results for the straight duct can also be used to examine

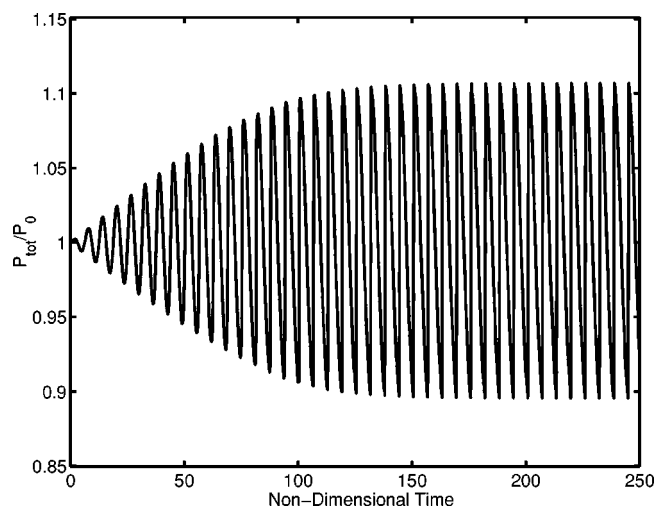


FIG. 2. Time evolution of pressure at closed end of a straight duct forced at the fundamental natural acoustic frequency starting from rest ($F_0=5E-4$, $\Omega=1.0$). Results show well-documented waveform steepening and saturation of total pressure amplitude as a result of nonlinear processes and viscous dissipation.

the impact of the number of modes included in the calculations upon the accuracy of the solution. Other than choosing a numerically stable temporal step size, the number of modes used to represent the solution exerts the most influence upon the accuracy of the solution, assuming that an adequate spatial step size is chosen to fully resolve the highest natural acoustic mode shape. Figure 4 displays results from calculations including different numbers of modes ranging from five to 20. The results clearly show that while the driven frequency is correctly captured with only five modes, the peak-to-peak pressure is significantly overpredicted and the waveform is rather distorted. Furthermore, Fig. 4 shows that as the number of modes included in the calculations increases, the waveform in between adjacent shock fronts is smoothed, i.e., the amplitudes of the higher frequency oscillations in between the shock fronts decrease.

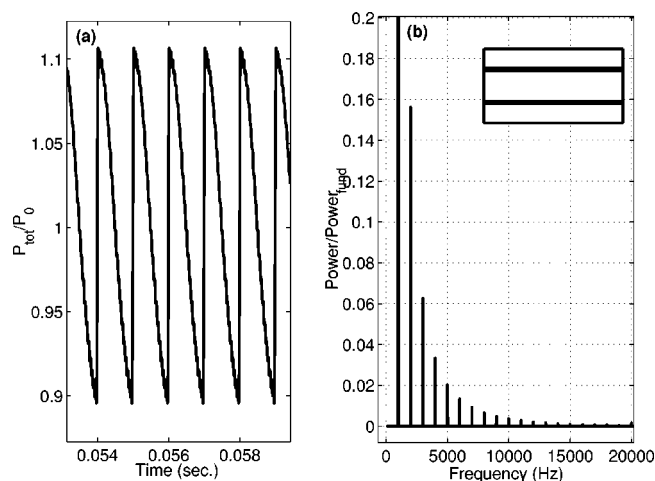


FIG. 3. (a) Time dependence of finite amplitude, limit cycle pressure oscillations driven in a straight duct ($F_0=5E-4$, $\Omega=1.0$, $c_0=330$ m/s, $\ell=0.165$ m), and (b), the corresponding FFT. Duct shape is shown in the upper right portion of (b). Note that the power in the FFT is normalized with respect to the power in the fundamental mode.

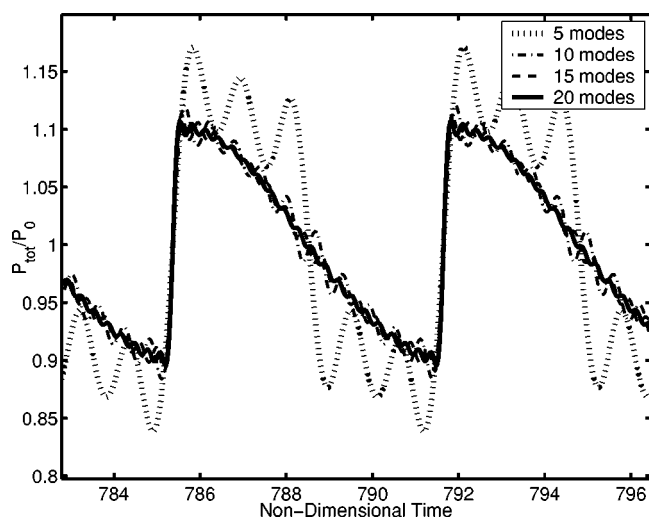


FIG. 4. Pressure time history at closed end of driven straight duct for calculations including different numbers of modes. The inclusion of only five modes significantly overpredicts overall pressure amplitude and does not accurately capture waveform.

B. Driven oscillations in horn-shaped ducts

This section describes the result of the study of forced oscillations in exponential horn-shaped ducts. The forcing coefficient equals that used in the previously discussed straight duct calculations ($F_0=5.0E-4$), and the forcing frequency ($\Omega=1.36$) corresponds to the frequency of maximum response for a duct with a flare constant (α) of 5.75. Given that the fundamental natural frequency of this shaped duct is $\Omega=1.355$ [calculated using Eqs. (3) and (24)], the results exhibit a hardening behavior, defined as an increase in the forcing frequency of maximum response with respect to the fundamental natural acoustic mode frequency. Additionally, given the value of the natural acoustic mode frequencies reported in Table I, agreement is achieved with recent theoretical work¹² predicting an increase in the forcing frequency of maximum response, $\omega_{\text{max response}} > \omega_1$ (i.e., hardening) when the condition $2\omega_1 > \omega_2$ is satisfied.

Figure 5(a) displays the time dependence of limit cycle pressure oscillations at the small end of a horn-shaped duct with a flare constant of 5.75. The normalized pressure amplitude reaches a maximum of over 2.5, which is considerably larger than the normalized amplitude of approximately 0.11 predicted for the straight duct. Figure 5(a) also presents a prediction obtained using only five modes, showing that the compression ratio is only slightly larger than that predicted by the 20-mode solution. Additionally, the waveform and frequency of the two calculations are very similar to one another, thus suggesting that the influence of higher harmonics for this duct is less important than for a straight duct. Figure 5(a) also shows that the shape of the waveform no longer resembles the saw-tooth shape found in straight ducts, the FFT of the waveform, shown in Fig. 5(b), indicates that the amplitudes of the higher harmonics are significantly smaller than that of the fundamental forcing frequency when compared with the amplitudes of the harmonics in straight ducts [see Fig. 3(b)]. Specifically, Fig. 5(b) shows that the power of the first harmonic in this horn-shaped duct is ap-

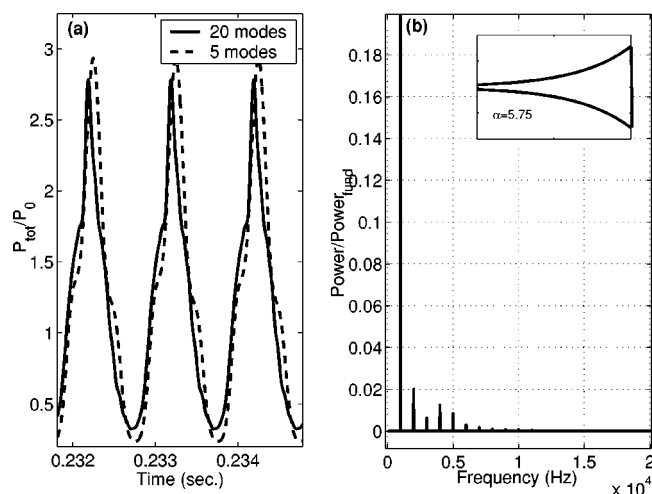


FIG. 5. Time dependence (a) and frequency content (b) of driven oscillations at the small end of horn-shaped duct ($\alpha=5.75$, $F0=5E-4$, $\Omega=1.36$, $c_0=330$ m/s, $\ell=0.224$ m). The frequency content in (b) corresponds to the calculations including 20 modes [solid line in (a)].

proximately 2 percent of the power of the fundamental. In contrast, the power of the first harmonic in the straight duct was larger than 15 percent of the power content of the fundamental mode.

As pointed out earlier and in previous studies,^{1,4} the natural acoustic mode frequencies of straight ducts are integer multiples of the fundamental. Hence, energy is readily transferred from the forced fundamental mode to its higher harmonics by nonlinearities because the frequencies of the harmonics equal the frequencies of the natural acoustic modes of the ducts. In contrast, for horn-shaped ducts, the natural acoustic mode frequencies of the system are generally not integer multiples of the fundamental frequency; see Table I. Consequently, energy transfer from the forced fundamental mode to the harmonics, which is by nature integer multiples of the forcing frequency, is not as efficient for horn-shaped ducts as for straight ducts. This “inefficient” energy transfer inhibits the formation of large amplitude harmonic oscillations, thus leading to appreciably less waveform steepening.^{1,4} This reduction in waveform “steeping” decreases local property gradients, thereby reducing viscous dissipation and “concentrating” the energy supplied by forcing in the driven mode.

While it was clearly demonstrated that with a given forcing amplitude it is possible to excite much larger amplitude oscillations in shaped ducts, it is still not clear whether an optimal duct shape can be found based on a desired performance parameter; i.e., is it possible that the compression ratio for a given volume and forcing amplitude is maximized for only one duct shape? This possibility was examined on a simplified framework by finding the flare constant α that maximizes the compression ratio for horn-shaped ducts. Figure 6 shows the maximum compression ratio for different flare constants ranging from 3.0 to 6.75. The results presented in Fig. 6 reveal that the compression ratio does not monotonically increase as the shape of the duct diverges from that of the straight geometry. In fact, a significant decrease in compression ratio occurs as the flare constant is

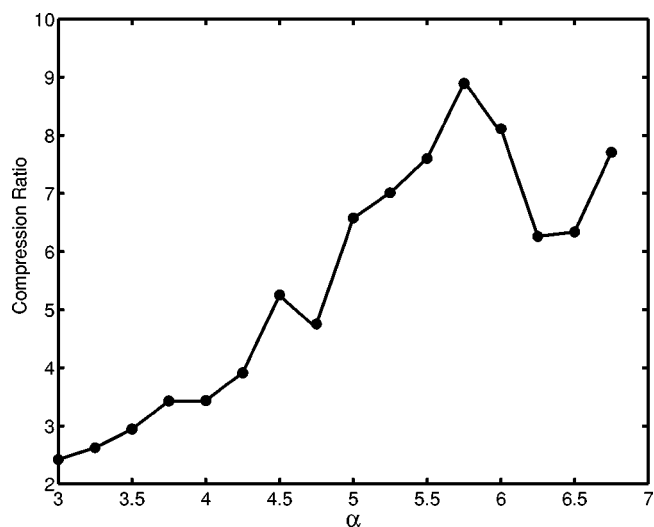


FIG. 6. Maximum compression ratio (P_{\max}/P_{\min}) as a function of flare constant (α). Each value of α corresponds to a different duct shape and forcing frequency. For values of α between 3.0 and 6.75, the compression ratio is maximized at $\alpha=5.75$.

increased from 5.75 to 6.25. These findings suggest that caution must be exercised in the design of acoustic compression devices because slight changes in duct shape can have a significant impact on overall performance.

C. Driven oscillations for $S(X) \neq A(X)$

As pointed out in the model development, the developed Galerkin method solution can model forced oscillations in ducts having arbitrary shapes within the model assumptions. To illustrate this point, the functions describing the spatial dependence of the acoustic modes of a straight duct were used as the trial functions for the Galerkin solution of ducts whose shapes were determined by the following equation developed by Hamilton *et al.*:¹²

$$S(X) = S_0 \{ \exp[a_m \cos(2m\pi X) + a_n \cos(2n\pi X) + \dots] \}, \quad (27)$$

where $S(X)$ is the shape of the duct and the subscripts m, n, \dots , denote the specific natural acoustic mode frequencies to be shifted by amounts that depend upon on a_m, a_n, \dots , respectively. Using this duct shape, it is possible to shift the frequencies of specific natural acoustic modes with respect to those of the straight duct. For example, to only shift the frequency of the second natural acoustic mode, $m=2$, $a_2 \neq 0$, and $a_n, \dots, = 0$. Figure 7 presents results for three different duct shapes, obtained using Eq. (27), corresponding to shifting of the second ($a_2 = -0.5, a_3 = 0$) and third ($a_2 = 0, a_3 = -0.5$) natural acoustic mode frequencies (one in each case) as well as the simultaneous shifting of the frequencies of both modes ($a_2 = -0.5, a_3 = -0.5$). Figure 7 shows that for the duct in which only the frequency of the second natural acoustic mode was shifted, a portion of the waveform of the excited oscillations exhibited very steep property gradients. However, unlike the saw-tooth waveform driven in the straight duct, the pressure continues to rise after this steep increase. Additionally, while the compression ratios for these ducts do not approach the compression ratio

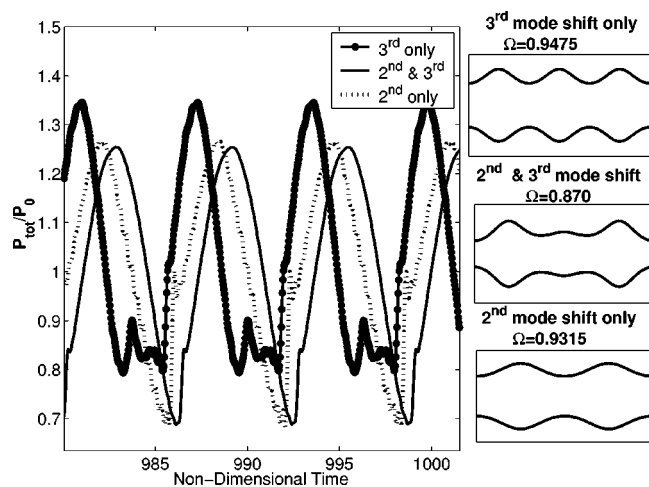


FIG. 7. Pressure time history at an end of three different duct shapes. For duct shapes displayed on the right, the top, middle, and lower ducts correspond to $a_2=0$, $a_2=-0.5$, $a_2=-0.5$, $a_3=-0.5$, and $a_2=-0.5$, $a_3=0$, respectively. The results show the effect of individual natural acoustic mode shifting as suggested by Hamilton *et al.* (Ref. 12).

attained in the horn-shaped ducts, it should be pointed out that cross-sectional area of the small end of the horn-shaped ducts must be significantly smaller than the cross-sectional area of the ends of the straight duct and the three ducts investigated here to keep the volume of both ducts the same. Nonetheless, the compression ratio of forced oscillations for all three ducts is significantly larger than that predicted for the straight duct.

IV. CONCLUSIONS

Finite amplitude oscillations driven in straight and shaped ducts were investigated using the approximate Galerkin method. The developed Galerkin method solutions provided approximate, time-dependent solutions of driven oscillations within closed ducts with reasonable accuracy and at a low computational cost. The characterization of driven oscillations within straight and shaped ducts in terms of compression ratio and frequency content was presented. In addition, it was found that the compression ratio of oscillations driven

in closed, horn-shaped ducts is sensitive to small changes in duct shape and scales nonmonotonically with flare constant. Finally, results were presented using trial functions based on a duct shape different from that of the actual duct shape in order to demonstrate the ability of this technique to characterize driven oscillations in a wide variety of duct shapes. This configuration also allowed exploration into the impact of the shifting of individual natural acoustic frequencies of a duct. These results showed that the individual shifting of natural acoustic frequencies can increase compression ratio as well as maximum attainable pressure amplitude in addition to significantly impacting waveform.

- ¹C. C. Lawrenson, B. Lipkens, T. S. Lucas, D. K. Perkins, and T. W. Van Doren, "Measurements of macrosonic standing waves in oscillating closed cavities," *J. Acoust. Soc. Am.* **104**, 623–636 (1998).
- ²G. W. Swift, "Thermoacoustic engines," *J. Acoust. Soc. Am.* **84**, 1145–1180 (1988).
- ³J. V. Foa, *Elements of Flight Propulsion* (Wiley, New York, 1960).
- ⁴Y. A. Ilinskii, B. Lipkens, T. S. Lucas, T. W. Van Doren, and E. A. Zabolotskaya, "Nonlinear standing waves in an acoustical resonator," *J. Acoust. Soc. Am.* **104**, 2664–2674 (1998).
- ⁵Y. D. Chun and Y. H. Kim, "Numerical analysis for nonlinear resonant oscillations of gas in axisymmetric closed tubes," *J. Acoust. Soc. Am.* **108**, 2765–2774 (2000).
- ⁶B. A. Finlayson, *The Method of Weighted Residuals and Variational Principles* (Academic, New York, 1972).
- ⁷W. Chester, "Resonant oscillations in closed tubes," *J. Fluid Mech.* **18**, 44–64 (1964).
- ⁸D. B. Cruikshank, Jr., "Experimental investigation of finite-amplitude acoustic oscillations in a closed tube," *J. Acoust. Soc. Am.* **52**, 1024–1036 (1972).
- ⁹J. Jimenez, "Nonlinear gas oscillations in pipes. I. Theory," *J. Fluid Mech.* **59**, 23–46 (1973).
- ¹⁰R. Saenger and G. Hudson, "Periodic shock waves in resonating gas columns," *J. Acoust. Soc. Am.* **32**, 961–970 (1960).
- ¹¹C. Vanhille and C. Campos-Pozuelo, "Numerical model for nonlinear standing waves and weak shocks in thermoviscous fluids," *J. Acoust. Soc. Am.* **109**, 2660–2667 (2001).
- ¹²M. F. Hamilton, Y. A. Ilinskii, and E. A. Zabolotskaya, "Linear and nonlinear frequency shifts in acoustical resonators with varying cross sections," *J. Acoust. Soc. Am.* **110**, 109–119 (2001).
- ¹³L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982).
- ¹⁴W. E. Boyce and R. C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, 4th ed. (Wiley, New York, 1986).

Acoustic attenuation in gas mixtures with nitrogen: Experimental data and calculations

Sally G. Ejakov^{a)}

Center for Sensor Materials, Michigan State University, East Lansing, Michigan 48824-1116

Scott Phillips

Commercial Electronics, Broken Arrow, Oklahoma 74012-2838

Yefim Dain and Richard M. Lueptow^{b)}

Department of Mechanical Engineering, Northwestern University, Evanston, Illinois 60208-3111

Jacobus H. Visser

Ford Research Laboratory, Mail Drop 3083/Scientific Research Laboratory, Dearborn, Michigan 48121-2053

(Received 31 July 2002; revised 20 December 2002; accepted 13 January 2003)

Attenuation in a gas results from a combination of classical attenuation, attenuation from diffusion, and attenuation due to molecular relaxation. In previous papers [J. Acoust. Soc. Am. **109**, 1955 (2001); **110**, 2974 (2001)] a model is described that predicts the attenuation from vibrational relaxation in gas mixtures. In order to validate this model, the attenuation was measured using a pulse technique with four transducer pairs, each with a different resonant frequency. The attenuation calculated using the model was compared to the measured values for a variety of gases including: air, oxygen, methane, hydrogen, and mixtures of oxygen/nitrogen, methane/nitrogen, carbon dioxide/nitrogen, and hydrogen/nitrogen. After the measured data is corrected for diffraction, the model matches the trends in the measured attenuation spectrum for this extensive set of gas mixtures. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1559177]

PACS numbers: 43.35.Ae [RR]

I. INTRODUCTION

Attenuation in gases results from several mechanisms that transfer the translational energy of the acoustic wave into other forms of energy. These mechanisms include classical effects related to viscosity and thermal conductivity, losses due to diffusion in gas mixtures, and losses due to excitation and relaxation of vibrational or rotational molecular energy levels in the gas.¹ Classical acoustic attenuation results from irreversible losses of acoustic energy to heat due to shear viscosity and thermal conductivity across temperature gradients related to the compression and rarefaction of the acoustic wave. Classical attenuation is well understood and can be predicted for ideal gases, both pure and in mixtures, as long as the viscosity and conductivity of the mixture are known. Attenuation due to diffusion occurs in mixtures when light gas molecules diffuse faster than heavier ones locally changing the mixture composition as the acoustic wave passes. This loss of entropy in the arrangement of molecules results in a reduction in the energy in the acoustic wave. Attenuation related to diffusion is largest for mixtures of gases having much different masses (e.g., hydrogen/nitrogen mixtures). The attenuation from diffusion can be calculated for gas mixtures, if the appropriate constants are known.² Exciting the energy states of the gas molecules and the relaxation of these states also causes attenuation. The

transfer of translational energy to internal modes occurs with a relaxation time that depends on the collisional dynamics of the internal vibrational modes available and is different than the time required to equilibrate the translational energy. The collisional dynamics in some gases result in relaxation times that correspond to relaxation frequencies ranging from a few Hz up to 10 MHz, where a significant increase in attenuation can be observed. On the other hand, rotational modes in the molecules typically have much shorter relaxation times and consequently usually affect the attenuation only at much higher frequencies.

The relaxation of vibrational modes in a molecule results in attenuation that is strongly dependent on the gases present. In gas mixtures with more than two species, the interactions of vibrational modes and their effect on attenuation were not fully quantified until recently. To remedy this, two of us, Dain and Lueptow, developed a model that describes molecular relaxation for mixtures of gases.^{3,4} In brief, this model numerically solves the Euler gas equations and the population equations for the energy states of the molecules (including collision rates and transition probabilities).⁵⁻⁷ More details of the calculations in particular as they apply to the molecular species reported in this paper are given in the Appendix. The calculations agree well with measurements reported in the literature for the cases where they can be compared.³ Unfortunately, the available measurements of attenuation in gases cover only limited species, report many results only at elevated temperatures, and contain some discrepancies even for single component gases.⁸⁻¹³ This study is focused on measurements of the at-

^{a)}Current address: Ford Motor Company, POEE Mail Drop 43, 21500 Oakwood Boulevard, Dearborn, MI 48124; electronic mail: gaffsall@msu.edu

^{b)}Email: r-lueptow@northwestern.edu

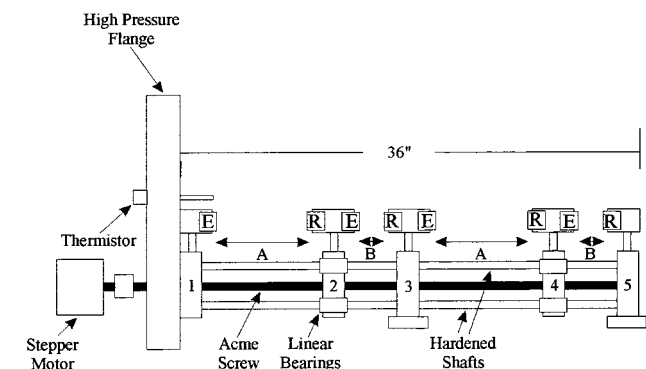


FIG. 1. A schematic diagram of the transducer traverse system mounted on the high-pressure flange. The stepper motor moves supports 2 and 4. This whole assembly was sealed inside a cylindrical chamber.

tenuation at room temperature for several binary mixtures of practical interest including methane/nitrogen, carbon dioxide/nitrogen, oxygen/nitrogen, and hydrogen/nitrogen mixtures. Our objective in this research was twofold. First we used an experimental apparatus similar to that used previously,¹⁴ but incorporating modern transducers and electronics technology, to accurately measure acoustic attenuation in gas mixtures. Second, we compare our experimental results with previous results and our model results to further validate our model.

II. EXPERIMENTAL DESCRIPTION

Measurements of the attenuation in various gases were carried out in a large cylindrical chamber with a diameter of 30.5 cm (12 inches) and a length of 91.5 cm (36 inches). Figure 1 shows the transducer traverse system mounted on the high-pressure flange that seals one end of the chamber. The supports (labeled 1–5) hold four sets of transducers, with supports 2, 3, and 4 each holding both an emitter (E) and a receiver (R). The acoustic measurements are made between an emitter on one support, for example, support 1, and a receiver on the next support, in this case support 2, with a separation distance A between the transducers. A stepper motor drives an Acme screw to displace supports 2 and 4 with a resolution of 0.001 cm (0.0005 inches) per step, while the other supports (1, 3, and 5) remain stationary. The transducer

positions are zeroed at the beginning of each test, and the separation between transducers is measured to within 0.003 cm (0.001 inches). In these measurements the distance between the emitter and receiver of each pair was adjusted from 0.762 to 19.812 cm (0.300–7.800 inches) using 75 steps.

At the beginning of each test, the chamber was evacuated to 0.0003 atm (0.05 psi), filled to 0.7 atm (10 psi) with the test gas, evacuated again to 0.0003 atm and then filled to a high pressure (30 atm) with a precise mixture of gases. The gas mixing for the chamber was controlled with mass flow controllers and a specially designed regulator that allows the test chamber to be pressurized, while the output pressure for the mass flow controllers was held constant providing the gas composition to within 0.01%. The nitrogen, air, oxygen, methane, carbon dioxide, and hydrogen used in these tests were all 99.99 or 99.999% pure.

Since attenuation depends on frequency divided by pressure (f/p), either frequency or pressure can be varied. Because the transducers are tuned to a particular resonant frequency, it is easier and more effective to vary the pressure in the chamber than to vary the transducer frequency. Thus, measurements of the attenuation were made at the initial high gas pressure, and then some gas was removed from the chamber and the test performed again. Measurements were made at 11 different pressures from 30–0.6 atm.

Four sets of transducers were used in a pitch–catch configuration. The piezoelectric transducer pairs had matched frequencies of 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz. Other details of the transducers are listed in Table I. After the gas filled the chamber to the correct pressure and the transducers had been moved to the appropriate separation distances using the stepper motor, the first emitter was excited with a burst of 10 pulses at the frequency of the transducer pair. (Bursts of 1 and 5 pulses give similar values for attenuation, but the amplitude is smaller, so 10 pulses were preferred.) Accounting for the speed of sound in the gas mixture being tested, the signal that corresponded to the original pulse measured at the receiver was recorded. In addition, the pressure and temperature in the test chamber along with other parameters were recorded. Each emitter was energized individually and after all four emitters had been pulsed, the

TABLE I. Data related to the transducers, including: transducer radius (R), transition point (R^2/λ), and the fit region used in the analysis for the individual gases for each transducer pair.

Frequency (kHz)	R (cm)	R^2/λ (cm) air	Fit region (cm) air	R^2/λ (cm) O ₂	Fit region (cm) O ₂	R^2/λ (cm) CH ₄	Fit region (cm) CH ₄	R^2/λ (cm) CO ₂	Fit region (cm) CO ₂	R^2/λ (cm) H ₂	Fit region (cm) H ₂
92 ^a	1.3	4.52	12.700– 19.812	4.42	12.700– 19.812	3.49	12.700– 19.812	5.56	7.620– 12.700	1.20	12.700– 19.812
149.1 ^b	0.9	3.51	7.620– 19.812	3.43	7.620– 19.812	2.71	7.620– 19.812	4.32	7.620– 12.700	0.93	7.620– 19.812
215 ^a	0.6	2.25	7.620– 19.812	2.20	7.620– 19.812	1.74	7.620– 19.812	2.77	7.620– 12.700	0.59	7.620– 19.812
1000 ^b	1.0	29.07	0.762– 7.620	28.42	0.762– 7.620	22.44	0.760– 7.620	35.74	0.760– 2.540	7.69	0.762– 7.620

^aManufactured by ITC.

^bManufactured by Etalon.

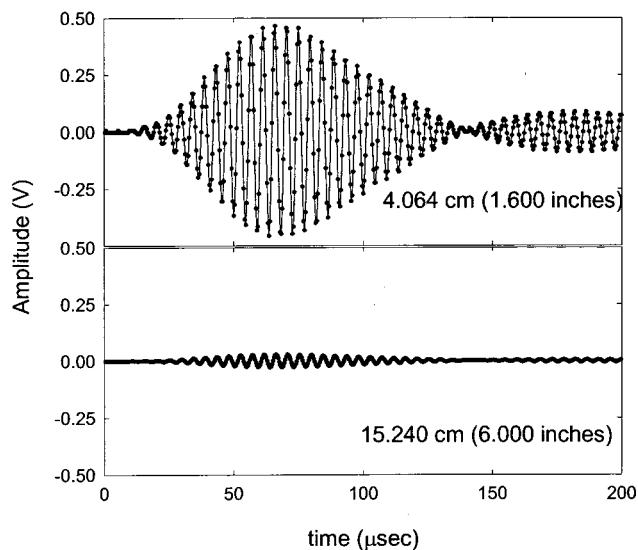


FIG. 2. Raw signals from the 215 kHz transducer in 100% methane at a chamber pressure of 3.18 atm at two different separation distances, 4.064 cm and 15.240 cm.

transducers were moved to a new separation distance and the measurement was repeated. When the range of separation distances had been covered, the pressure was reduced for the next set of tests, and the test protocol was repeated.

Measurements were made in several gases, both pure gases (oxygen, methane, and hydrogen) and mixtures (air, oxygen/nitrogen, methane/nitrogen, carbon dioxide/nitrogen, and hydrogen/nitrogen). (Pure carbon dioxide has such a strong attenuation that the resulting pulses at the receiver were below our detection threshold, so data could not be obtained in this case.) The mixtures were tested in concentration steps of 20% from 20–80%. The chamber was at ambient temperature, with an average temperature during a test between 292.6 and 298.6 K. While this temperature range is relatively small, it affects the attenuation to some extent. Using our model of attenuation, we estimate that the attenuation can change by a maximum of 10% over this temperature range, depending on the gases present. However, the temperature range for any given test was smaller, varying by 1.0–3.5 K.

III. DATA ANALYSIS

The harmonic acoustic pressure P decreases with distance z from the emitter according to

$$P = P_0 e^{-\alpha z}, \quad (1)$$

where P_0 is the amplitude of the acoustic pressure at the emitter, and α is the attenuation. In our tests the voltage produced in the receiving transducer by the sound wave is proportional to the acoustic pressure. Therefore, the attenuation can be found from the slope of the logarithm of the voltage amplitude plotted as a function of separation distance according to Eq. (1). This equation is exact for plane waves, but requires a correction for diffraction of a sound wave emitted from a transducer of finite size, discussed shortly.

For each gas mixture, transducer pair, and separation distance, the raw data consists of the receiver voltage versus

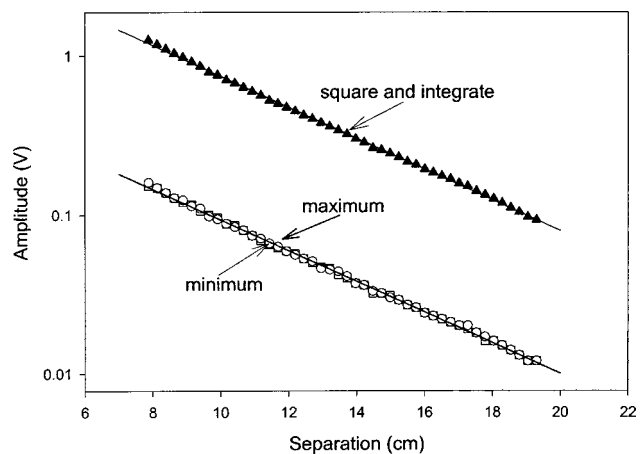


FIG. 3. Amplitude vs separation distance for 100% methane, the 215 kHz transducers, and a pressure of 3.18 atm. The amplitude was evaluated as the maximum or largest value of the data set (squares), the minimum or absolute value of the most negative value in the data set (circles), and the sum of each point in the data set squared (triangles). The curves are least-squares fits to the data.

time, as shown in Fig. 2 for the 215 kHz transducer with 100% methane at a pressure of 3.18 atm (46.76 psi). The entire data set of 512 points was recorded over 40 μ s for the 1 MHz transducer, 200 μ s for the 149.1 and 215 kHz transducers, and 400 μ s for the 92 kHz transducer. The form of the wave packet is typical of the transducer response and does not show the first reflection, which arrived 182 μ s after the original wave for the 4.064 cm separation distance. Even though the emitter was excited with only 10 pulses, ringing of the emitter results in two to three times that number of peaks in the voltage being recorded at the receiver. (The second series of pulses at about 150 μ s are anomalous signals possibly related to transducer ringing or reflections from the mounting system. They were omitted from the analysis.) The decrease in amplitude resulting from attenuation as the separation distance is increased from 4.064 to 15.240 cm is evident in Fig. 2.

Three methods were evaluated to measure the amplitude of the signal at the receiver: the maximum of the signal (largest value observed), the minimum of the signal (absolute value of the most negative value observed), and the integral of the square of the signal, which is equivalent to the sound intensity. No curve fitting was used for the maximum and minimum as the raw data has enough resolution to provide points very close to the extrema. For all three methods to evaluate the amplitude, the log of the voltage amplitude (which is proportional to acoustic pressure) plotted as a function of separation distance should yield a linear relationship with a slope of α , in accordance with Eq. (1). Figure 3 compares the amplitude of the signal evaluated using the three methods as a function of separation distance for a typical data set. Although the magnitude of the amplitude depends on the method of evaluation, the slope, which corresponds to the attenuation α , is very similar for all three techniques. Consequently, the maximum of the signal was used to determine the attenuation, since it is most convenient and provides identical results as shown in Fig. 3.

In order to accurately determine the physical attenua-

tion, it is necessary to correct for any other effects that also change the amplitude of the acoustic wave with separation distance such as spreading and diffraction. The sound field has many maxima and minima that approximately average out in the near field region close to the emitting transducer so that no correction is needed. However, a correction is needed in the far field as the sound decreases in amplitude due to spreading and diffraction as the distance from the emitter increases.¹⁵ It is important to measure the attenuation either completely in the near field or completely in the far field to avoid difficulty in accounting for the diffraction.¹⁶ The transition between these two regions occurs at approximately R^2/λ , where R is the transducer radius and λ is the wavelength of the sound.¹⁷ The values for this transition point for each transducer pair evaluated for pure gases used in these measurements are given in Table I. In many cases, the transition between the near and far fields occurs within the range of separation distances (0.762–19.812 cm) in our experiments.

For each gas composition and transducer pair it is necessary to choose the range of separation distances corresponding to either the far or near field that should be used to find the attenuation, α . The ranges that were used are shown in Table I for the pure gases. Mixtures of these gases with nitrogen were analyzed over the range used for the pure gas. The data measured in the far field was used for the three lower frequency transducers. For the 1 MHz transducer the data in the near field was used, but the range of data was further restricted to separation distances less than 7.620 cm, because at larger separation distances the strong attenuation at this frequency reduced the signal to such a low amplitude that it was no longer measurable. Because carbon dioxide has greater attenuation than the other gases, the maximum separation was further restricted in these cases.

For the three lower frequency transducers, it is necessary to correct the data for diffraction in the far field. We considered two diffraction correction methods. Pinkerton calculated the amplitude, $A(z)$, on the axis of the coaxial transducers as

$$A(z) = A_0 e^{-\alpha z} \left\{ \sin \left[\frac{1}{2} k (\{z^2 + R^2\}^{1/2} - z) \right] \right\}, \quad (2)$$

where A_0 is the amplitude at the emitter, $k = 2\pi/\lambda$ is the wave number, and z is the separation between the emitter and the receiver.¹⁶ A more precise diffraction correction was developed by Khimunin¹⁸ for harmonic waves based on Williams' formula for average pressure on a receiver.¹⁹ Rogers and Van Buren used a simplified version of this expression integrating the acoustic amplitude over the surface of the emitter and the receiver, both of radius R , to find a short wave approximation ($ka \gg 1$) for the total average amplitude

$$A(z) = A_0 e^{-\alpha z} \left\{ \left[\cos \left(\frac{2\pi}{s} \right) - J_0 \left(\frac{2\pi}{s} \right) \right]^2 + \left[\sin \left(\frac{2\pi}{s} \right) - J_1 \left(\frac{2\pi}{s} \right) \right]^2 \right\}^{1/2}, \quad (3)$$

where $s = 2\pi z/kR^2$, J_0 and J_1 are zero and first order Bessel functions, and z is the separation between the transducers.²⁰ Both of these corrections assume that diffraction and attenuation are independent processes. The exponential is related to

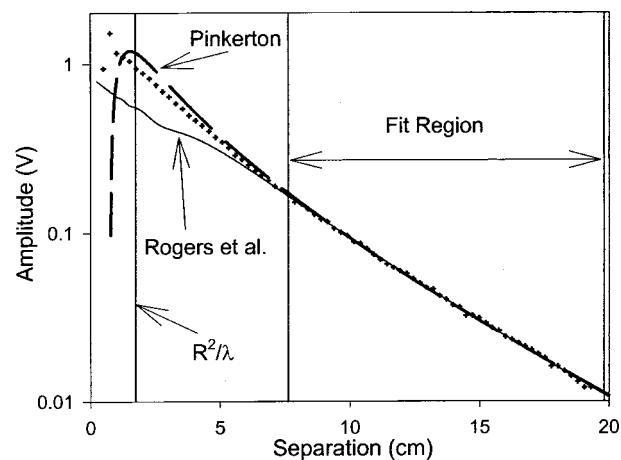


FIG. 4. Amplitude of the sound wave as a function of separation distance for 100% methane, the 215 kHz transducers and 3.18 atm. The dots show the measurements. The diffraction correction method of Pinkerton [Eq. (2)], $\alpha = 0.375$, is the dashed curve, and the diffraction correction method of Rogers *et al.* [Eq. (3)], $\alpha = 0.397$, is the solid curve.

the attenuation, while the term in brackets is the diffraction correction. Ideally, a complex wave number should be used in Eqs. (2) and (3) to incorporate the effect of acoustic attenuation on diffraction. In addition, Eqs. (2) and (3) were developed for harmonic waves but not for bursts, which were used in these experiments. Nevertheless, both methods provided good results. An analysis incorporating a complex wave number (and thereby attenuation) into a diffraction correction for bursts rather than harmonic waves in a relaxing medium will be addressed in a separate paper.²¹

The two diffraction corrections are compared with the raw data in Fig. 4 for tests of 100% methane using the 215 kHz transducer. In both cases, the attenuation α in Eqs. (2) and (3) was adjusted to provide the best fit with the data in the fit region noted in Table I. In the fit region both corrections match the data well, but at shorter separation distances, near the transition between the near and far field, the simpler Pinkerton form follows the data better. This result was confirmed by calculating the attenuation for all of the chamber pressures and transducers for the gas mixtures measured. The diffraction correction was applied by dividing the intensity at each separation distance by the calculated correction at that distance and then fitting the resulting curve to find α . Applying the Pinkerton correction to the data in the standard attenuation spectrum form of $\alpha\lambda$ as a function of f/p results in an overlap between the data from different transducers at the same f/p (as will be evident in Figs. 5–12, which will be discussed later). The Rogers and Van Buren correction does not result in overlap of the data for the different transducers when plotted as an attenuation spectrum, even when the radius R in Eq. (3) is replaced by an effective radius that is some fraction of R . Thus, the Pinkerton correction for diffraction was used for all of the data that is presented in the next section.

IV. RESULTS

In this section we present the results of attenuation measurements for several pure gases and gas mixtures, most of

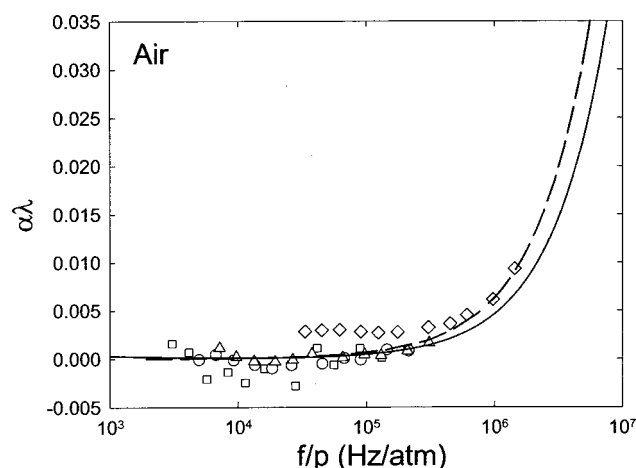


FIG. 5. Results for air using the Pinkerton diffraction correction at an average temperature of 293.1 K. Data points are for the 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz transducers (squares, circles, triangles, and diamonds, respectively). The solid curve is the classical attenuation and the dashed curve is the empirical model from Bass *et al.* (Ref. 1).

which have not been previously measured. These results are useful in understanding the effect of gas composition on acoustic attenuation. In addition, we compare the experimental results with the sum of the theoretical model of Dain and Lueptow³ for vibrational relaxation, the classical attenuation, and the diffusional attenuation. This allows us to determine the effectiveness of the vibrational relaxation model in predicting acoustic attenuation. The classical attenuation was calculated using the Stokes and Kirchoff formulation, and the attenuation due to diffusion was calculated from Bhatia's formulations.² More details on the vibrational relaxation calculations are included in the Appendix. In all cases we plot the attenuation α nondimensionalized by the wavelength λ as a function of the frequency divided by the pressure, f/p , which can be called an attenuation spectrum.

A. Air

The acoustic attenuation spectrum for air is shown in Fig. 5. It is immediately evident that the different transducers

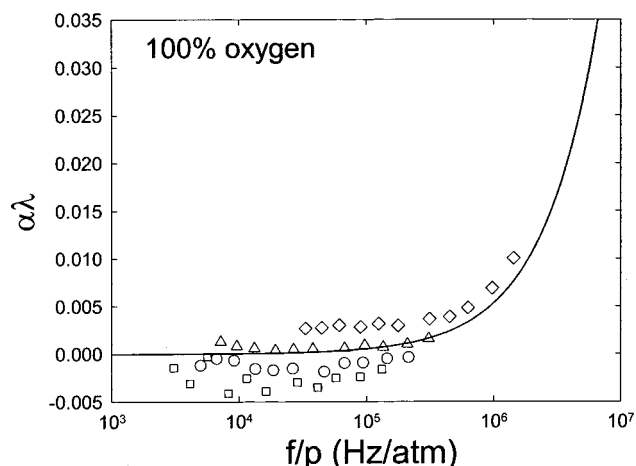


FIG. 6. Results for 100% oxygen using the Pinkerton diffraction correction at an average temperature of 293.3 K. Data points are for the 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz transducers (squares, circles, triangles, and diamonds, respectively). The solid curve is the classical attenuation. Vibrational relaxation is negligible.

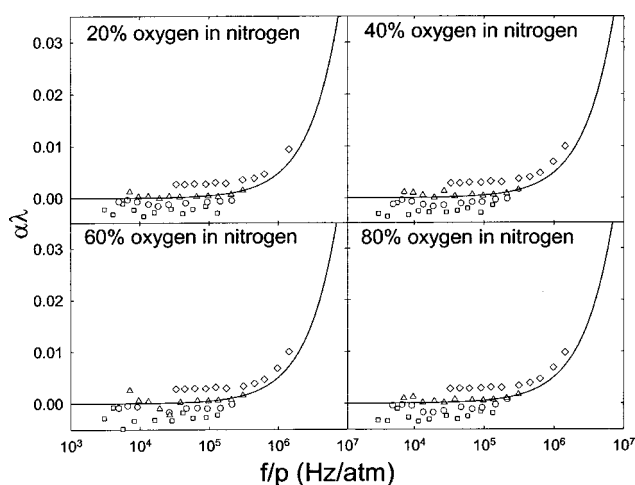


FIG. 7. Results for mixtures of 20%, 40%, 60%, and 80% oxygen in nitrogen using the Pinkerton diffraction correction at average temperatures of 293.6, 292.8, 292.8, and 293.9 K, respectively. Data points are for the 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz transducers (squares, circles, triangles, and diamonds, respectively). The solid curves are the sum of the classical and diffusional attenuation. Vibrational relaxation is negligible.

having frequencies of 92 kHz to 1 MHz (indicated by different symbols) give similar values. The values from the 1 MHz transducer (diamonds) are slightly higher, possibly because the results are based on measurements in the near field rather than in the far field, as with the other transducers. Nevertheless, the agreement between results from transducers differing by an order of magnitude in frequency indicates that the experimental apparatus and procedure are robust. Many previous experiments in air have been used to develop an empirical formula for the acoustic attenuation in air at a variety of temperatures and humidities.¹ The empirical model accounts for classical attenuation and attenuation from relaxation of vibrational and rotational modes. The curves in Fig. 5 indicate both the classical component only (solid curve) and the full form of the empirical fit for air (dashed curve). The difference between these curves comes primarily from the relaxation of rotational modes in the air. The increase in $\alpha\lambda$ at higher frequencies agrees well with the classical model, although the empirical model fits the data even better. The negative attenuation evident at lower frequencies arises from the diffraction correction and the error in the experiments. Comparable anomalous negative attenuation has occurred previously in similar measurements.²² Even so, the empirical model matches the data quite well.

B. Oxygen

Pure oxygen and mixtures of oxygen and nitrogen were also tested. Figure 6 shows the data for pure oxygen, while Fig. 7 shows the results for various mixtures of oxygen and nitrogen. The relaxation frequency of oxygen has been found to be at about 3 Hz at atmospheric pressure,²³ which is below our measurement capability. Our data agrees well with the classical calculations indicated by the curves in the figures. These plots provide more evidence of the reliability of the experimental technique. The model calculations from Dain and Lueptow³ are consistent with the observation that the

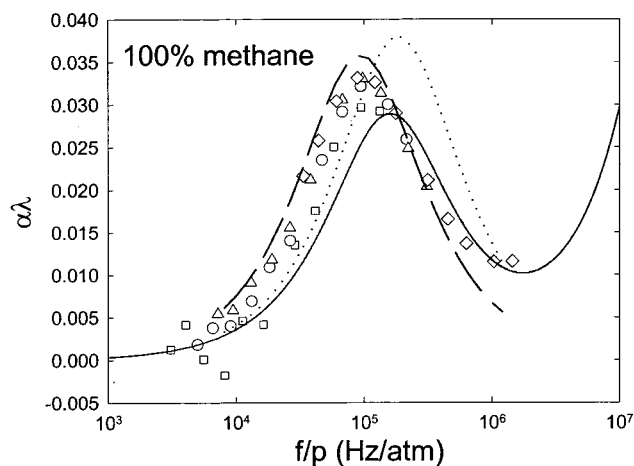


FIG. 8. Results for 100% methane with the Pinkerton diffraction correction at an average temperature of 293.9 K. Data points are for the 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz transducers (squares, circles, triangles, and diamonds, respectively). The solid curve is the calculations based on the vibrational relaxation model of Dain and Lueptow (Ref. 3) summed with the classical attenuation. The dotted curve represents the experimental results from Gravitt *et al.* (Ref. 9), and the dashed curve the results from Edmond and Lamb (Ref. 24).

classical attenuation should be the dominant factor with negligible relaxational attenuation in mixtures of oxygen and nitrogen.

C. Methane

The study of methane provides a case where a strong relaxation frequency is present in the range of frequencies measured. The attenuation spectrum for 100% methane is shown in Fig. 8. Again the different transducers agree fairly well. The attenuation for 100% methane has been studied previously. Studies by Edmonds and Lamb²⁴ and by Gravitt, Whetstone, and Lagemann⁹ were carried out using acoustic resonance tubes in a similar frequency range and at similar

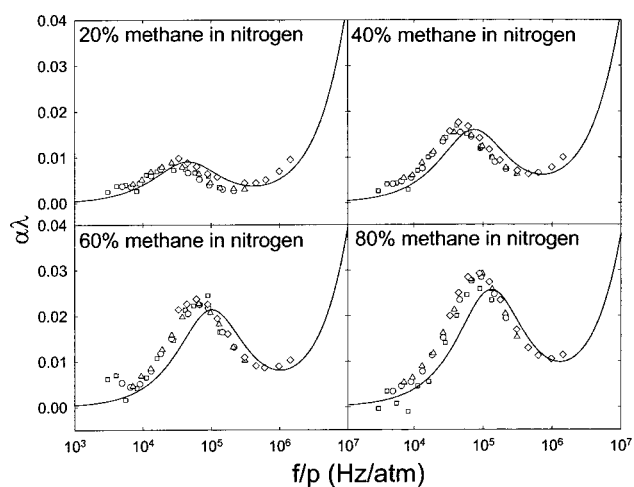


FIG. 9. Results for mixtures of 20%, 40%, 60%, and 80% methane in nitrogen with the Pinkerton diffraction correction at average temperatures of 293.2, 293.0, 293.4, and 295.0 K, respectively. Data points are for the 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz transducers (squares, circles, triangles, and diamonds, respectively). The solid curves are the calculations based on the vibrational relaxation model of Dain and Lueptow (Ref. 3) summed with the classical and diffusional attenuation.

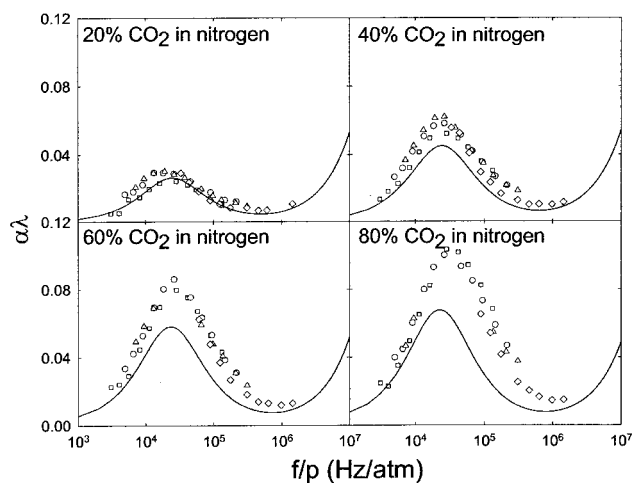


FIG. 10. Results for mixtures of 20%, 40%, 60%, and 80% CO₂ in nitrogen with the Pinkerton diffraction correction at average temperatures of 292.6, 293.7, 293.5, and 294.0 K, respectively. Data points are for the 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz transducers (squares, circles, triangles, and diamonds, respectively). The solid curves are the calculations based on the vibrational relaxation model of Dain and Lueptow (Ref. 3) summed with the classical and diffusional attenuation.

temperatures. Fits to these measurements are included in Fig. 8 as dashed and dotted curves, respectively. Our measurements agree better with the results of Edmond and Lamb. The variation in attenuation between researchers may result from the sensitivity of attenuation in gases to impurities. In particular small concentrations of water can cause a large shift in the relaxation peak of the attenuation spectrum in other gases.^{1,25,26} The calculations of Dain and Lueptow³ (solid line) agree fairly well with the measurements, although the amplitude of the attenuation peak is slightly lower than the measured peak.

The attenuation spectrum for mixtures of methane and nitrogen are shown in Fig. 9. The results from the different transducers at the same f/p overlap quite well. In addition, the Dain and Lueptow model predicts the attenuation spectrum very accurately. The relaxation peak shifts to a higher frequency and increases in magnitude as the fraction of methane increases. This is related to the increasing dominance of methane relaxation modes as the fraction of methane increases.

D. Carbon dioxide

The attenuation spectra for mixtures of carbon dioxide in nitrogen are shown in Fig. 10. The different transducers again give fairly similar results. Like the relaxation frequency in methane, the amplitude of the peak increases with increasing concentration of carbon dioxide, but the shift in frequency is much smaller than it is for methane/nitrogen mixtures. Also note that the magnitude of the attenuation is substantially larger for carbon dioxide than for methane.

Carbon dioxide has been studied so extensively that it is used in many textbooks as an example of attenuation in a gas.² Leonard²⁷ and Fricke²⁸ measured the attenuation in pure carbon dioxide at the same frequency and temperature range as the present experiment. Leonard used a pulsed technique to measure the attenuation and found the peak attenu-

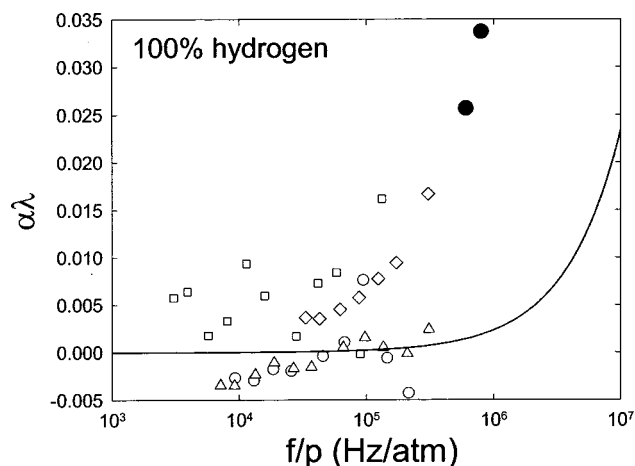


FIG. 11. Results for 100% hydrogen with Pinkerton diffraction correction at an average temperature of 293.8 K. Data points are for the 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz transducers (squares, circles, triangles, and diamonds, respectively). The solid curve is the sum of the classical and diffusional attenuation. The filled circles are data from Winter and Hill (Ref. 30).

ation of $\alpha\lambda = 0.125$ at $f/p = 30$ kHz/atm. Fricke measured attenuation in a cubical resonance chamber²⁹ and found the peak attenuation of $\alpha\lambda = 0.115$ at $f/p = 20$ kHz/atm. For our experiments at 100% carbon dioxide, the attenuation was so strong that the signal arriving at the receiver was too small to be reliably detected even for very small separation distances. However, considering the trend with increasing CO₂ concentration in nitrogen, one could expect 100% CO₂ to have a relaxation peak at a frequency slightly higher than the value of $f/p = 28$ kHz/atm and an amplitude higher than the value of $\alpha\lambda = 0.10$, that we measured for 80% CO₂ in N₂. Furthermore, the slight upward shift in the measured relaxation frequencies and the noticeable increase in the magnitude of the relaxation peaks in Fig. 10 are consistent with previous measurements for small amounts of carbon dioxide in nitrogen.²²

The calculations for carbon dioxide from our model provide a qualitative representation of the experimental results. The similarity of the frequency for the relaxation peak in the attenuation spectrum with increasing CO₂ concentration is fairly well reproduced. However, the model underestimates the amplitude of the relaxation peak, especially for high CO₂ concentrations. This discrepancy could be the result of the linear structure of the carbon dioxide molecule, which was probably not adequately taken into account in the model for the collisional dynamics.

E. Hydrogen

The acoustic attenuation spectrum for 100% hydrogen is shown in Fig. 11. In this case the values for different transducers do not agree as well as for the other gas mixtures. One problem is that the speed of sound in hydrogen (1306 m/s) is much higher than that in any of the other gases. This results in reflections arriving more quickly at the receiver and being more likely to interfere with the attenuation measurements. More noise may also be present in the system from the excitation of the emitting transducer.

Winter and Hill measured the attenuation in pure hydrogen at high frequencies using a pulse technique.³⁰ Their re-

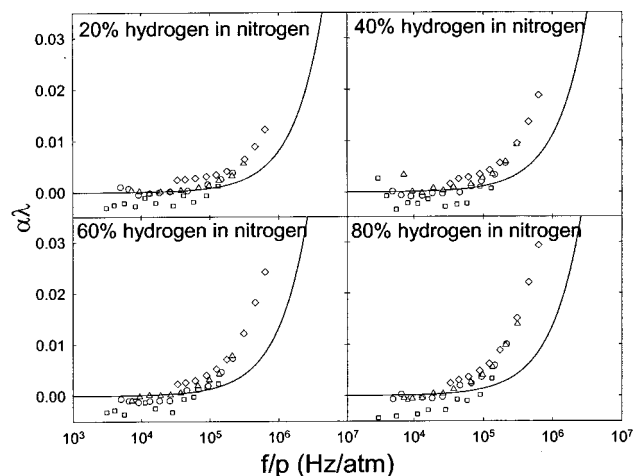


FIG. 12. Results for mixtures of 20%, 40%, 60%, and 80% hydrogen in nitrogen with the Pinkerton diffraction correction at average temperatures of 298.6, 298.2, 297.9, and 297.8 K, respectively. Data points are for the 92 kHz, 149.1 kHz, 215 kHz, and 1 MHz transducers (squares, circles, triangles, and diamonds, respectively). The solid curves are the classical and diffusional attenuation. The vibrational relaxation is negligible.

sults show a large increase in attenuation per wavelength ($\alpha\lambda$) above 1 MHz as indicated by filled circles in Fig. 11. The slight excess energy absorption above the classical case may be related to rotational relaxation in hydrogen with a peak near 20 MHz that affects attenuation above 1 MHz.

Figure 12 shows the attenuation spectrum for mixtures of hydrogen/nitrogen. The scatter observed in the mixtures is similar to that seen for mixtures of oxygen and nitrogen. The calculations for classical attenuation plus the attenuation due to diffusion match the experiments for lower frequencies but fall below the experimental results at high frequencies, especially for higher concentrations of hydrogen. The discrepancy between the calculations and the data is presumably due to the rotational relaxation of hydrogen at high frequencies.

V. CONCLUSIONS

Acoustic attenuation in gases is challenging to measure and predict. Nevertheless, we have been able to measure the attenuation in pure gases and binary gas mixtures. The measurements are in good agreement with previous data. Air and mixtures of oxygen and nitrogen measured using our experimental technique agree well with previous measurements. Our calculations based on the Dain and Lueptow model agree well with experiments for methane and mixtures of methane and nitrogen. In particular, the calculations reproduce the shift in frequency and magnitude of the relaxation peak with increasing methane concentration. Also the calculation for mixtures of carbon dioxide and nitrogen qualitatively agree with the data, showing that the frequency of the relaxation peak in the attenuation spectrum does not shift significantly with increasing CO₂ concentration, while the magnitude does increase. The model also does well in predicting negligible vibrational relaxation for hydrogen. Because the model does not include rotational relaxation, it does not match the measurements for attenuation at higher frequencies in hydrogen. Although the Dain and Lueptow

TABLE II. The collisional diameter (σ), the force constants (ϵ_{LJ}), the vibrational modes, the degeneracies (g), and the vibrational amplitude coefficients for gases that are necessary for the relaxational attenuation calculations with the Dain and Lueptow model (Ref. 3).

Gas	σ (Å)	ϵ_{LJ} (cal mol ⁻¹)	Normal modes of vibration (cm ⁻¹)	g	Vibrational amplitude coefficients (amu ⁻¹)
O ₂	3.548	175	$v = 1554$	1	0.0625
N ₂	3.546	159	$v = 2331$	1	0.0714
CH ₄	3.759	286	$v_1 = 2915$	1	0.9921
			$v_2 = 1534$	2	0.9921
			$v_3 = 3019$	1	0.9923
			$v_4 = 1306$	3	0.8368
H ₂	2.761	75.5	$v = 4160$	1	1.0
CO ₂	3.99	378	$v_1 = 1333$	1	0.05
			$v_2 = 667$	2	0.05
			$v_3 = 2349$	1	0.05

model should be applicable to gas mixtures of three or more components, further testing is necessary to validate the model in these cases.

ACKNOWLEDGMENTS

This work was supported by NASA, DOE, and the Oklahoma Center for the Advancement of Science and Technology. We thank G. Mozurkewich and B. Ghaffari of Ford Motor Company for helpful discussions.

APPENDIX

Acoustic attenuation in a gaseous medium results from viscous dissipation, irreversible heat conduction, diffusion of gas components, and molecular relaxation. The theoretical curves in this paper are based on appropriate models for each of these phenomena. The total attenuation is the sum of the attenuation due to all three mechanisms, each addressed separately below. All calculations were carried out for a temperature of 292 K and a pressure of 1 atm.

Classical attenuation: The attenuation related to viscous dissipation and irreversible heat conduction are based on the classical formulation by Stokes and Kirchhoff² using the shear viscosity, the thermal conductivity, and the specific heats for the gas mixtures calculated according to commercial software.³¹

Diffusional attenuation: The attenuation due to diffusion of the gas components is calculated using Eq. 14.3.19 in Bhatia² with the mutual diffusion coefficients from Lide,³² and neglecting thermal diffusion. For the gas mixtures studied here, attenuation from diffusion is negligible although we have included it in the calculations. An exception occurs in the case of mixtures of hydrogen and nitrogen.

Relaxational attenuation: The theory of relaxational attenuation for vibrational modes of a polyatomic gas mixture is based on the acoustic equations for a plane wave process in a continuous medium accompanied by semimacroscopic population equations for the number of gas molecules in a given energy state.³ The energy equation in the formulation represents the fluctuation of the total energy of the gas mixture including internal molecular energy states. These equations yield a linear system of ordinary differential equations describing multiple relaxation processes in gas mixtures. The equations, presented for a three-component mixture by Dain

and Lueptow,³ can be easily generalized for the mixtures under consideration in the present paper. We assume that energy exchange occurs between all vibrational modes.

Table II provides collisional diameter (σ), the force constants (ϵ_{LJ}), the vibrational modes, the degeneracies (g), and the vibrational amplitude coefficients for oxygen, nitrogen, water, methane, hydrogen, and carbon dioxide, that are necessary for the model. The normal modes are expressed using the spectroscopic convention in terms of inverse wavelength.

We assume that at room temperature only one-quantum collisional reactions are possible. The calculation of transition probabilities $P_{0 \rightarrow 0}^{1 \rightarrow 0}(j, k)$ and $P_{0 \rightarrow 1}^{1 \rightarrow 0}(j, k)$ is based on the approximate formulas of transition probabilities for polyatomic gases derived by Tanzcos.⁶ The depth of the potential well ϵ_{LJ} and the vibrational amplitude coefficients are provided in Lambert.³³ The vibrational amplitude coefficients for carbon dioxide were adjusted using known vibration relaxation frequency and experimental attenuation curve for pure CO₂. The adjustable values of the collision diameter in the Lennard-Jones potential related to the gases of interest were obtained by the method outlined by Hirshfelder *et al.*³⁴ The temperature dependent collision diameter for low energy collisions, σ , was calculated using the kinetic theory formula (Eqs. 8.4–8.5 of Ref. 34) for the gas viscosity data and tabulated values of the kinetic integral.³⁴ The attenuation due to rotational relaxation was not included, because its contribution is quite similar to the classical contribution at the frequencies considered,^{1,35} and except for hydrogen/nitrogen mixtures, it is small in the range of frequencies for which measurements were made.

¹H. E. Bass, L. C. Sutherland, J. Piercy, and L. Evans, "Absorption of sound by the atmosphere," in *Physical Acoustics*, edited by W. P. Mason (Academic, Orlando, 1984), Vol. XVII, pp. 145–232.

²A. B. Bhatia, *Ultrasonic Absorption* (Dover, New York, 1984).

³Y. Dain and R. M. Lueptow, "Acoustic attenuation in three-component gas mixtures—Theory," *J. Acoust. Soc. Am.* **109**, 1955–1964 (2001).

⁴Y. Dain and R. M. Lueptow, "Acoustic attenuation in a three-gas mixture: Results," *J. Acoust. Soc. Am.* **110**, 2974–2979 (2001).

⁵M. N. Kogan, *Rarefied Gas Dynamics* (Plenum, New York, 1969).

⁶F. Tanzcos, "Calculation of vibrational relaxation times of the chloromethanes," *J. Chem. Phys.* **25**, 439–447 (1956).

⁷K. F. Herzfeld and T. H. Litovitz, *Absorption and Dispersion of Ultrasonic Waves* (Academic, New York, 1959).

- ⁸W. Tempest and H. D. Parbrook, "The absorption of sound in argon, nitrogen and oxygen," *Acustica* **7**, 354–362 (1957).
- ⁹J. C. Gravitt, C. N. Whetstone, and R. T. Lagemann, "Thermal relaxation absorption of sound in the deuterated methanes at 26 °C," *J. Chem. Phys.* **44**, 70–72 (1966).
- ¹⁰T. G. Winter and G. L. Hill, "High-temperature ultrasonic measurements of rotational relaxation in hydrogen, deuterium, nitrogen and oxygen," *J. Acoust. Soc. Am.* **42**, 848–858 (1967).
- ¹¹L. B. Evans, "Vibrational relaxation in moist nitrogen," *J. Acoust. Soc. Am.* **51**, 409–411 (1972).
- ¹²M. C. Henderson, "Vibrational relaxation in nitrogen and other gases," *J. Acoust. Soc. Am.* **34**, 349–350 (1962).
- ¹³H.-J. Bauer and R. Schotter, "Collision transfer of vibrational energy from nitrogen and methane to the carbon dioxide molecule," *J. Chem. Phys.* **51**, 3261–3270 (1969).
- ¹⁴D. Telfair and W. H. Pielemeier, "An improved apparatus for supersonic velocity and absorption measurements," *Rev. Sci. Instrum.* **13**, 122–126 (1942).
- ¹⁵J. Krautkrämer and H. Krautkrämer, *Ultrasonic Testing of Materials* (Springer-Verlag, New York, 1990), pp. 58–92.
- ¹⁶J. M. M. Pinkerton, "On the pulse method of measuring ultrasonic absorption in liquids," *Proc. Phys. Soc. London* **62**, 286–299 (1949).
- ¹⁷H. W. Lord, W. S. Gately, and H. A. Evensen, *Noise Control for Engineering* (Robert E. Krieger Publishing Co., Malabar, FL, 1980), pp. 71–73.
- ¹⁸A. S. Khimunin, "Numerical calculation of the diffraction corrections for the precise measurement of ultrasound absorption," *Acustica* **27**, 173–181 (1972).
- ¹⁹A. O. Williams, "The piston source at high frequencies," *J. Acoust. Soc. Am.* **23**, 1–6 (1951).
- ²⁰P. H. Rogers and A. L. Van Buren, "An exact expression for the Lommel diffraction correction integral," *J. Acoust. Soc. Am.* **55**, 724–728 (1974).
- ²¹Y. Dain and R. M. Lueptow, "Diffraction and attenuation of a tone burst in a mono-relaxing medium," *J. Acoust. Soc. Am.*, to appear (2003).
- ²²A. J. Zuckerwar and W. A. Griffin, "Resonant tube for measurement of sound absorption in gases at low frequency/pressure ratios," *J. Acoust. Soc. Am.* **68**, 218–226 (1980).
- ²³J. G. Parker and D. N. Ritke, "Vibrational relaxation times of oxygen at high pressure," *J. Acoust. Soc. Am.* **52**, 1380–1384 (1972).
- ²⁴P. D. Edmonds and J. Lamb, "Vibrational relaxation times of a number of polyatomic gases derived from measurements of acoustic absorption," *Proc. Phys. Soc. London* **72**, 940–948 (1958).
- ²⁵V. O. Knudsen and E. Fricke, "The absorption of sound in CO₂, N₂O, COS, and in CS₂, containing added impurities," *J. Acoust. Soc. Am.* **12**, 255–259 (1940).
- ²⁶M. C. Henderson, K. F. Herzfeld, J. Bry, R. Coakley, and G. Carriere, "Thermal relaxation in nitrogen with wet carbon dioxide as impurity," *J. Acoust. Soc. Am.* **45**, 109–114 (1969).
- ²⁷R. W. Leonard, "The absorption of sound in carbon dioxide," *J. Acoust. Soc. Am.* **12**, 241–244 (1940).
- ²⁸E. F. Fricke, "The absorption of sound in five triatomic gases," *J. Acoust. Soc. Am.* **12**, 245–254 (1940).
- ²⁹V. O. Knudsen and E. F. Fricke, "The absorption of sound in carbon dioxide and other gases," *J. Acoust. Soc. Am.* **10**, 89–97 (1938).
- ³⁰T. G. Winter and G. L. Hill, "High-temperature ultrasonic measurements of rotational relaxation in hydrogen, deuterium, nitrogen, and oxygen," *J. Acoust. Soc. Am.* **42**, 848–858 (1967).
- ³¹Physical Property Data Service (PPDS2) for Windows (National Engineering Laboratory, Glasgow, or Technical Database Services, Inc., New York City, 1998).
- ³²D. R. Lide, *Handbook of Chemistry and Physics* (CRC Press, New York, 2001), pp. 6–192, 193.
- ³³J. D. Lambert, *Vibrational and Rotational Relaxation in Gases* (Clarendon, Oxford, 1977).
- ³⁴J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954).
- ³⁵G. L. Gooberman, *Ultrasonics: Theory and Application* (Hart Publishing Co., Inc., New York, 1968), pp. 113–138.

Scattering of the fundamental shear horizontal mode from steps and notches in plates

A. Demma, P. Cawley, and M. Lowe

Department of Mechanical Engineering, Imperial College, London SW7 2BX, United Kingdom

(Received 28 May 2002; revised 20 December 2002; accepted 30 December 2002)

The scattering of the SH0 mode from discontinuities in the geometry of a plate has been studied. Both finite element and modal decomposition methods have been used to study the reflection and transmission characteristics from a thickness step in a plate, obtaining very good agreement. The significance of nonpropagating modes in the scattering from steps in plates has been specifically investigated. A method to approximate the reflection from rectangular notches by superimposing the reflection from a step down (start of the notch) and a step up (end of the notch) has been proposed. It is demonstrated that it is possible to use this method to obtain the reflection from a notch of any depth and at any frequency. The effect of frequency on the reflection from notches has been examined. The limits of this method in approximating cracklike defects have also been studied.
© 2003 Acoustical Society of America. [DOI: 10.1121/1.1554694]

PACS numbers: 43.35.Cg, 43.20.Fn [YHB]

I. INTRODUCTION

The detection of corrosion defects in plates is an important topic. Several nondestructive testing (NDT) techniques are currently used for the detection and sizing of defects in plates. Among them, ultrasonic techniques offer good flexibility due to the numerous test configurations available depending on the specific problem. Conventional ultrasonic tests inspect the region of structure immediately adjacent to the transducer, therefore limiting the single test range to a small area. Alternatively, it is possible to use guided waves, which can potentially propagate along a plate structure with minimal attenuation of the ultrasonic signal, to increase the single test range. Moreover, guided waves produce stresses through the thickness of the plate so enabling the examination of the entire thickness of the plate.

Practical ultrasonic guided wave testing of plates can be done using a simple pulse-echo arrangement with a single transducer which generates the desired wave and receives the echoes caused by the defects present in the structure. Using this test configuration a line is monitored. It is also possible to create a rapid scan of plate structures by using an array of transducers positioned at one location on the plate to be tested.¹ However, guided wave testing is complicated by the nature of the modes. At any given frequency there are at least two modes and they can be dispersive, so in the most general case the ultrasonic signal will be multimodal and dispersive. It is possible to generate a single mode signal but this can be converted to a multimode signal when interacting with discontinuities in the structure. It is also wise to use the single mode signal in a nondispersive frequency region so that the signal is not distorted as it travels along the structure. Therefore an optimization of the guided wave inspection technique is commonly required.²⁻⁴

In general, two different families of guided waves can exist in a free isotropic plate: Lamb waves characterized by the fact that the particle displacement is in the plane of propagation and SH waves in which the particle displacement is perpendicular to the plane of propagation. SH waves can be very useful in practical testing because there is no

mode conversion to Lamb waves when the polarization is along the long defect dimension. SH waves are also very similar to torsional waves in pipes so in most cases the results obtained for SH waves in plates can be extended to torsional waves in pipes. This is of great interest especially because torsional waves are established in use in practical pipe testing.⁵

The subject of this paper is to examine the reflection of the fundamental SH0 mode from thickness changes, notches, and cracks in plates. The interaction of guided elastic waves in plates and pipes with discontinuities such as thickness steps has been already studied.⁶⁻⁸ Koshiba *et al.*⁶ proposed a combined finite element and analytical technique for the analysis of the scattering of SH waves by steps in plates. Ditre⁷ studied the scattering of guided SH waves from steps in plates in terms of energy reflection and he also explained the multimode reflection when more than one propagating mode can potentially travel along the waveguide. Engan⁸ investigated torsional wave scattering from diameter changes in a circular solid rod using the modal decomposition method. Several researchers have studied the reflection characteristics of guided waves from notches in plates and pipes.^{6,9-13} Both Lamb waves^{9,10} and SH waves^{6,11} have been investigated in terms of their interaction with notches in plates. The examination of the reflection coefficient as a function of the notch width has identified the important phenomenon of the interference between the reflection from the two sides of a notch.^{9,10,13}

The aim of this paper is to investigate the possibility of predicting the reflection from a notch by superimposing the reflections from a thickness step down (start of the notch) and a thickness step up (end of the notch). The significance of nonpropagating modes in the reflection from discontinuities and the relationship between crack and notch reflections are also discussed.

In this paper both finite element and modal decomposition methods are used to investigate the effect of a thickness step on the propagation of the fundamental shear horizontal mode, the modal decomposition enabling a fuller understanding of the effect of nonpropagating modes on the scat-

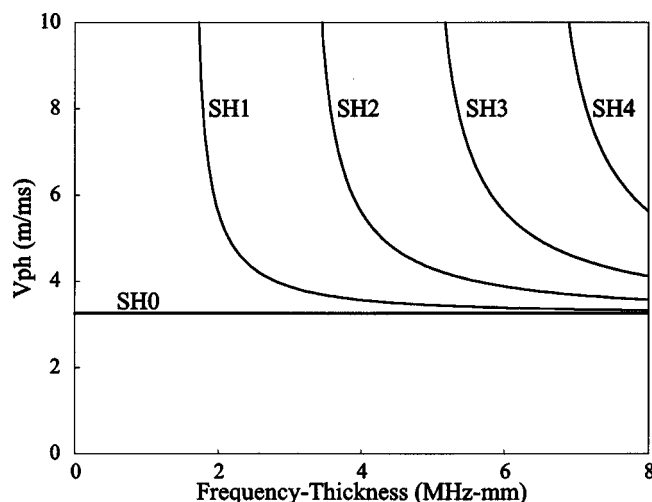


FIG. 1. Phase velocity dispersion curves for a steel plate in a vacuum. Only shear horizontal propagating modes are traced.

tering characteristics. The modulus and phase of the reflection coefficient at a step-down (when the thickness decreases) and step-up (when the thickness increases) were used to simulate the signal reflected from a rectangular notch. Therefore we reproduced the interference phenomenon between the reflection from a step-down and a step-up, and we then compared our predictions with the finite element results obtained for notch cases with varying axial extent. The effect of frequency on the reflection from geometrical discontinuities in plates has also been considered. Finally, the limits of the method proposed in this paper to predict the reflection coefficient from notches have been examined. In particular, a study of the reflection from notches when the axial extent tends to zero (cracklike case) has been performed, and a verification of the convergence of the notch case to the crack case as the axial extent decreases has been shown.

II. SH WAVES IN PLATES

Shear horizontal (SH) waves in an isotropic plate in vacuum can be expressed explicitly as suggested by Auld.¹⁴ Figure 1 shows the phase velocity dispersion curves for a steel plate in vacuum. The curves scale linearly with frequency and thickness so that the use of the frequency-thickness scale on the abscissa allows these curves to be used for a plate of any thickness. They were calculated using the program Disperse,¹⁵ developed at Imperial College. SH waves can be either symmetric or antisymmetric but we did not differentiate the two families of SH modes in this paper and simply used a counter variable to distinguish the different modes. The modes with even counter variable are symmetric and the modes with odd counter variable are antisymmetric. The fundamental SH mode, existing at zero frequency, is the SH0 (symmetric) mode. The properties of this mode are not frequency dependent: it is completely non-dispersive at all frequencies and its phase velocity is the bulk shear velocity. The next mode appearing is the SH1 (antisymmetric) mode and its cut-off frequency is at about 1.6 MHz mm.

This study considered an incident SH0 mode in all cases and the region under examination was 0–0.55 MHz mm (e.g., 0–100 kHz for a 5.5 mm plate). This relatively low frequency range is of particular interest to the authors because of the immediate generalization of this study to the case of pipes with notches where a low frequency torsional wave is used in order to have a long range for a single test (from 10 to 100 m depending on the application).⁵ Moreover, it is of general interest to study the effects of nonpropagating modes on the reflection from defects and this is best viewed when only one propagating mode can exist in the wave guide.

A nonpropagating mode can be thought of as a local vibration of the structure which does not propagate away from the point where it is generated. Partial wave theory gives extra physical insight on how to relate the phenomenon of nonpropagating modes to the frequency of the wave.¹⁴ In fact, if we consider a propagating mode and we decrease the frequency, the angle of incidence of the partial wave decreases, becoming zero at the cut-off frequency. At the cut-off frequency, the partial waves simply reflect back and forth across the thickness of the waveguide and there is no variation of the stress and displacement field along the direction of propagation. The wave number, which is real in a propagating SH mode, is zero at the cut-off and becomes a purely imaginary number at lower frequencies. Consequently, the displacement field will be different below and above the cut-off frequency. SH modes have only one component of displacement. The direction of propagation is here defined as z , the particle displacement is in the y direction and x is the through thickness direction. Therefore the displacement field can be described by:

$$u_y = A_y e^{i(k_{\text{real}}(z - \omega t))} e^{-k_{\text{imag}} z}, \quad (1)$$

where A_y is the amplitude of the displacement, i is the imaginary unit ($i = \sqrt{-1}$), k_{real} and k_{imag} are, respectively, the real and imaginary parts of the wave number, ω is the circular frequency, z is the direction of propagation, and t is the time. From Eq. (1) it is clear that the displacement, which is a sinusoidal wave when the wave number is real, becomes a decreasing exponential curve for a nonpropagating SH mode¹⁴ so its effect decreases exponentially with distance from the point where the nonpropagating mode is localized.

Figure 2 shows the attenuation curves for the nonpropagating modes in a plate in a vacuum. An infinite number of nonpropagating modes exist at any given frequency, but only the first five of them are plotted in Fig. 2. The ordinate is the attenuation (in dB mm/m) of the decay function and it is given by:

$$\text{attn} = t \cdot 20 \log \frac{1}{e^{-k_{\text{imag}}}}, \quad (2)$$

where t is the thickness of the plate and k_{imag} is the imaginary wave number in Eq. (1). The attenuation of each mode decreases as the frequency increases and it becomes zero at the cut-off frequency. Moreover, at any frequency value the attenuation increases with the counter variable of the SH mode. The SH0 mode is not traced in Fig. 2 because it is the

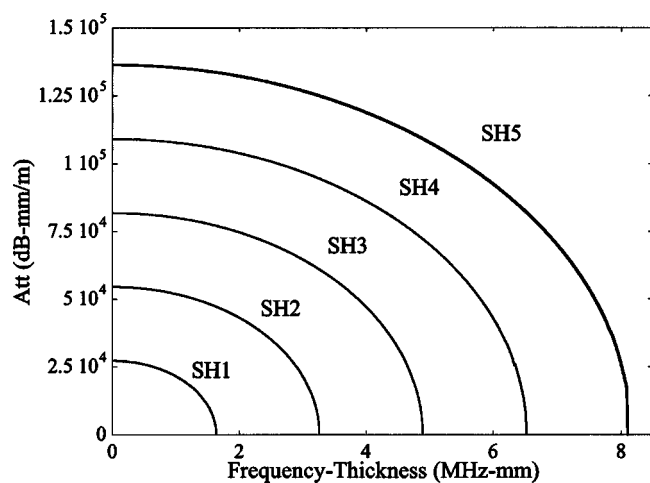


FIG. 2. Attenuation curves for a steel plate in a vacuum. Only shear horizontal nonpropagating modes are traced.

only mode which has real wave number at all frequencies. Another important characteristic of the SH0 mode is that its displacement is constant through the thickness [see Fig. 3(a)]. Figure 3(b) shows the mode shape of the antisymmetric SH1 mode. The mode shapes of the higher order modes are characterized by an increasing number of zero crossings through the thickness (zero for SH0, one for SH1 and so on).

The stress field of SH modes is also relatively simple. This is characterized by two components of stress:

$$\tau_{yz} = c_{44} \frac{\partial u_y}{\partial z}, \quad \tau_{xy} = c_{44} \frac{\partial u_y}{\partial x}, \quad (3)$$

where c_{44} is the shear modulus. The τ_{yz} stress component is directly proportional to the u_y displacement so its mode shape is the same as the mode shape for u_y . The τ_{xy} component is not relevant in this study.

III. FINITE ELEMENT MODELS

The finite element (FE) method has been extensively and successfully used to study the interaction between guided waves and defects in structures.^{6,9,12,16-21} In general a three-dimensional (3-D) solid model is required to perform a numerical analysis of the interaction between guided waves and discrete defects. However, 3-D models are computationally expensive so when possible we use simplified models.¹⁶

Many studies have been done on Lamb wave interaction with defects using two-dimensional models with the assumption of plane strain.^{9,10,17} However, plane strain elements model displacements which are solely in the assumed plane

of strain. In the case of SH waves, the displacement is normal to the plane, so such an approximation cannot be taken. Nevertheless, an approximate approach using an axisymmetric idealization has been found to work very satisfactorily. The implementation of the approximate method uses a two-dimensional (2-D) axisymmetric model of a pipe with large diameter (a pipe with infinite diameter approximates a plate). The waves travel in the axial direction and the displacements are in the circumferential direction. Thus the SH0 mode in a plate is approximated by an analysis of the T(0,1) mode in a large thin walled pipe. Further details of the use of the 2-D axisymmetric analysis of cylinders with large diameter to approximate the 3-D analysis of wave propagation in plates is given in Ref. 22. The advantage is that the axisymmetric analysis allows displacements in the direction normal to the element (the circumferential direction); this degree of freedom is included in most implementations of axisymmetric elements in order to allow analyses of problems with non-zero circumferential harmonics. Therefore a standard finite element program can be used without the need to write specific code. Both the axial extent and the depth through the thickness of a defect can be varied using this model. The axisymmetric nature of the model implies that the defects extend over the full circumference of the pipe so they are equivalent to notches in a plate that are infinitely long in the direction normal to the propagation.

In all cases a finite length of the system was modeled and a geometrical discontinuity was introduced at some distance along it. The input wave signal was excited by prescribing time-varying tangential (u_y) displacements at one end of the model. Since the SH0 mode shape is constant with frequency, pure mode excitation was obtained by simply imposing the mode shape at the center frequency.²³ The displacement distribution was constant through the thickness in this case [see Fig. 3(a)]. Identically sized linear (four-noded) quadrilateral axisymmetric elements were used. Using finite element analysis we needed to model a plate with a specific thickness but the results obtained from the finite elements can be generalized in terms of frequency-thickness product. We modeled a pipe with 5.5 mm wall thickness and 1.5 m radius and performed a convergence test to verify that this radius was large enough for the system to approximate a plate; it was found that the reflection coefficient from a notch in the 1.5 m radius case was the same as that from a notch of the same depth in a pipe with 4.5 m radius, so confirming that the reflection coefficient had converged to that for a plate.

Three different types of discontinuity were modeled:

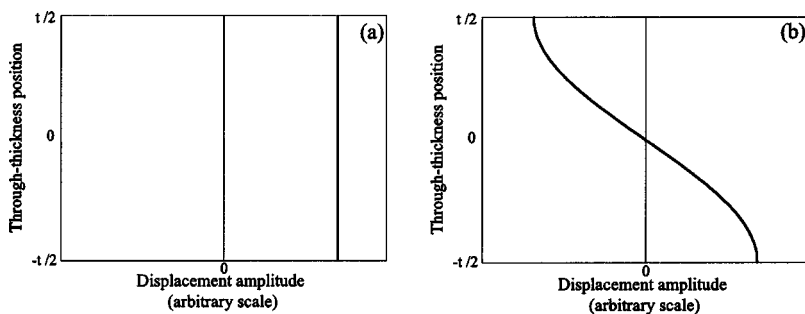


FIG. 3. Displacement mode shapes in a plate for SH0 (a) and SH1 (b).

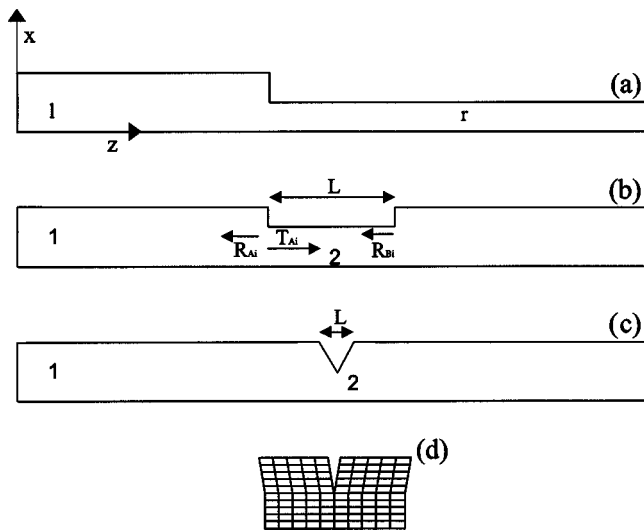


FIG. 4. (a) Schematic of step; (b) rectangular notch; (c) V-notch; (d) detail of finite element mesh for V-notch.

- (1) thickness step [see Fig. 4(a)]. This was used to study the scattering behavior when there is a change in the thickness of the waveguide. From this model we derived modulus and phase of both reflection and transmission at a thickness step. We modeled both the case in which the thickness decreases (step down) and the case in which the thickness increases (step up). The information derived from this set of models was further used to simulate the reflection from a rectangular notch.
- (2) rectangular notch [see Fig. 4(b)]. This model was employed to study the effect of both frequency and axial extent on the reflection from notches. The results obtained from this model were also used to validate the reconstruction of the reflection from a rectangular notch superimposing the reflections from down and up steps.
- (3) V-notch [see Figs. 4(c) and (d)]. This model was used to study the relationship between crack and notch reflections, the rectangular notch having the limitation of being impractical to use to simulate notches with very small axial extent (this is explained further below).

We now present two subsections in which we describe the FE model for thickness steps and notches, respectively.

A. FE model for thickness step

The thickness step was modeled by creating two regions with different thickness but keeping the dimensions of the elements constant. Only rectangular steps were modeled. The models represented a 1 m length using 2000 elements along

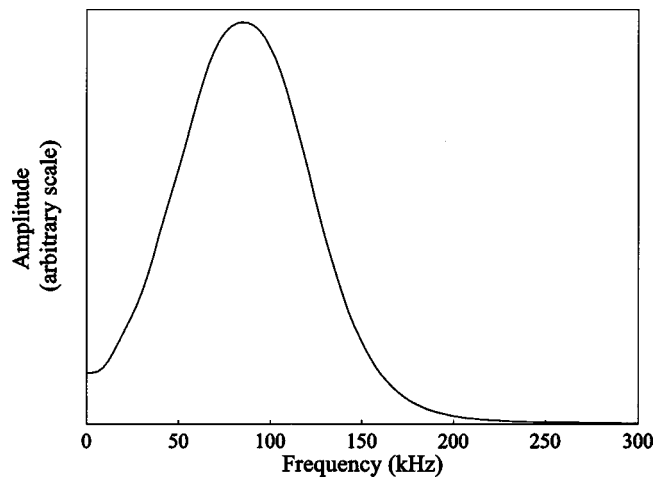


FIG. 5. Frequency spectrum of Hanning windowed linearly chirped toneburst at 100 kHz center frequency.

the length and 10 elements through the thickness. A convergence study showed that the discretization used was satisfactory. A summary of the FE models that we studied for the step cases is in Table I. The steps were placed at 0.6 m from the excitation point and the monitoring points were at 0.4 m and 0.8 m from the excitation point so the total path of the reflected (first monitoring point) and transmitted signal (second monitoring point) was 0.8 m in all cases. The excitation was a 10 cycle, Hanning windowed, linear chirp at a center frequency of 100 kHz where the carrier signal $u(t)$ of duration T , center frequency ω_c , and amplitude A can be described by:

$$u(t) = A \sin\left(\frac{2\pi\omega_c t^2}{T}\right) \left[1 - \cos\left(\frac{2\pi t}{T}\right)\right]. \quad (4)$$

This gave more energy at low frequency (see typical spectrum signal in Fig. 5) compared to a similar windowed toneburst. A typical time record from the simulation is also shown in Fig. 6. This shows the incident signal on its way toward the step defect and then the reflected SH0 mode. The time trace in Fig. 6 is for the 50% step down model.

Both modulus and phase of the reflected and transmitted signals were calculated. The modulus was obtained by dividing the amplitude of the reflected and transmitted signals by the amplitude of the incident signal in the frequency domain. An extra FE model was run to obtain a phase reference. This was a model of a 1 m long, 5.5 mm thick plate without defects which was excited as in the case of the plate with a step defect. The signal was monitored at 0.8 m from the excitation point so we had a reference signal that had trav-

TABLE I. Summary FE models for thickness step.

Model	Thickness of left region (mm)	Number elements through thickness left region	Thickness of right region (mm)	Number elements through thickness right region
Step down 20%	5.5	10	4.4	8
Step up 20%	4.4	8	5.5	10
Step down 50%	5.5	10	2.75	5
Step up 50%	2.75	5	5.5	10

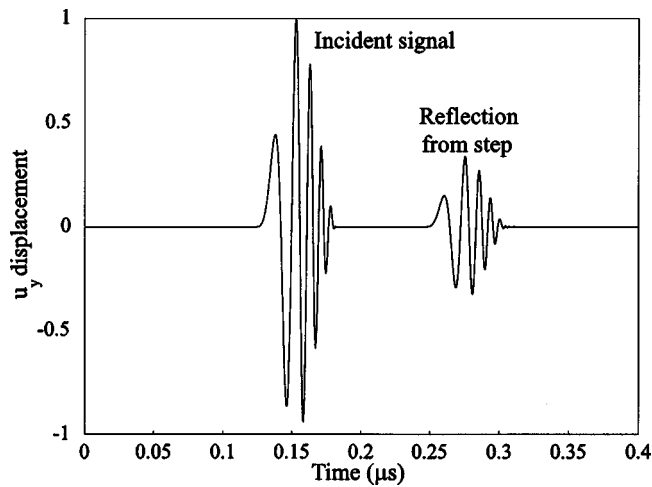


FIG. 6. Predicted time record for a 5.5 mm plate with 50% thickness step and SH0 mode incident at 100 kHz center frequency.

eled along the same path length as the reflected and transmitted signals in the plate with a step. Using a simple FFT algorithm it was possible to obtain the phase of the signals in the frequency domain. The phase of the reference signal was subtracted from the phase of the signals reflected and transmitted at a step so giving the phase shift caused by the scattering at the step.

B. FE model for notch

In both rectangular notch and V-notch models, a length of the plate was modeled and a notch was introduced at some distance along it. The input wave was excited by prescribing a 10 cycle, Gaussian windowed toneburst with 100 kHz center frequency at one end of the plate. The tangential displacements (u_y) were monitored at a location between the excitation end and the defect, so both the incident wave and the reflected wave from the discontinuity were detected.

The models represented a 2.4 m long, 5.5 mm thick plate using 2000 elements along the length and 10 elements through the thickness. A finer mesh with 20 elements through the thickness and 6000 elements along the length was also studied and gave almost identical results, indicating that convergence was satisfactory. Only the modulus of the reflection was calculated for the notch case and this was done exactly as in the step case. As already mentioned, two types of notch were studied:

- (1) rectangular notch
- (2) V notch.

The rectangular notches [see Fig. 4(b)] were created by removing elements. We studied two different defect depths (20% and 50% of the thickness). Several FE models with different axial extent have been studied to investigate the effect of axial extent on the reflection. The axial extent of the elements (inside and outside the notch) was kept constant in order to avoid numerical impedance differences. Therefore a fine mesh was needed to obtain the reflection coefficient at relatively small notch axial extent. Zero-length notches (crack case) were also modeled and this was done by disconnecting adjacent elements; thus although the elements on

each side of the notch had nodes at coincident locations, they were not connected.¹² We also wanted to verify the convergence of the notch to the crack when the axial extent is small. A reduction of the axial extent of the notch using the mesh for the rectangular notch implies the use of elements with shorter length and consequently a refinement of the mesh. This is computationally expensive and can become simply impossible at very small defect axial extents. Therefore, in order to verify the convergence of the notch to the crack case, we used a V-notch model.

Using the V-notch model, we reduced the axial extent of the notch by simply decreasing the distance between the opposite faces of the notch. The V-notch was modeled with a 50% defect depth and varying axial extent L [see Fig. 4(c)]. In this case we used the same number of elements along the length and through the thickness in all of the models, the shape and dimension of each element depending on the extension of the V-notch. The shape change of each element from one model to another was small enough to neglect its effects on the accuracy of the results [see Fig. 4(d)]. We investigated the effect of changing the axial extent L of the head of the notch, and we checked the convergence of this model to the crack case when the axial extent was decreased.

IV. MODAL DECOMPOSITION

The presence of a step (up or down) in the thickness of a plate causes scattering and consequent reflection and transmission of the incident signal. In this work we considered a case in which SH0 was incident on a step discontinuity and only SH0 could potentially propagate in the plate (SH0 excited at a frequency lower than the SH1 cut-off). As a first approximation for the scattering problem we assume that only the SH0 mode contributes to the reflection and transmission of the wave. Under this assumption the magnitude of the reflection coefficient would be constant with frequency. The value of reflection obtained at a step down in a plate in this case would be:

$$R_{A1} = \frac{1 - \alpha}{1 + \alpha}, \quad (5)$$

in which $\alpha = (t_2/t_1)$, where t_1 is the larger thickness and t_2 is the smaller thickness (i.e. $\alpha < 1$ for both down and up steps). The transmission coefficient past the step would be:

$$T_{A1} = \frac{2}{1 + \alpha}. \quad (6)$$

From a step up in a plate we would have:

$$R_{B1} = \frac{\alpha - 1}{1 + \alpha}, \quad (7)$$

$$T_{B1} = \frac{2\alpha}{1 + \alpha}. \quad (8)$$

Clearly, only the magnitude of the reflection and transmission has been considered here. Equations (5)–(8) are the ratios between the amplitude of the signal after the interaction with the step discontinuity and the amplitude of the incoming signal. When the signal is transmitted through a step discon-

tinuity, the thickness of the waveguide changes so a power normalization procedure is required to obtain a power normalized transmission coefficient. The power normalization procedure is in general needed when comparing two modes in the same waveguide or one mode in two different waveguides as explained by Auld²⁴ and Pavlakovic²³ (the power normalization is not necessary in the reflection case because the thickness of the waveguide does not change). In the case described above, the power normalized transmission coefficients can be obtained by multiplying the transmission coefficients in Eqs. (6) and (8) by $\sqrt{\alpha}$ and $1/\sqrt{\alpha}$, respectively. In this simplified theory, the phase of the reflection and transmission coefficient is zero in all cases except in the step up reflection case where there is a π phase shift which is accounted for by the sign reversal in the numerator between Eqs. (5) and (7).

However, the assumption that only one mode contributes to the scattering phenomenon at a step does not fulfill the boundary conditions and, in order to obtain an exact solution of the scattering problem, all of the propagating and non-propagating modes should be included in the solution. In practice this is impossible but it has been demonstrated that it is possible to obtain accurate solutions for scattering problems when considering a finite (and reasonable) number of modes.^{8,25,26} Le Clezio *et al.*²⁵ used modal analysis to investigate the interaction of the A_0 and S_0 Lamb waves with vertical cracks in plates. Vogt *et al.*²⁶ applied the modal solution to the scattering of longitudinal and torsional waves at a point when a free waveguide enters an embedding material. Engan⁸ used modal decomposition to study the scattering of torsional waves in a circular rod with a free surface that has an abrupt diameter change. The problem discussed by Engan is very similar to the problem of steps in plates when SH0 is incident. Therefore in this paper we used the approach proposed by Engan to tackle our scattering problem.

We consider here a plate with a discontinuity in thickness, changing from t_l on the left side to t_r on the right side [see Fig. 4(a)]. In this paper subscripts l and r will be used to refer to the left and right sides of the waveguide. In the most general case SH waves are propagating in both directions on either side of the step. The incident modes have amplitude a and the scattering propagating and nonpropagating modes have amplitude b . It is then possible to define the scattering problem using matrix notation:

$$\begin{bmatrix} b_l \\ b_r \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_l \\ a_r \end{bmatrix}. \quad (9)$$

If we consider n modes, the scattering matrix can be described by:

$$\begin{bmatrix} b_{l1} \\ b_{l2} \\ \vdots \\ b_{ln} \\ b_{r1} \\ b_{r2} \\ \vdots \\ b_{rn} \end{bmatrix} = \begin{bmatrix} [S_{11}] & [S_{12}] \\ [S_{21}] & [S_{22}] \end{bmatrix} \begin{bmatrix} a_{l1} \\ a_{l2} \\ \vdots \\ a_{ln} \\ a_{r1} \\ a_{r2} \\ \vdots \\ a_{rn} \end{bmatrix}, \quad (10)$$

where the $[S]$ terms are $n \times n$ matrices and, for example, a_{lj} is the amplitude of the j th incident mode on the left side of the waveguide. We here considered the most general case when there are incoming waves from both sides; the equations were further simplified for the practical implementation.

As already indicated, to solve the scattering problem we need to fulfill the boundary conditions at the step. These are that the stress is zero at the free surface of the discontinuity and both velocity and stress are continuous functions across the interface. As mentioned in Sec. II, the only component of stress which is relevant in our study is the τ_{yz} and this can be derived from the u_y displacement when we know the characteristics of the material [see Eq. (3)]. The velocity can also be derived from the expression for the u_y displacement:

$$v = \frac{du_y}{dt}. \quad (11)$$

Since only one component of stress and displacement will be considered from now on, we omit the subscripts y and z . Therefore the expressions for displacement, stress, and velocity of the incoming wave will be:

$$u_l^a = \sum_j a_{l,j} e^{i(\omega t - k_{l,j} z)},$$

$$\tau_l^a = -i c_{44} \sum_j k_{l,j} a_{l,j} e^{i(\omega t - k_{l,j} z)}, \quad (12)$$

$$v_l^a = i \omega \sum_j a_{l,j} e^{i(\omega t - k_{l,j} z)},$$

where $a_{l,j}$ is the amplitude of the incoming mode of order j in the left part of the structure, $u_{l,j}$ is the power-normalized displacement, ω is the circular frequency and $k_{l,j}$ is the complex wave number. The expressions for u_l^b , τ_l^b , and v_l^b (outgoing wave in the left part) are the same as those for u_l^a , τ_l^a , and v_l^a , respectively, except for the sign in front of the wave number, because they are traveling in the opposite direction from right to left after the scattering (see Engan⁸ for more details). By using the same procedure, it is possible to derive expressions for displacement, stress, and velocity in the right part of the waveguide. Therefore, the boundary conditions mentioned above can be written as:

$$i \omega u_l^a + i \omega u_l^b = i \omega u_r^a + i \omega u_r^b, \quad x < t_r, \quad (13)$$

$$\tau_l^a + \tau_l^b = \tau_r^a + \tau_r^b, \quad x < t_r, \quad (14)$$

$$\tau_l^a + \tau_l^b = 0, \quad x > t_r.$$

Expansion in eigenmodes yields:

$$i \omega \sum_j a_{l,j} u_{l,j} + i \omega \sum_j b_{l,j} u_{l,j} = i \omega \sum_j a_{r,j} u_{r,j} + i \omega \sum_j b_{r,j} u_{r,j}, \quad x < t_r, \quad (15)$$

$$\begin{aligned}
& -ic_{44} \sum_j k_{l,j} a_{l,j} u_{l,j} + ic_{44} \sum_j k_{l,j} b_{l,j} u_{l,j} \\
& = ic_{44} \sum_j k_{r,j} a_{r,j} u_{r,j} - ic_{44} \sum_j k_{r,j} b_{r,j} u_{r,j}, \quad x < t_r, \\
& \quad \quad \quad (16)
\end{aligned}$$

$$-ic_{44} \sum_j k_{l,j} a_{l,j} u_{l,j} + ic_{44} \sum_j k_{l,j} b_{l,j} u_{l,j} = 0, \quad x > t_r.$$

These equations already contain the solution for the scattering problem but we need to simplify the expressions in order to obtain a simple matrix form which can easily be implemented in a computer algorithm. In particular, since the modes considered in the computation were power normalized, it is wise to obtain an explicit expression for the power flow.⁸

From Eq. (15), taking the complex conjugate, multiplying by $\frac{1}{2}ic_{44}k_{r,m}u_{r,m}$ and integrating through the thickness of the plate, we obtain:

$$\begin{aligned}
& \frac{1}{2} \sum_j \omega c_{44} k_{r,m} a_{l,j}^* \int_0^{t_r} u_{l,j} u_{r,m} dt \\
& + \frac{1}{2} \sum_j \omega c_{44} k_{r,m} b_{l,j}^* \int_0^{t_r} u_{l,j} u_{r,m} dt \\
& = \frac{1}{2} \sum_j \omega c_{44} k_{r,m} a_{r,j}^* \int_0^{t_r} u_{r,j} u_{r,m} dt \\
& + \frac{1}{2} \sum_j \omega c_{44} k_{r,m} b_{r,j}^* \int_0^{t_r} u_{r,j} u_{r,m} dt. \quad (17)
\end{aligned}$$

Introducing the normalized wave number $Q_{r,m} = (k_{r,m}/k_0)$ and simplifying the expression, we obtain:

$$Q_{r,m}^* \sum_j L_{j,m} a_{l,j} + Q_{r,m}^* \sum_j L_{j,m} b_{l,j} = P_{r,m}^* a_{r,m} + P_{r,m}^* b_{r,m}, \quad (18)$$

where k_0 is the wave number of the SH0 mode, $L_{j,m} = (c_{44}/2)\omega k_0 \int_0^{t_r} u_{l,j} u_{r,m} dt$ and $P_{r,m}$ is the power flow which is equal to one for propagating modes and $(-i)$ for non-propagating modes.⁸

From Eq. (16) after multiplying by $\frac{1}{2}i\omega u_{l,m}$ and integrating over the total cross section we can write:

$$\begin{aligned}
& - \sum_j \frac{1}{2} \omega c_{44} k_{l,j} a_{l,j} \int_0^{t_l} u_{l,j} u_{l,m} dt \\
& + \sum_j \frac{1}{2} \omega c_{44} k_{l,j} b_{l,j} \int_0^{t_l} u_{l,j} u_{l,m} dt \\
& = \sum_j \frac{1}{2} \omega c_{44} k_{r,j} a_{r,j} \int_0^{t_r} u_{r,j} u_{l,m} dt \\
& - \sum_j \frac{1}{2} \omega c_{44} k_{r,j} b_{r,j} \int_0^{t_r} u_{r,j} u_{l,m} dt. \quad (19)
\end{aligned}$$

Again, introducing the quantity P , we obtain:

$$\sum_j Q_{r,j} L_{m,j} a_{r,j} - \sum_j Q_{r,j} L_{m,j} b_{r,j} = -P_{l,m} a_{l,m} + P_{l,m} b_{l,m}. \quad (20)$$

In order to derive the S-parameters¹⁴ for the defined problem, we write Eqs. (18) and (20) as:

$$|Q_{r,m}| \sum_j L_{j,m} b_{l,j} - b_{r,m} = -|Q_{r,m}| \sum_j L_{j,m} a_{l,j} - a_{r,m}, \quad (21)$$

$$P_{l,m} b_{l,m} + \sum_j Q_{r,j} L_{m,j} b_{r,j} = P_{l,m} a_{l,m} + \sum_j Q_{r,j} L_{m,j} a_{r,j}. \quad (22)$$

It is then possible to obtain a simple matrix in which we separate the inputs (a) from the unknown values (b):

$$[M_b] \begin{bmatrix} b_l \\ b_r \end{bmatrix} = [M_a] \begin{bmatrix} a_l \\ a_r \end{bmatrix} \quad (23)$$

with solution:

$$\begin{bmatrix} b_l \\ b_r \end{bmatrix} = [M_b^{-1}] [M_a] \begin{bmatrix} a_l \\ a_r \end{bmatrix}. \quad (24)$$

The scattering matrix will then be:

$$[S] = [M_b^{-1}] [M_a]. \quad (25)$$

Using Disperse¹⁵ it was possible to obtain the mode shapes and wave numbers for each of the computed modes in the frequency-thickness range 0–0.55 MHz mm. The mode shapes were extracted using 201 points through the thickness of the left side of the waveguide and we verified that the number of points was satisfactory by performing a convergence study. Subsequently the modes were power normalized according to Engan⁸ and the coefficients of the scattering matrix were calculated. The input vector was in this case a vector with all zeros except the element accounting for the SH0 excitation in the left side of the waveguide. In order to estimate the accuracy of the method, we performed convergence tests by increasing the number of modes considered in the computation and we also verified the conservation of energy. The energy check was done by estimating the error:

$$\epsilon = 1 - (|b_{l,SH0}|^2 + |b_{r,SH0}|^2), \quad (26)$$

where $b_{l,SH0}$ and $b_{r,SH0}$ are, respectively, the power normalized reflection and transmission coefficients for the SH0 mode (only propagating modes can carry energy).

V. RESULTS

The numerical results are presented below. We have investigated various cases describing the scattering phenomenon in plates with defects. The section is divided in two subsections in which we describe the results for the thickness step and the notch cases.

A. Plate with thickness step

Using the results of the FE simulations we studied the effect of up and down steps in a plate where the SH0 mode was incident. A relatively low frequency range (0–0.55 MHz mm) was investigated (e.g., 0–100 kHz in a 5.5 mm plate). Figures 7(a) and 8(a) show the modulus of the reflection

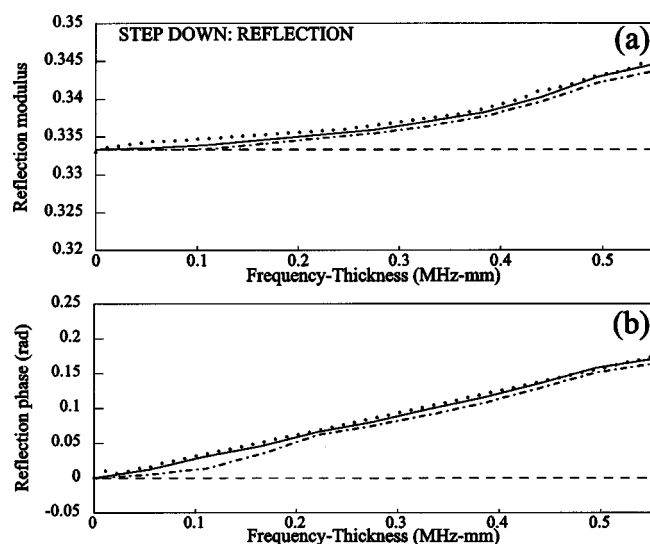


FIG. 7. Modulus (a) and phase angle (b) of reflected fundamental SH0 mode from a thickness step down of 50% of total thickness. The plot shows the results obtained with FE (dots) and modal decomposition using only SH0 (dashed line), SH0 and the first nonpropagating mode (double dashed line), SH0 and the first five nonpropagating modes (solid line).

tion and transmission coefficients respectively when there is a down step of 50%. In this frequency range the modulus is almost constant with frequency in both cases [note the very narrow range on the y axis of Fig. 7(a)]. The value of the reflection and transmission modulus can be approximated using Eqs. (5) and (6), respectively. A power normalization procedure was executed when the signal was transmitted through the step (because of the thickness change in the waveguide) in order to be able to compare the results obtained from the finite element analysis with the results obtained from the modal decomposition method (see Sec. IV). Figures 9(a) and 10(a) show the modulus of the reflection

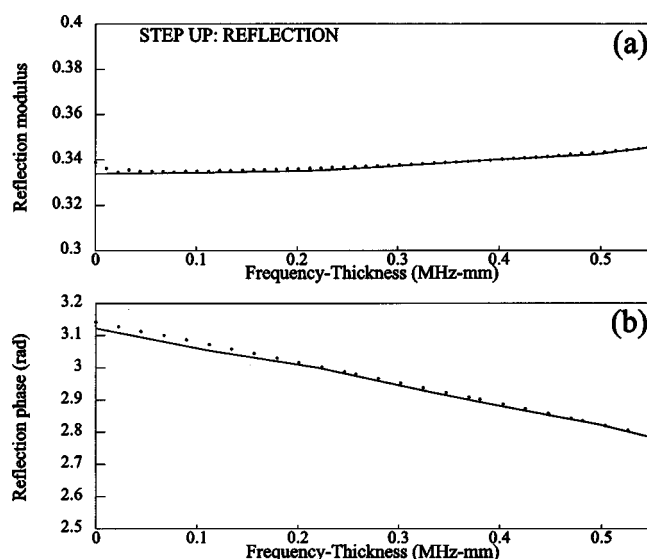


FIG. 9. Modulus (a) and phase angle (b) of reflected fundamental SH0 mode from a thickness step up of 50% of total thickness. The plot shows the results obtained with FE (dots) and modal decomposition using SH0 and five nonpropagating modes (solid line).

and transmission coefficients, respectively, when there is an up step of the thickness (the reflection and transmission modulus are power normalized).

Figures 7(b) and 8(b) show the phase shift of the reflected and transmitted signal, respectively, when there is a down step of 50% of the thickness. It is clear that the phase shift is frequency dependent in both cases and it varies monotonically with frequency. At zero frequency the value of phase shift tends to zero for both the transmitted and the reflected waves. In the case of the up step, the phase shift at zero frequency is zero for the transmitted signal [see Fig. 10(b)] and π for the reflected signal [see Fig. 9(b)].

The numerical results of the modal decomposition method are also shown. The modal decomposition method

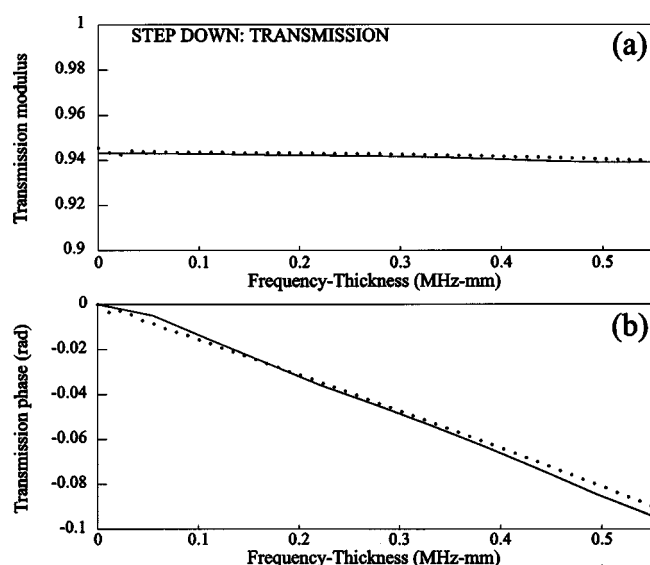


FIG. 8. Modulus (a) and phase angle (b) of transmitted fundamental SH0 mode from a thickness step down of 50% of total thickness. The plot shows the results obtained with FE (dots) and modal decomposition using SH0 and five nonpropagating modes (solid line).

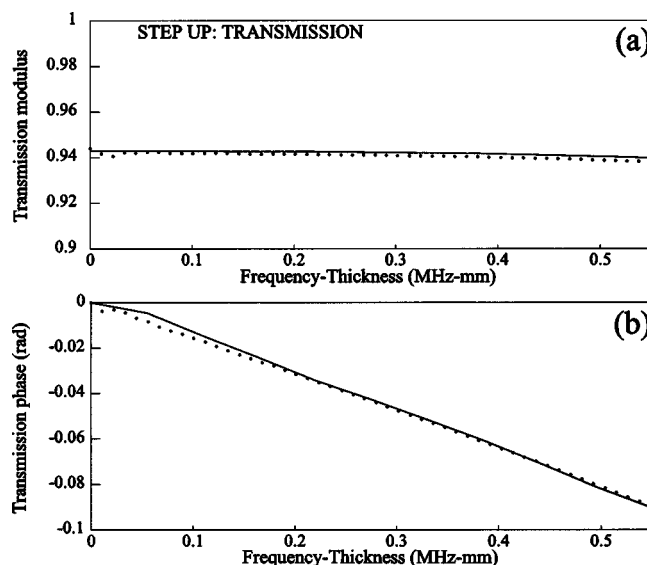


FIG. 10. Modulus (a) and phase angle (b) of transmitted fundamental SH0 mode from a thickness step up of 50% of total thickness. The plot shows the results obtained with FE (dots) and modal decomposition using SH0 and five nonpropagating modes (solid line).

using only SH0 leads to the same results as those obtained using the simplified theory (see Sec. IV). The reflection coefficient obtained using Eq. (5) is constant with frequency and, at low frequency-thickness, it is reasonably close to the reflection behavior predicted using finite elements, diverging at high frequency-thickness values [dashed line in Fig. 7(a)]. It is clear from Fig. 7(a) that the modulus of the reflection coefficient obtained from the finite element prediction can be approximated by simply using the SH0 and SH1 modes and neglecting the higher order nonpropagating modes in the computation. The reflection modulus is also shown for the case in which five nonpropagating modes were considered, confirming the good agreement between the finite element and the modal decomposition results.

The results for the phase information of reflection when there is a step down [see Fig. 7(b)] clearly shows that this varies with the number of modes included in the computation. As already mentioned in Sec. IV, the phase shift is zero if we take into account only the SH0 mode [see dashed line in Fig. 7(b)]. The phase shift obtained from the finite element predictions is approximated quite well by using two modes (SH0 and SH1) in the modal decomposition algorithm. A convergence test for the phase of the reflection and transmission was performed and its value converged when five nonpropagating modes were considered in the modal decomposition algorithm. The calculations showed that more modes tend to be required for the phase calculation to converge than for the amplitude to converge. The phase angles predicted introducing five nonpropagating modes in the modal decomposition algorithm matched very well with the results obtained from the FE models. The good agreement between the FE predictions and the modal decomposition method when introducing five nonpropagating modes in the computation was confirmed for the transmission through a down step (Fig. 8) and both reflection and transmission in the up step case (Figs. 9–10).

The energy criterion described in Sec. IV was satisfied even when only a few modes were included in the computation. The ϵ value was always less than 10^{-4} in the cases plotted in Figs. 7, 8, 9, and 10.

Tests were also done on 20% up and down steps but for brevity these are not shown here. Again good agreement between the results obtained using FE and the results from the modal decomposition were obtained. The number of modes to be included in the computation depended on the specific case being studied. More than 10 nonpropagating modes were needed to obtain the convergence of the algorithm when studying a 20% thickness step. It could be interesting to investigate ways of predicting the number of nonpropagating modes required in the analysis, but that has not yet been pursued.

B. Plate with rectangular notch

In this section, finite element predictions for the rectangular notch case are compared with superposition of the up and down step predictions.

The geometry of the notch is such that there will be a series of consecutive reflections from the notch decreasing in modulus with the number of times the wave travels back and

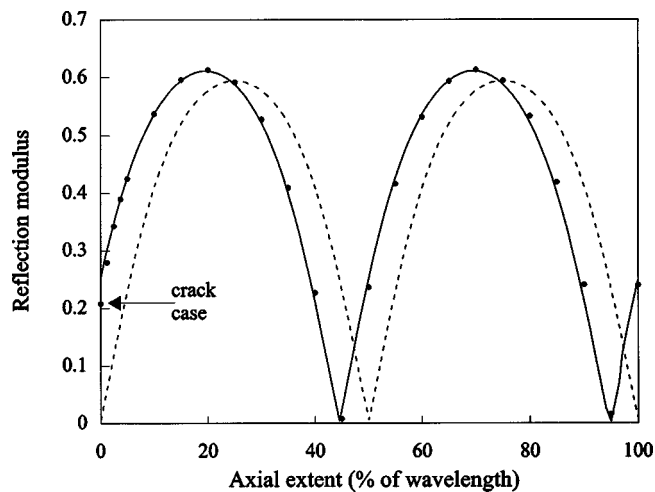


FIG. 11. Variation of reflection ratio with axial extent of the notch. Results are for a plate with SH0 at 0.55 MHz mm incident on a notch with 50% thickness depth. The dots indicate the FE results obtained for the rectangular notch case. The dashed line predicts the notch behavior using the simplified theory (no phase shift) and the solid line reproduces the notch reflection behavior using the complete theory with the phase shift information.

forth inside the notch. If the effect of the nonpropagating modes is neglected, it is possible to use the simplified theory explained in Sec. IV. The first reflection is given by Eq. (5). The second reflection is due to the signal transmitted through the down step (start of notch), reflected at the up step (end of notch), and transmitted through the up step (start of notch). The amplitude of the second reflection is:

$$R_{A2} = T_{A1} \cdot R_{B1} \cdot T_{B1}, \quad (27)$$

and its time delay is $2L/V$ where L is the axial extent of the notch and V is the SH0 velocity. The amplitude of the j th reflection is given by:

$$R_{Aj} = T_{A1} \cdot (R_{B1})^{2j-3} \cdot T_{B1}, \quad (28)$$

with time delay $j \cdot 2 \cdot L/V$. The total reflection is:

$$R_{TOT} = R_{A1} + R_{A2} + \dots + R_{Aj}. \quad (29)$$

If we do not consider the phase shift due to the contribution of the nonpropagating modes, the behavior of the total reflection coefficient from a 50% depth rectangular notch with varying axial extent is described by the dashed line in Fig. 11. In this prediction, the reflection coefficient has minima when the axial extent is $(n\lambda/2)$ and it has maxima when the axial extent is equal to $(n\lambda/2 + \lambda/4)$, where λ is the wavelength and n is an integer.

It is also possible to take into account the phase shift caused by the up and down steps. If we do this using the finite element or modal decomposition predictions from down and up steps in an algorithm which adds up the first four reflections with their modulus and phase, we obtain the solid curve in Fig. 11. This was traced at a frequency-thickness product of 0.55 MHz mm (100 kHz for a 5.5 mm plate). This curve is shifted to the left with respect to the curve obtained by neglecting the phase shifts at the steps.

In order to verify the accuracy of the predictions and identify its limits, we studied the reflection from rectangular notches with varying axial extent using finite element analy-

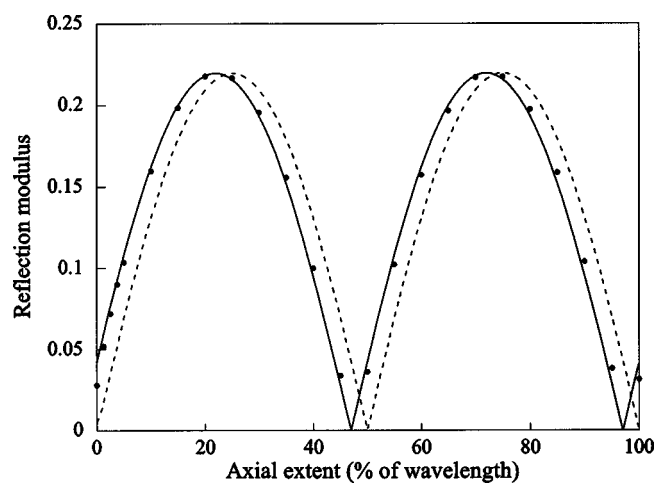


FIG. 12. Variation of reflection ratio with axial extent of the notch. Results are for a plate with SH0 at 0.55 MHz mm incident on a notch with 20% thickness depth. The dots indicate the FE results obtained for the rectangular notch case. The dashed line predicts the notch behavior using the simplified theory (no phase shift) and the solid line reproduces the notch reflection behavior using the complete theory with the phase shift information.

sis and the results are also shown in Fig. 11 for comparison. Clearly there is very good agreement between the predictions using the amplitude and phase information obtained from the step models and the notch cases studies using FE. In contrast, the reflection coefficient predicted neglecting the phase information does not match the results obtained from the FE model, demonstrating that the phase shift must be considered in this frequency regime. However, it is clear from Figs. 7(a), 8(a), 9(a), and 10(a) that the simplified theory (no phase shift), can be a relatively good approximation at very low frequency where the phase shift is small enough to be neglected.

Figure 12 shows similar results for 20% notch depth, very good agreement between the FE results for the notches and the predictions using the phase shift information again being obtained.

The effect of frequency on the reflection from notches is shown in Fig. 13. The reflection coefficient was plotted for a 20% depth notch with varying axial extent at three different frequencies (0.44 MHz mm, 0.55 MHz mm, and 0.66 MHz mm). It is clear from this plot that the effect of increasing the frequency (if working at relatively low frequency) is a shift of the curves toward the ordinate axis. This is due to the fact that the amplitude of the reflection and transmission coefficient is almost constant with frequency and the phase shift changes monotonically with frequency.

The result obtained for the crack case (zero axial extent) is highlighted in Fig. 11. The prediction for the crack is below the solid curve; this is the only point that does not match well with the curve. In order to verify the predictions for the crack case, we studied the effect of changing the axial extent of a V-notch. Figure 14 shows the modulus of reflection plotted versus V-notch opening $[L$ in Fig. 4(c)] on a logarithmic scale. The reflection coefficient decreases monotonically with the axial extent and, when the axial extent is less than 0.001% of the wavelength, it settles to the value of reflection coefficient obtained for the crack case in Fig. 11

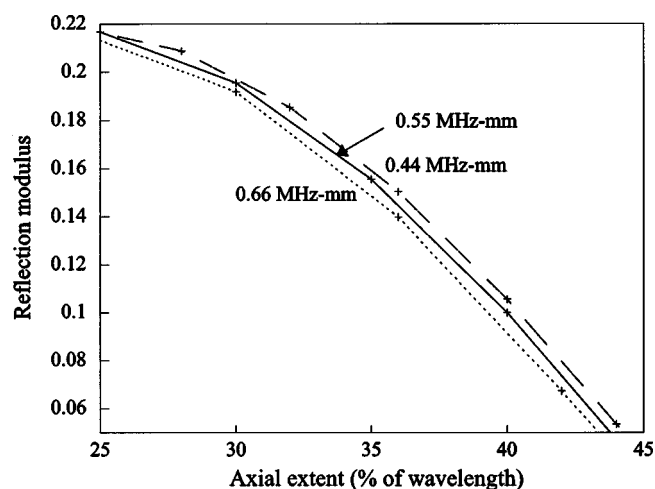


FIG. 13. Variation of reflection ratio with axial extent of the notch. Results are for plate with SH0 incident on a notch with 50% thickness depth. The dotted, solid, and dashed lines indicate the FE results obtained for a rectangular notch case at 0.66 MHz mm, 0.55 MHz mm, and 0.44 MHz mm, respectively. The crosses indicate the values at which the reflection coefficients were computed using the FE analysis.

(shown as dashed line in Fig. 14). Therefore the convergence of the notch case to the crack case at low axial extent of the notch was verified.

Figures 11 and 12 show that the crack case cannot be approximated using the predicted curve for the notch case. This is due to the presence of the nonpropagating modes. After the scattering of the incoming signal (SH0) on the left side of the notch (step down), both propagating (SH0 only in our case) and nonpropagating modes are present. If the axial extent of the notch is zero (crack case), the local vibration generated on the left side interacts with the right side and vice versa. A different interference pattern than the one depicted for a notch case (at relatively large axial extent) is to be expected in the crack case. In theory, for a rigorous analysis of the crack case using the approach proposed in this paper, the input signal at the second step discontinuity should be written considering both propagating and nonpropagating

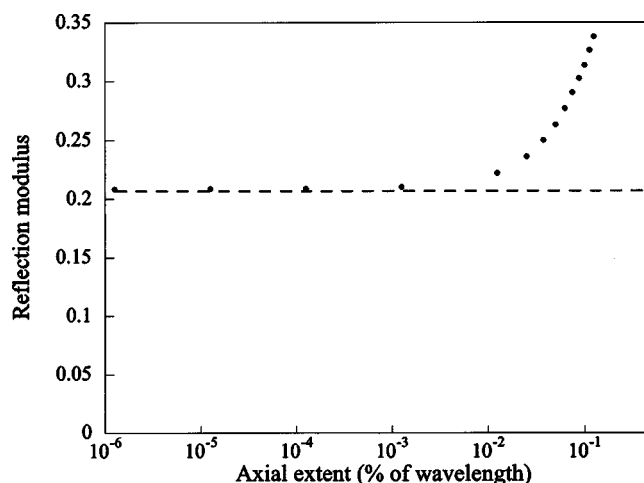


FIG. 14. Variation of reflection ratio with axial extent L of the V-notch. Results are for a 5.5 mm plate with SH0 incident at 100 kHz on a notch with 50% thickness depth. The dashed line is the crack prediction in Fig. 11 and the dots are the FE predictions.

modes. In practice, the input of the modal decomposition was SH0 in all cases so the effect of nonpropagating modes generated on one side of the crack and interacting with the opposite side was not taken into account. This indicates that it will be very difficult to model the reflection due to cracks in this way and it would be better to adopt the approach used by Le Clezio *et al.*²⁵ They developed their model specifically for the crack geometry and they assumed from the outset a large number of participating modes.

From Fig. 11 it is also evident that the notch cases characterized by relatively low axial extent are reasonably well approximated using our algorithm. It would be expected that, as the axial extent diminishes, the nonpropagating modes generated at the down step interact with the up step and vice versa. This is because the discontinuities are not far enough apart for the nonpropagating wave field to be considered to be completely attenuated. Moreover, as the axial extent reduces, the number of nonpropagating modes interacting with both discontinuities increases. This suggests that the reflection coefficient at very low axial extent of the notch cannot be approximated using our algorithm. However, in practice, the lowest axial extent case that was studied for the rectangular notch (1.25% of wavelength) was still reasonably well approximated using our method, suggesting that the nonpropagating modes only become significant at lower axial extents. In fact at 1.25% of the wavelength the first nonpropagating mode has already decayed by 25% of its maximum value (at zero distance) and the higher order modes are effectively completely attenuated due to their higher attenuation value (as shown in Fig. 2).

VI. CONCLUSIONS

The scattering of the fundamental SH wave from discontinuities in plate waveguides has been discussed. Both the finite element and the modal decomposition methods have been used to study the effect of thickness steps in plates, and the finite element method has also been used to study the reflection from cracks and notches in plates. An alternative method for predicting the reflection from notches has also been proposed. This is based on the superposition of the reflections from a down step (start of the notch) and an up step (end of the notch). The limits of the proposed method have been investigated. The convergence of the notch case to the crack case as the axial extent of the notch decreases has also been verified.

The modulus of reflection and transmission coefficients from a step in a plate when SH0 is incident at relatively low frequency is almost constant with frequency-thickness product and it can be approximated by simply considering SH0. The phase of the reflection and transmission coefficient varies monotonically with the frequency-thickness product and it is necessary to consider the effect of non-propagating modes in the modal solution to reproduce its frequency-dependent behavior. Very good agreement between the finite element and modal solutions has been shown.

The reflection from a rectangular notch can be approximated by superimposing the reflection from a down step followed by an up step, the phase information of the reflection

and transmission at the steps being very important to reproduce the interference phenomenon at the notch.

The superposition of the reflection for down and up steps cannot be used with the same success to approximate the reflection from cracks, because the interaction of nonpropagating modes with both sides of the notch has not been taken into account (the nonpropagating modes generated at one side of the notch are assumed to decay before reaching the opposite side of the notch). However, the results obtained at notch axial extent as low as 1.25% of the wavelength still agreed very well with the finite element solution, indicating that the interaction of nonpropagating modes with the sides of the notch is only significant at very small axial extents.

- ¹P. Wilcox, M. Lowe, and P. Cawley, "Lamb and SH wave transducer arrays for the inspection of large areas of thick plates," in *Review of Progress in Quantitative NDE*, edited by D. Thompson and D. Chimenti (Plenum, New York, 2000), Vol. 19, pp. 1049–1056.
- ²D. Alleyne and P. Cawley, "Optimization of Lamb wave inspection techniques," *NDT & E Int.* **25**, 11–22 (1992).
- ³J. Ditri, J. Rose, and G. Chen, "Mode selection criteria for defect optimization using Lamb waves," in *Review of Progress in Quantitative NDE*, edited by D. Thompson and D. Chimenti (Plenum, New York, 1992), p. 2109.
- ⁴T. Ghosh, T. Kundu, and P. Karpur, "Efficient use of Lamb modes for detecting defects in large plates," *Ultrasonics* **36**, 791–801 (1998).
- ⁵D. Alleyne, B. Pavlakovic, M. Lowe, and P. Cawley, "Rapid, long range inspection of chemical plant pipework using guided waves," *Insight* **43**, 93–96, 101 (2001).
- ⁶M. Koshiha, H. Hasegawa, and M. Suzuki, "Finite-element solution of horizontally polarized shear wave scattering in an elastic plate," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **34**, 461–466 (1987).
- ⁷J. Ditri, "Some results on the scattering of guided elastic SH waves from material and geometric waveguide discontinuities," *J. Acoust. Soc. Am.* **100**, 3078–3087 (1996).
- ⁸H. Engan, "Torsional wave scattering from a diameter step in a rod," *J. Acoust. Soc. Am.* **104**, 2015–2024 (1998).
- ⁹M. Lowe and O. Diligent, "Low frequency reflection characteristics of the S0 Lamb wave from a rectangular notch in a plate," *J. Acoust. Soc. Am.* **111**, 64–74 (2002).
- ¹⁰M. Lowe, P. Cawley, J.-Y. Kao, and O. Diligent, "Low frequency reflection characteristics of the A0 Lamb wave from a rectangular notch in a plate," *J. Acoust. Soc. Am.* **112**, 2612–2622 (2002).
- ¹¹J. Rose and X. Zhao, "Anomaly throughwall depth measurement potential with shear horizontal guided waves," *Mater. Eval.* **59**, 1234–1238 (2001).
- ¹²D. Alleyne, M. Lowe, and P. Cawley, "The reflection of guided waves from circumferential notches in pipes," *J. Appl. Mech.* **65**, 635–641 (1998).
- ¹³A. Demma, P. Cawley, M. Lowe, and A. Roosenbrand, "The reflection of the fundamental torsional mode from cracks and notches in pipes," *J. Acoust. Soc. Am.* (submitted).
- ¹⁴B. Auld, *Acoustic Fields and Waves in Solids*, Vol. 2 (Krieger, Malabar, FL, 1990).
- ¹⁵B. Pavlakovic, M. Lowe, D. Alleyne, and P. Cawley, "DISPERSE: A general purpose program for creating dispersion curves," in *Review of Progress in Quantitative NDE*, edited by D. Thompson and D. Chimenti (Plenum, New York, 1997), Vol. 16, pp. 185–192.
- ¹⁶M. Lowe, D. Alleyne, and P. Cawley, "The mode conversion of a guided wave by a part-circumferential notch in a pipe," *J. Appl. Mech.* **65**, 649–656 (1998).
- ¹⁷M. Lowe, "Characteristics of the reflection of Lamb waves from defects in plates and pipes," in *Review of Progress in Quantitative NDE*, edited by D. Thompson and D. Chimenti (Plenum, New York, 1998), p. 113.
- ¹⁸D. Alleyne and P. Cawley, "The interaction of Lamb waves with defects," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**, 381–397 (1992).
- ¹⁹D. Alleyne, "The Nondestructive Testing of Plates Using Ultrasonic Lamb Waves," Ph.D. thesis, University of London (1991).
- ²⁰F. Moser, L. Jacobs, and J. Qu, "Modeling elastic wave propagation in waveguides with finite element method," *NDT & E Int.* **33**, 225–234 (1999).
- ²¹N. Kishore, I. Sridhar, and N. Iyengar, "Finite element modelling of the

- scattering of ultrasonic waves by isolated flaws," *NDT & E Int.* **33**, 297–305 (2000).
- ²²P. Wilcox, M. Evans, O. Diligent, M. Lowe, and P. Cawley, "Dispersion and excitability of guided acoustic waves in isotropic beams with arbitrary cross section," in *Review of Progress in Quantitative NDE*, edited by D. Thompson and D. Chimenti (American Institute of Physics, New York, 2002), Vol. 21, pp. 203–210.
- ²³B. Pavlakovic, "Leaky guided ultrasonic waves in NDT," Ph.D. thesis, University of London (1998).
- ²⁴B. Auld, *Acoustic Fields and Waves in Solids*, Vol. 1 (Krieger, Malabar, FL, 1990).
- ²⁵E. Le Clezio, M. Castaings, and B. Hosten, "The interaction of the S0 Lamb mode with vertical cracks in an aluminium plate," *Ultrasonics* **40**, 187–192 (2002).
- ²⁶T. Vogt, M. Lowe, and P. Cawley, "The scattering of guided waves in partly embedded cylindrical structures," *J. Acoust. Soc. Am.* **113**, 1258–1272 (2003).

Acoustic streaming measurements in annular thermoacoustic engines

Stéphane Job^{a)}

Laboratoire d'Acoustique (UMR CNRS 6613), Université du Maine, 72000 Le Mans, France

Vitaliy Gusev

Laboratoire de Physique de l'Etat Condensé (UMR CNRS 6087), Laboratoire d'Acoustique (UMR CNRS 6613), Université du Maine, 72000 Le Mans, France

Pierrick Lotton and Michel Bruneau

Laboratoire d'Acoustique (UMR CNRS 6613), Université du Maine, 72000 Le Mans, France

(Received 28 December 2001; revised 23 December 2002; accepted 6 January 2003)

Experiments with an annular thermoacoustic engine employing quasiadiabatic interaction between traveling acoustic waves and an inhomogeneously heated porous material indicate the presence of a closed-loop mass flux. A qualitative modeling of the enthalpy flux in the thermoacoustic core provides an opportunity to estimate the thermal convection associated with this mass flux, by using temperature measurement at different positions in the system. The estimated acoustically induced mass flux is in accordance with recent theoretical results. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1555076]

PACS numbers: 43.35.Ud, 43.25.Zx, 43.20.Mv [MFH]

I. INTRODUCTION

Thermoacoustic (TA) interaction, the interaction between acoustic waves and temperature oscillations, is a phenomenon developed in acoustic boundary layers near rigid walls.¹ The TA interaction occurring in acoustic resonators subjected to inhomogeneous heating may become unstable and leads to the generation of acoustic oscillations.¹ This effect is enhanced by using the interaction between an inhomogeneously heated stack of solid plates placed in the resonator and the gas oscillations for the transformation of thermal energy into mechanical energy. The idea of a traveling wave annular TA engine was put forward more than 20 years ago.² However, the first quasiadiabatic annular TA engine was built only a few years ago.³ Recently, a closed-loop TA engine which employed quasi-isothermal sound propagation through the stack (Stirling cycle through the regenerator) was reported.⁴ This experiment⁴ demonstrated that circulation of the gas is induced by the traveling wave in a TA Stirling engine. More recently, the presence of closed-loop induced acoustic streaming has been reported in an annular Stirling cooler driven by a TA regenerator (also called the prime-mover stack in comparison with the cooler stack).⁵ This closed-loop mass flux contributes to the enthalpy transport in annular devices, and should be suppressed in order to increase the efficiency of a TA engine.^{4,5} The experiments with externally induced mass flux (of nonacoustical nature) through TA stack have been also reported.^{6,7} Acoustically induced mass flux has been theoretically predicted in annular Stirling engine.⁸ However, neither measurements nor estimates of the acoustically induced streaming mass flux have been reported in an annular acoustic resonator driven by a quasiadiabatic TA prime mover.

In the present paper, results of temperature measurements at some characteristic positions in the quasiadiabatic stack driving an annular acoustic resonator are presented, both above and below the threshold for TA instability. It is demonstrated that, similarly to the case of a TA engine with a quasi-isothermal stack,⁴ a closed-loop circulation of the gas is excited in a quasiadiabatic stack. A simple theoretical model provides an estimate of the acoustically induced mass flux. This value matches results from recently developed theory.⁹ Theoretical estimation⁹ of closed-loop acoustic streaming in annular TA devices can be very helpful, for example, to choose an appropriate working fluid that minimizes acoustically induced mass flux.

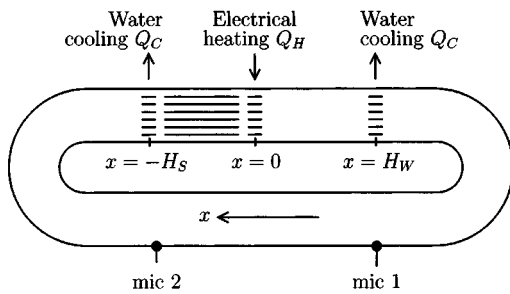
The annular TA engine under study is described in Sec. II and observations of some of acoustical and thermal phenomena are presented. An experimental method to determine the acoustically induced mass flux from temperature measurements is presented in Sec. III. Measured values of this mass flux are compared to the available theoretical estimation⁹ and its improvement in Sec. IV before concluding in Sec. V. Improvements of theoretical results presented in Ref. 9 are exposed in the Appendix.

II. EXPERIMENTAL SETUP: DESCRIPTION AND OBSERVATIONS

The annular TA engine used in our experiments is presented in Fig. 1(a). This resonator, which consists of a torus-shaped waveguide made of stainless-steel (AISI 316L) cylindrical pipes, is composed of two straight pipes (length 0.6 m) and two curved pipes (curvature radius 0.14 m). All pipes have circular cross section (inner radius $r_i \approx 26.5$ mm, outer radius $r_o \approx 30.1$ mm, and wall thickness 3.6 mm) and the total axial length of the waveguide is $L \approx 2.24$ m. This waveguide is filled with air at ambient temperature $T_0 \approx 293$ K

^{a)}Electronic mail: stephane.job@univ-lemans.fr

(a) Annular thermoacoustic prime-mover



(b) Thermoacoustic core

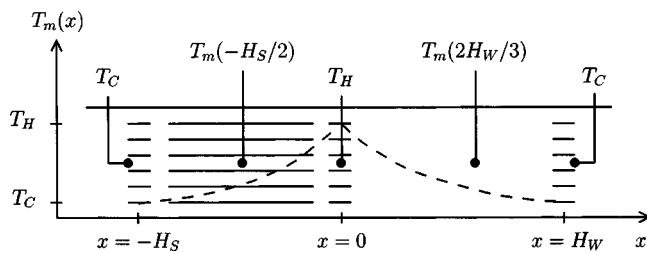


FIG. 1. (a) Schematic view of the annular TA engine ($L \approx 2.24$ m) and the two piezoresistive microphones (mic. 1 and mic. 2). (b) Enlarged view of the TA core ($-H_S < x < H_W$, $H_S \approx 0.15$ m and $H_W \approx 0.38$ m) with the ceramic stack ($-H_S < x < 0$) and five of the six temperature probes. Temperature distribution is schematically represented in dashed line.

and atmospheric pressure $P_0 \approx 10^5$ Pa. In the following, ambient condition is related to temperature T_0 and pressure P_0 .

Under specific conditions^{1,2,10} presented below in more detail, the TA interaction leads to the development of an instability, called the TA instability: an acoustic field can be generated and sustained in the waveguide. Due to the annular geometry of the waveguide, the axial wave number of acoustic oscillations is attempted to be close to axial resonance conditions $k_n \approx 2n\pi/L$, where n is a strictly positive integer. Moreover, if the wave number is lower than the wave number of the first transversal mode ($k_0 = 1.84/r_i$), only plane wave propagates in the cylindrical pipe. Considering dimensions of the experimental setup, one can estimate that the first 24 ($k_n < k_0$ leads to $n \leq 24$) axial resonance modes propagate in plane wave. Experimentally, the measured frequency of acoustic oscillations is equal to $f_1 \approx 152$ Hz: it corresponds to the first resonance mode $k_1 = 2\pi f_1/a_C \approx 2\pi/L$ of the annular pipe, where $a_C \approx 340$ m/s is the air sound velocity at cold temperature, and acoustic oscillations propagate in plane wave at this frequency.

Above the onset of the TA instability, the presence of acoustic oscillations also induces a mass flux which results from various nonlinear processes.^{1,4,6-11} This acoustically induced mass flux produces a closed-loop circulation of the gas around the annular waveguide that carries the enthalpy by forced thermal convection.

The TA interaction occurs near rigid walls inside acoustic boundary layers.¹⁰ A ceramic porous material (length $H_S \approx 0.15$ m), made of *Cordierite* and usually used in catalytic exhaust pipe, is placed in the waveguide. This ceramic, called the stack, contains a multitude of square cross-section channels (characteristic transversal width $D_S \approx 0.9$ mm, approximate wall thickness 0.1 mm) parallel to the waveguide

axis. The stack is used to increase locally the volume of boundary layer in the fluid. Considering that the thermal boundary layer ($\delta_\kappa = \sqrt{2\kappa/\omega_1} \approx 0.21$ mm) is more than four times smaller than the characteristic transversal width D_S , TA interaction employs a quasiadiabatic regime¹⁰ ($\kappa \approx 2.1 \cdot 10^{-5}$ m²/s is the thermal diffusivity in air at ambient condition, and $\omega_1 = 2\pi f_1$ is the angular frequency).

A temperature difference is imposed between stack ends: its left side ($x = -H_S$) is maintained at cold temperature $T_C = T_0$ and its right side ($x = 0$) is maintained at hot temperature $T_H > T_C$. The origin of axial coordinate x is chosen to be the position of hot temperature T_H , and the axial coordinate is directed clockwise in Fig. 1(a). Spatial extent of the inhomogeneously heated part of the waveguide ($0 < x < H_W$, where $H_W \approx 0.38$ m) is controlled by imposing the cold temperature T_C at the position $x = H_W$. The inhomogeneously heated region $-H_S < x < H_W$, presented in Fig. 1(b), is called the TA core. The heated region of the stack is denoted by the subscript S , and the heated region of the waveguide by the subscript W .

Temperatures at cold positions ($x = -H_S$ and $x = H_W$) are imposed by using two cold heat reservoirs (at cold temperature T_C), which consist of two water-cooled copper shells set around the waveguide. Two cold heat exchangers made of copper wire grids (wire approximately 0.2 mm in diameter and approximately 1 mm spaced) are placed in contact with cold reservoirs perpendicularly to the waveguide axis. One of the heat exchangers ($x = -H_S$) is in contact with the stack. Copper grids, ensuring a uniform radial temperature distribution in the cross section of the pipe, are designed to achieve approximately the same characteristic transversal width as in the stack.

A hot heat reservoir (temperature T_H at position $x = 0$), consisting of eight electrical cartridges (125 W each) soldered in a stainless-steel block, is placed in contact with external surface of the waveguide. The electrical power Q_H supplied to the hot heat reservoir is extracted by measuring electrical voltage and current supplied to the cartridges. A hot heat exchanger is placed in contact with the stack, in front of the hot heat reservoir, and firmly in contact with internal surface of the waveguide. The hot heat exchanger is an appropriately coiled stainless-steel ribbon (width 2.5 cm and thickness 10 μ m). The distance between two adjacent walls in the hot heat exchanger, chosen to be close to the characteristic transversal width in the stack, is approximately 1 mm.

When an axial distribution of the mean temperature $T_C < T_m(x) < T_H$ of the fluid is imposed in the TA core ($-H_S < x < H_W$), the TA interaction can be seen from a simplified point of view as the conversion of a part of the supplied heat in mechanical energy of acoustic oscillations.¹⁰ The TA interaction mainly occurs in the stack. This interaction leads to the amplification of an acoustic wave traveling through the stack in the direction of positive temperature gradient.^{1,2,10} When the ratio T_H/T_C of hot to cold temperature reaches the threshold value $(T_H/T_C)_{th} \approx 1.9$, the TA amplification in the stack compensates acoustic dissipation in the waveguide and the system becomes unstable. Any initial acoustical perturbation (noise for example) is first amplified at resonance fre-

quency and then sustained in the waveguide. The threshold condition is achieved by applying an electrical load of $Q_{th} \approx 130$ W.

Considering that acoustic oscillations propagate in plane waves, the acoustic field can be expressed in the form $p(x, t) = \text{Re}[\tilde{p}(x)e^{-i\omega_1 t}]$, where

$$\tilde{p}(x) = [\tilde{p}^+(x_0)e^{+ik_W(x-x_0)} + \tilde{p}^-(x_0)e^{-ik_W(x-x_0)}] \quad (1)$$

is the acoustic pressure in frequency domain. The acoustic wave number k_W , which takes into account dissipative and dispersive effects in the cold part of the waveguide, is defined as

$$k_W \approx k_C[1 + (1+i)c(\delta_\nu/r)]. \quad (2)$$

In Eq. (2), $k_C = \omega_1/a_C$ is the acoustic wave number at cold temperature in absence of sound interaction with walls, $\delta_\nu = \sqrt{2\nu/\omega_1} \approx 0.17$ mm is the viscous boundary layer, $\nu \approx 1.5 \cdot 10^{-5}$ m²/s is the kinematic viscosity in air at ambient condition, $c = (1/2)[1 + (\gamma - 1)/\sqrt{\sigma}] \approx 0.74$ is the Kirchhoff's constant,^{1,12} $\gamma \approx 1.4$ is the polytropic coefficient, and $\sigma = \nu/\kappa \approx 0.71$ is the Prandtl number of air.

Two piezoresistive PCB microphones placed at $x_{m1} \approx 0.95$ m and $x_{m2} \approx 1.53$ m allow the measurement of the acoustic field. Each microphone has been calibrated in phase and in amplitude by comparison with a reference B&K microphone. By measuring acoustic pressure at two separated positions $x = x_{m1}$ and $x = x_{m2}$, one can extract frequency of acoustic oscillations and both contrapropagating wave complex amplitudes $\tilde{p}^+(x_0)$ and $\tilde{p}^-(x_0)$ anywhere in the cold part of the waveguide ($x_0 < -H_S$ or $x_0 > H_W$). Clockwise and anticlockwise traveling waves amplitudes $\tilde{p}^+(-H_S) \equiv p^+$ are then extracted as a function of heat supplied Q_H . As shown in Fig. 2(a), the acoustic pressure amplitude increases when heat load Q_H is increased above the TA instability threshold condition. It is estimated that the amplitude of the clockwise-traveling wave $|\tilde{p}^+|$ exceeds the amplitude of the wave traveling in the opposite direction $|\tilde{p}^-|$ by more than a factor of 2, as shown in Fig. 2(b). However, the amplitude of the anticlockwise traveling wave is not completely negligible. This can be explained by the fact that though only the clockwise-traveling wave is amplified in the heated stack,^{1,2,10} a significant part of this wave is reflected in the system. Relative phase $\phi(p^-/p^+) = \phi(p^-) - \phi(p^+)$ measurements, also presented in Fig. 2(b), confirm that the interaction between contrapropagating waves in the annular TA resonator is not trivial: the acoustic field generated in the system is neither a standing wave nor a purely traveling wave acoustic field.^{3,5}

Temperature measurements are obtained by using six thermocouples (1 mm in diameter type-K probes). Five of them are set in the TA core ($x = -H_S$, $-H_S/2$, 0 , $2H_W/3$, H_W) on the axis of the waveguide as shown in Fig. 1(b). The two cold thermocouples ($x = -H_S, H_W$) are placed in contact with the two cold heat exchangers outside the TA core. The hot temperature thermocouple ($x = 0$) measures the temperature at the center of the hot heat exchanger. The sixth temperature probe (not shown in Fig. 1) controls that the temperature far outside the TA core (in the cold part of the waveguide) remains constant (i.e., at cold temperature T_C).

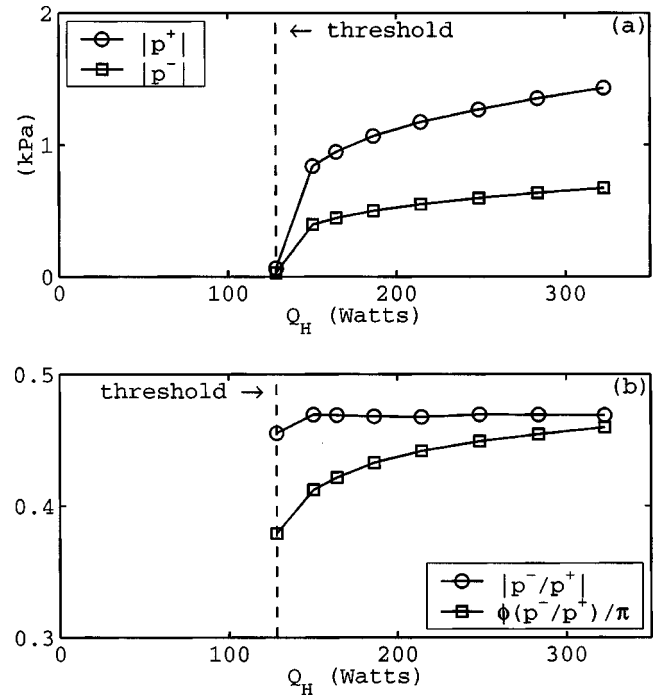


FIG. 2. (a) Measured amplitudes of clockwise (\tilde{p}^+) and anticlockwise (\tilde{p}^-) waves, plotted versus heat load Q_H . (b) Modulus and phase of the ratio $R(-H_S) = \tilde{p}^-/\tilde{p}^+$ plotted versus heat load Q_H . Instability onset occurs at $Q_H \approx 130$ W.

The six thermocouples used are provided from the same manufacturing series and are certified by the manufacturer. It has been checked that all thermocouples give the same measurement (within an error of less than 1 K) of the ambient temperature T_0 . As a consequence, absolute temperature measurement may be considered within an error of the order of 1 K. Despite this accuracy, the relative error in absolute temperature measurements is overestimated to $\pm 5\%$. In the whole range of experiments, this relative error corresponds at least to ± 15 K (at ambient temperature T_0) and at the most to ± 30 K (at the highest measured temperature), which is much greater than the usual accuracy of type-K probes (less than 1 K). This overestimation of the error in temperature measurement is introduced to take into account the lack of other sources of errors. For example, it is assumed in the following that the temperature does not depend on radial coordinate, although it may vary in the cross section of the waveguide. In fact, this variation is less than or equal to 3 K,¹³ i.e., less than the considered relative error of $\pm 5\%$ in temperature measurement.

Results of temperature measurements are presented in Fig. 3, below and above the onset of TA instability. As shown in Fig. 3(a), where absolute temperature measurements are plotted as functions of the heat load Q_H , it is first checked that cold temperature T_C (measured outside the TA core) and cold exchangers temperatures $T_m(-H_S)$ and $T_m(H_W)$ (not shown in Fig. 3) remain constant and equal to ambient temperature T_0 in the whole range of experiments. It is then noted that below the threshold condition, hot temperature T_H , temperature measured at $x = -H_S/2$ (in the stack), and temperature measured at $x = 2H_W/3$ (in the heated part of the waveguide) increase when Q_H increases. At the threshold

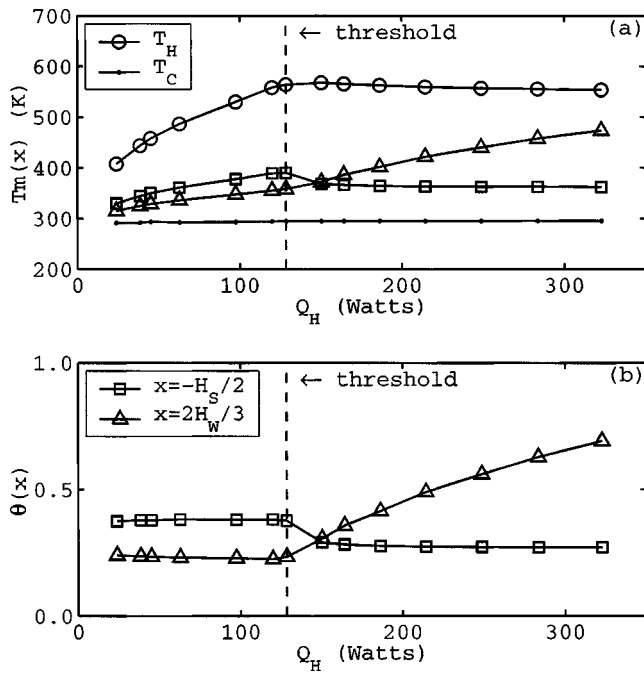


FIG. 3. (a) Absolute measured temperatures above and below the threshold condition for TA instability onset plotted vs heat load Q_H . (b) Normalized temperatures $\theta(x) = [T_m(x) - T_C]/[T_H - T_C]$ calculated from inset (a). In inset (a), circles are hot temperature measurements ($x=0$) and dots are cold temperature measurements (far outside the TA core). In inset (a) and (b), squares are temperatures measured inside the stack ($x = -H_S/2$) and triangles are temperatures measured inside the heated part of the waveguide ($x = 2H_W/3$). Error bounds in temperature measurements are not shown.

condition, acoustical phenomena (i.e., acoustic waves and acoustically induced mass flux) are generated: the thermal behavior of the system is modified. Above the threshold, temperature T_H remains approximately constant whereas Q_H increases, and an evolution of the temperature profile in the TA core is observed. This demonstrates that there is a significant influence of streaming on temperature distribution: the thermal forced convection led by the acoustically induced mass flux transports cold air from the cold part of the waveguide through the TA core.

Temperature profile evolution in the TA core is described in Fig. 3(b) by using a normalized temperature defined in the form

$$\theta(x) = [T_m(x) - T_C]/[T_H - T_C]. \quad (3)$$

The latter quantity $\theta(x)$ is also used in the thermal model presented in Sec. III. The normalized temperature $\theta(x)$ is evaluated at characteristic positions $x = -H_S/2$ and $x = 2H_W/3$ and plotted in Fig. 3(b) as a function of heat load Q_H . The decreasing of the normalized temperature $\theta(x)$ at $x = -H_S/2$ and its increasing at $x = 2H_W/3$ indicate that a mass flux carrying the enthalpy parallel to the waveguide axis is induced above the threshold of the TA instability. Moreover, the direction of the bend indicates that the mass flux is directed toward the clockwise direction. The fact that the direction of the mass flux is the same as the direction of the amplified traveling wave is in accordance with available theoretical results for annular TA engines driven by quasidiabatic prime mover.⁹

III. MASS FLUX MEASUREMENT

In order to measure the mass flux, the following qualitative model of the axial temperature distribution in the TA core is proposed. The aim of this model is to provide an estimation of the order of magnitude of the acoustically induced mass flux. It is not intended to ensure accurate quantitative measurements. In the frame of this approach, simplifying assumptions have been used. The temperature distribution inside the cylindrical tube is assumed to be independent of the radial coordinate. This assumption has been confirmed experimentally¹³ by using additional thermocouples (thinner and more sensitive than actual temperature probes) in the radial direction of the stack. The temperature variation in the cross section is less than or equal to 3 K.¹³ Neglecting radial contributions, the enthalpy flux can be written in the form

$$J = -A \partial_x T_m + C_p M T_m, \quad (4)$$

where $\partial_x \equiv \partial/\partial x$ denotes the partial x derivative, $C_p \approx 10^3$ J/(kg K) the specific heat of air, M the mass flux (expressed in kg/s) averaged over the cross section of the waveguide, and A a constant. The second term in the right-hand side (r.h.s.) of Eq. (4) takes into account the forced thermal convection due to the mass flux M . The first term in the r.h.s. of Eq. (4) describes the thermal conduction, so the constant A can be considered as an area scaled effective thermal conductivity

$$A \equiv k_{st} \pi (r_o^2 - r_i^2) + k_{air} \phi \pi r_i^2 + k_{ce} (1 - \phi) \pi r_i^2, \quad (5)$$

where $k_{st} \approx 16$ W/(m K) at $T \approx 373$ K and $k_{st} \approx 25$ W/(m K) at $T \approx 973$ K (intermediate values are obtained using a linear fit), $k_{air} \approx 0.025$ W/(m K), and $k_{ce} \approx 1$ W/(m K) are, respectively, the thermal conductivities of steel, air, and ceramic. In the frame of our qualitative approach, C_p , k_{air} , and k_{ce} are assumed to be temperature-independent constants. The porosity (the volume of fluid divided by the total volume of the stack region) is taken equal to $\phi \approx 0.8$ in the region $-H_S < x < 0$ and to $\phi = 1$ in the region $0 < x < H_W$.

The heat transfer from the external surface of the waveguide to the surroundings is modeled by using the Newton's law of cooling:¹⁴

$$\partial_x J + h(T_m - T_C) = 0, \quad (6)$$

where h is a phenomenological coefficient of heat transfer (mainly due to natural convection and radiation). The first term of the left-hand side (l.h.s.) of Eq. (6) describes the axial variation of the enthalpy flux caused by heat leakage to surroundings in radial direction (second term of l.h.s.). Finally, substituting Eq. (4) into Eq. (6), an equation for the temperature distribution is obtained

$$A \partial_{xx}^2 \theta - C_p M \partial_x \theta - h \theta = 0, \quad (7)$$

where $\theta(x)$ is the normalized mean temperature of the fluid defined in Eq. (3).

In order to find a tractable analytical solution, coefficients in Eq. (7) (which in fact all depend on temperature) are approximated by their mean values, estimated at a characteristic temperature $T_{char} = (T_H + T_C)/2$. In the frame of our qualitative thermal model, coefficients in Eq. (7) are thus

assumed to be independent of the local temperature and consequently independent of the spatial coordinate. For an imposed temperature T_H , it is then straightforward to find the solution (depending on position x and on parameter h and M) of the linear differential equation Eq. (7) with constant coefficients. This solution satisfies the boundary conditions $\theta(x=0)=1$, $\theta(x=-H_S, H_W)=0$, both in the stack region ($-H_S < x < 0$) and in the waveguide region ($0 < x < H_W$). Consequently, if coefficients A and h are known, mass flux M can be extracted by mapping the solution to Eq. (7) to temperature measurements.

It is important to keep in mind that in the frame of our one-dimensional thermal theory, the quantity M involved in Eq. (7) is the acoustically induced mass flux averaged over the cross section of the waveguide. However, the mass flux actually depends on the radial coordinate in the form of a parabolic laminar flow:^{9,15} $M(r) \approx M(r=0)[1 - (r/r_i)^2]$. Assuming the expected velocity of the streaming⁹ is of the order of magnitude of $v \sim 10^{-2}$ m/s and using the thermal diffusivity of air ($\kappa \approx 2.1 \cdot 10^{-5}$ m²/s), the characteristic distance of thermocouple sensitivity is of the order of $\kappa/v \sim 2$ mm $\ll r_i$. Temperature measurements at $r=0$ thus provide the local measurement $M(r=0)$ of the mass flux, and the correct value of the mass flux averaged over the cross section of the waveguide is equal to $M(r=0)/2$.

It is interesting to analyze the order of magnitude of involved thermal effects. First, the thermal conduction in the steel wall of the pipe [first term in the r.h.s. of Eq. (5)] is approximately 100 times greater than thermal conduction in air or in ceramic stack [second and third terms in the r.h.s. of Eq. (5)]. Second, in the whole range of our experiments, it has been checked that thermal conduction [first term in the l.h.s. of Eq. (7)] and heat leaks [third term in the l.h.s. of Eq. (7)] are of the same order of magnitude. Third, above the threshold condition (when a mass flux is generated), it has been checked that the forced convection [second term in the l.h.s. of Eq. (7)] is at least twice greater (for the minimum value of the mass flux measured near the threshold condition) and at the most ten times greater (for the maximum measured mass flux) than both other contributions [first or third terms in the l.h.s. of Eq. (7)]. Thereby, forced thermal convection due to the mass flux M is the dominant effect described in Eq. (7).

Although the mass flux depends on the radial coordinate (in the form of a parabolic laminar flow), the enthalpy flux carried by the forced thermal convection is only directed toward the axial coordinate. Similarly, because of the high thermal conductivity of steel, and because temperature gradients are only imposed in the axial direction, radial temperature gradients in the wall of the pipe are weak. Enthalpy flux carried by thermal conduction in steel is negligible in the radial direction compared to its contribution in the axial direction. Moreover, thermal conduction in air and in ceramic is negligible both in the axial and in the radial direction compared to axial thermal conduction in steel and thermal forced convection in the gas. Finally, except for heat leaks from external surface of the waveguide to the surroundings [Eq. (6)], thermal contributions in radial direction

do not have significant influence on temperature distribution and can be neglected in Eq. (4).

The following procedure is applied to evaluate the mass flux M . First, phenomenological parameters h_S and h_W are measured. For this purpose, an absorbing partition is installed in the cold part of the waveguide in order to prevent the TA instability. Thus, $M \equiv 0$ and by mapping the solutions of Eq. (7), respectively, to the measured values of the temperature at $x = -H_S/2$ and $x = 2H_W/2$, values of coefficients h_S and h_W are found as a function of T_{char} . It appears that both these coefficients weakly depend on the characteristic temperature. Orders of magnitude of these coefficients are $h_S \sim 1.4$ W/(m K) and $h_W \sim 0.3$ W/(m K) at ambient temperature, and they vary no more than 10% in the whole range of our experiments. Second, the absorbing partition is removed to allow the TA instability development. By mapping the theoretical solution of Eq. (7) to the experimental temperature measured at $x = 2H_W/3$ above the threshold, and by using previous measurements of h_W (as a function of T_{char} , i.e., as a function of Q_H) the dependence of the mass flux M on the supplied heat Q_H can be extracted.

The mass flux M and the coefficients h_S and h_W being known, theoretical temperature distributions satisfying Eq. (7) can be evaluated. Two examples of theoretical fit of the axial temperature distribution are given in Fig. 4, where circles represent measured values. Solid line and dashed line in Fig. 4 represent mapped solutions of Eq. (7) to temperatures, respectively, measured above and below the threshold condition for TA instability. The bending of the temperature curve above threshold condition indicates the presence of the mass flux carrying enthalpy parallel to the waveguide axis in a clockwise direction.

Finally, the use of mass flux measurements as function of Q_H and the use of acoustic wave amplitude measurements as functions of Q_H (see Fig. 2), provide the dependence of M as a function of $|\tilde{p}^+|$, for example. The acoustically induced mass flux per unit area $M/(\pi r_i^2)$ (mass flux M normalized to the cross-section area πr_i^2 of the pipe) is plotted as a function of $|\tilde{p}^+|^2$ in Fig. 5, where circles are measured values. Error bounds in Fig. 5 correspond to minimal and maximal values of the mass flux obtained from the latter mapping process, when the $\pm 5\%$ relative error is applied to measured values of all temperatures.

IV. MASS FLUX THEORETICAL PREDICTION

In a previous paper,⁹ a theory was proposed to predict the acoustically induced mass flux in an annular TA engine. The purpose of this paper was to find the dominant physical behavior of the acoustically induced mass flux. First, the stack was assumed to be acoustically thin (the dimensionless parameter $\mu \equiv k_c H_S \ll 1$ was supposed to be small). Second, the porosity was set equal to one ($\phi = 1$). From an experimental point of view, these assumptions appear too constraining (experimentally: $\mu \approx 0.42$ and $\phi \approx 0.8$). In order to take into account these experimental constraints, an improvement of the previous theory⁹ is proposed in the Appendix. Furthermore, another improvement should be considered concerning the contribution of the anticlockwise traveling

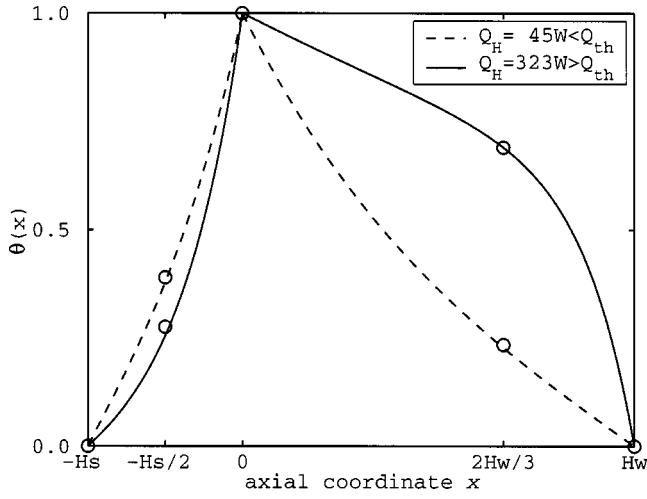


FIG. 4. Two examples of axial temperature distribution $\theta(x)$ in the TA core, below (dashed line $Q_H \approx 45$ W) and above (solid line $Q_H \approx 323$ W) the threshold condition ($Q_{th} \approx 130$ W). Curves are calculated using solution of Eq. (7), where constant A is known, coefficients h_S and h_W are determined from temperature measurements below the onset (respectively, at $x = -H_S/2$ and $x = 2H_W/3$ when $M=0$), and the mass flux M is extracted below the onset from temperature measurements at $x = 2H_W/3$. Circles are measured values, error bounds being not shown.

wave. Although the measured ratio of contrapropagating waves $|R| = |\tilde{p}^-/\tilde{p}^+| \approx 0.47$ is not negligible, the influence of the anticlockwise wave was not analyzed in the model described in Ref. 9.

In accordance with the theory⁹ and the Appendix, terms proportional to the thickness $\mu \ll 1$ of the stack are first neglected (porosity $\phi \approx 0.8$ is taken into account). The theoretical expression for the acoustically induced mass flux per unit area can be derived in the form of Eq. (A14)

$$M_0 \approx \Delta |\tilde{p}^+|^2. \quad (8)$$

From Eq. (8), the predicted mass flux M_0 is proportional to the square of acoustic pressure amplitude, where the proportionality coefficient Δ is given by

$$\Delta = \left(\frac{2/\phi}{2.37} \right) \frac{(c+d/2)}{12\pi\rho_C a_C^3} \left(\frac{D_S}{\delta_v} \right) \left(\frac{L}{H_S} \right) \left(\frac{\beta+1}{\beta-1} \right) \times \frac{[(T_H/T_C)^{(\beta-1)/2} - 1]}{H_S^{-1} \int_{-H_S}^0 dx T_N^{\beta+1}(x)}. \quad (9)$$

The density of air at cold temperature is $\rho_C \approx 1.2$ kg/m³. The phenomenological parameter $\beta \approx 0.73$ ^{1,12} takes into account that kinematic viscosity and thermal diffusivity of air depend on temperature: $\nu \propto \kappa \propto T_m^{\beta+1}(x)$. The coefficient $d = (1/\sqrt{\sigma})[2/(1+\beta)/(1+\sqrt{\sigma}) - (\gamma-1)] \approx 0.28$ is the Kramer's constant.^{1,12}

An improvement of the theory is achieved by considering correction terms proportional to the thickness μ of the stack. The mass flux is obtained in the form of Eq. (A17)

$$M_1 \approx (\Delta + \Delta') |\tilde{p}^+|^2, \quad (10)$$

where the additional coefficient Δ' is:

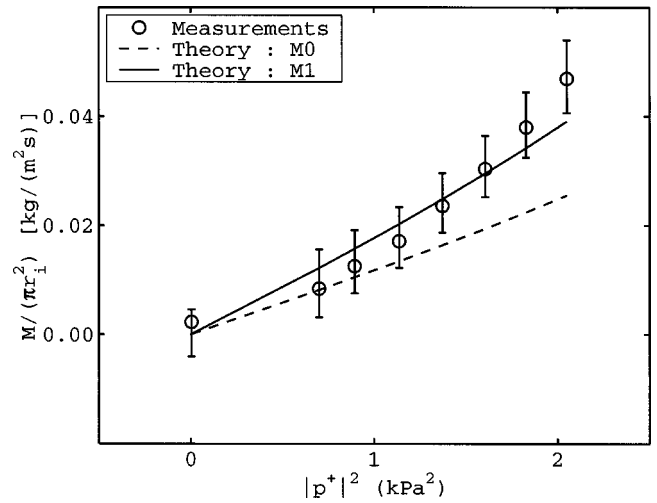


FIG. 5. Acoustically induced mass flux normalized to the cross section of the waveguide $M/(\pi r_i^2)$ plotted versus square of the clockwise traveling wave measured amplitude $|\tilde{p}^+|^2$. Circles with error bounds are estimated values from temperature measurements. Lines are the theoretical estimations given in Sec. IV. The dashed line is the lowest approximation M_0 of the mass flux induced by the clockwise acoustic wave. The solid line corresponds to the improved approximation M_1 of the acoustically induced mass flux.

$$\Delta' = \left(\frac{2}{2.37} \right) \frac{(3-2c)}{6\rho_C a_C^3} \left(\frac{D_S}{\delta_v} \right) \frac{\int_{-H_S}^0 dx T_N^{[(\beta-1)/2]}(x)}{\int_{-H_S}^0 dx T_N^{\beta+1}(x)}. \quad (11)$$

Theoretical estimates of mass fluxes M_0 and M_1 are plotted as functions of $|\tilde{p}^+|^2$ and compared to measured value in Fig. 5. Orders of magnitude of experimental and theoretical estimates are in a satisfactory agreement, the most important finding being that the acoustically induced mass flux dependence on acoustic wave amplitude $|\tilde{p}^+|$, which is close to quadratic, has been experimentally demonstrated. The lowest approximation M_0 of the theoretical acoustically induced mass flux appears to be insufficient to fit experimental data. In contrast, the improved theoretical prediction M_1 obtained by considering correction terms proportional to the thickness of the stack is in a better agreement with experimental data. Nevertheless, the influence of the anticlockwise wave has been neglected: this can explain why the predicted mass flux M_1 is still slightly lower than the experimental finding at high acoustic pressure amplitude.

Thus, it can be concluded that the qualitative thermal model presented in this paper provides a reasonable explanation for the experimental observations. It should be mentioned that in the region between the heat exchangers there also exists additional acoustic streaming that does not contribute to the cross-sectional average mass flux but may participate in the heat transport.⁴ Our experimental findings and their interpretation demonstrate that in the system under consideration this additional acoustic streaming contribution to heat transfer between the heat exchangers is not significant. This correlates with our theoretical estimates based on the theory of the Rayleigh-type streaming.^{11,16} It is predicted that, even though in our experiments the characteristic velocity of this streaming is of the order of $M/(\pi r_i^2)/\rho_C$, the enthalpy flux carried by this streaming is nevertheless at least

two orders of magnitude weaker than the one carried by the annular streaming. There are two reasons for this. First, the Rayleigh-type streaming velocity near the walls of the tube and near the axis of the waveguide are in opposite directions. The enthalpy flux is significantly canceled in the cross section and $M=0$. Second, the measured radial temperature gradients¹³ are estimated to be small in our experiments.

V. CONCLUSION

The acoustically induced mass flux in the annular thermoacoustic engine using quasiadiabatic stack, found to be in the range of 0.01 kg/(m² s) to 0.05 kg/(m² s), matches recent theoretical results⁹ and their improvements. This needs to be included in modeling annular thermoacoustic engines and should be taken into account in the optimization of thermoacoustic devices.

ACKNOWLEDGMENTS

The authors are indebted to E. Egon for technical contribution. This work was supported by the Délégation Générale de l'Armement (DGA) under Contract No. 99.34.072/DSP and through the DGA-CNRS Ph.D. fellowship of the first author (S. Job).

APPENDIX: IMPROVED THEORY OF ACOUSTIC STREAMING IN ANNULAR THERMOACOUSTIC ENGINES

According to recent theoretical results⁹ devoted to the study of an annular TA engine driven by a stack composed of parallel plates, the acoustically induced mass flux is expressed in terms of a density of streaming source divided by a hydrodynamic resistance. Dominant contributions of both these quantities are due to the stack. For a stack composed of parallel plates spaced by the distance D_S , the acoustically induced mass flux per unit area is given by

$$M \approx \left[\int_{-H_S}^0 dx s(x) \right] / \left[(2/D_S)^2 \int_{-H_S}^0 dx \nu_m(x) \right], \quad (\text{A1})$$

where $\nu_m(x) \propto T_m^{(\beta+1)}(x)$ is the kinematic viscosity at temperature $T_m(x)$. The denominator represents the hydrodynamic resistance. The density of streaming source $s(x)$ is found by substituting Eq. (20) in Eq. (17) of Ref. 9. In the quasiadiabatic regime, $f_v \approx (1+i)[\delta_v/D_S]$ and $D_S \gg \delta_v$. Considering terms proportional to $|f_v|$, the density of streaming source is given by

$$s(x) = \frac{1}{6\omega^2 \rho_m} \text{Re} [T_N^{-1} \partial_x T_N |\partial_x \tilde{p}|^2 + \partial_{xx}^2 \tilde{p} \partial_x \tilde{p}^* + k_C^2 T_N^{-1} \tilde{p} \partial_x \tilde{p}^* (1 - 3f_v^*)] + o(|f_v|), \quad (\text{A2})$$

where $T_N \equiv T_m/T_C$ is a normalized temperature and $\rho_m \propto 1/T_N$ is the density of the fluid. The quantity $f_v \propto T_m^{(\beta+1)/2}$ depends on the temperature and consequently on the position x . The dependence of $\tilde{p}(x)$, $T_N(x)$, $\rho_m(x)$, and $f_v(x)$ on x is taken into account but has been omitted in Eq. (A2) to simplify the expression.

For a stack composed of square channels (transversal width D_S), the volume of the boundary layers in the quasiadiabatic regime is approximately twice compared to a stack of parallel plates, and the hydrodynamic resistance is 2.37 greater.¹⁷ The mass flux per unit area induced in a stack of square channels is then

$$M_{\text{square}} \approx (2/2.37) M_{\text{plates}}. \quad (\text{A3})$$

It is now necessary to find the axial dependence of the acoustic field $\tilde{p}(x)$. In the quasiadiabatic regime, the acoustic wave propagation equation^{1,9,10,12} is

$$\partial_{xx}^2 \tilde{p} + [1 - (1/2)(\beta+1)(2c+d)f_v] T_N^{-1} \partial_x T_N \partial_x \tilde{p} + k_C^2 T_N^{-1} (1 + 2cf_v) \tilde{p} = 0 + o(|f_v|), \quad (\text{A4})$$

where c and d are, respectively, the Kirchhoff's and the Kramer's constants.^{1,12} The acoustic field in the stack region is sought in the same way as in Ref. 12. Equation (A4) is rewritten by neglecting contributions proportional to $|f_v| \ll 1$, that is

$$\partial_x [T_N \partial_x \tilde{p}] = -k_C^2 \tilde{p} + o(|f_v|^0). \quad (\text{A5})$$

Equation (A5) is then integrated over x , leading to

$$T_N(x) \partial_x \tilde{p}(x) = T_N(x_0) \partial_x \tilde{p}(x_0) - k_C^2 \int_{x_0}^x dx_1 \tilde{p}(x_1) + o(|f_v|^0), \quad (\text{A6})$$

and, once again, Eq. (A6) is integrated over x , giving

$$\tilde{p}(x) = \tilde{p}(x_0) + \partial_x \tilde{p}(x_0) \int_{x_0}^x dx_1 \frac{T_N(x_0)}{T_N(x_1)} - k_C^2 \int_{x_0}^x \frac{dx_1}{T_N(x_1)} \int_{x_0}^{x_1} dx_2 \tilde{p}(x_2) + o(|f_v|^0). \quad (\text{A7})$$

Choosing $x_0 \equiv -H_S$, the temperature $T_N(x_0)$ reduces to $T_N(-H_S) = 1$. Each integration over $-H_S < x < 0$ introduces a correction of the order of μ . The first integral in the r.h.s. of Eq. (A6) introduces a correction of the order μ , while the first term of the r.h.s. is of the order of μ^0 . Equation (A7) is an integral formulation of the differential Eq. (A4). The first term of the r.h.s. of Eq. (A7) corresponds to the order of magnitude μ^0 , the second one to μ , and the last one to μ^2 . By substituting $\tilde{p}(x)$ in the r.h.s. with $\tilde{p}(x_0)$, Eq. (A7) becomes a solution of Eq. (A4), approximated at the order μ^2 . Substituting Eq. (A4) in Eq. (A2) leads to

$$s(x) = \frac{1}{6\omega^2 \rho_C} \text{Re} [(1/2)(\beta+1)(2c+d)f_v \partial_x T_N |\partial_x \tilde{p}|^2 - k_C^2 (2cf_v + 3f_v^*) \tilde{p} \partial_x \tilde{p}^*] + o(|f_v|). \quad (\text{A8})$$

Using the quantity $R \equiv \tilde{p}^-(-H_S) / \tilde{p}^+(-H_S)$ and using the lowest order of magnitude of Eqs. (1), (A6), and (A7), one can check that

$$\left| \frac{k_c^2 \tilde{p} \partial_x \tilde{p}^*}{\partial_x T_N |\partial_x \tilde{p}|^2} \right| \sim \left| \left(\frac{1+R}{1-R} \right) \frac{k_c T_N}{\partial_x T_N} \right| \sim \mu. \quad (\text{A9})$$

The second term in square brackets of Eq. (A8), proportional to the thickness $\mu \ll 1$ of the stack, represents a small correction of the first one.

The porosity $\phi < 1$ of the stack must now be considered. First, the hydrodynamic resistance of the stack increases by a factor $(1/\phi)$. The mass flux per unit area in the waveguide region is then

$$M(\phi < 1) = \phi M(\phi = 1). \quad (\text{A10})$$

Second, the acoustic pressure $\tilde{p}(x)$ in the l.h.s. of Eqs. (A6) and (A7) represents the acoustic pressure inside the stack ($-H_S < x < 0$), and the quantity

$$T_N(x) \partial_x \tilde{p}(x) \approx i \omega \rho_c \tilde{v}(x) + o(|f_v|^0), \quad (\text{A11})$$

is proportional to the particle velocity $\tilde{v}(x)$. Therefore, the porosity of the stack induces a jump of the acoustic pressure gradient at the extremity of the stack, the acoustic field being measured at $x = -H_S - 0$ (just outside the left side of the stack)

$$\partial_x \tilde{p}(-H_S + 0) = (1/\phi) \partial_x \tilde{p}(-H_S - 0). \quad (\text{A12})$$

In Ref. 9, the stack is first assumed to be acoustically thin ($\mu \ll 1$): contributions proportional to μ are neglected in Eqs. (A6), (A7), and (A8). As a consequence, Eq. (A6) reduces to

$$T_N(x) \partial_x \tilde{p}(x) \approx (1/\phi) \partial_x \tilde{p}(-H_S) \approx (ik_c/\phi) \tilde{p}^+(1-R), \quad (\text{A13})$$

where the second line of Eq. (A13) has been derived using the expression of acoustic pressure defined in Eqs. (1) and (2) (neglecting terms proportional to $|f_v|$). Finally, neglecting the anticlockwise wave compared to the clockwise one ($|R| \ll 1$), the theoretical acoustically induced mass flux finally becomes

$$M_0 \approx \Delta |\tilde{p}^+|^2, \quad (\text{A14})$$

where the coefficient Δ is given in Eq. (9). The theoretical mass flux M_0 correspond to the one given in Ref. 9, subscript 0 indicating that Eq. (A14) is the theoretical value at the lowest order.

An improvement is derived by considering terms proportional to the length μ of the stack, the anticlockwise wave still being neglected ($|R| \ll 1$). Acoustic field is expressed at the necessary order of magnitude

$$p(x) \approx \tilde{p}^+, \quad (\text{A15})$$

$$T_N(x) \partial_x \tilde{p}(x) \approx (ik_c/\phi) \tilde{p}^+, \quad (\text{A16})$$

and the mass flux becomes

$$M_1 \approx (\Delta + \Delta') |\tilde{p}^+|^2, \quad (\text{A17})$$

where the coefficient Δ' is given in Eq. (11). Subscript 1 indicates that Eq. (A17) corresponds to the improved theory.

Concerning the contribution of the anticlockwise wave, it would be possible to obtain a better estimation of the acoustically induced mass flux. Although this improvement is tractable, it is beyond the scope of our qualitative approach. In fact, taking into account the anticlockwise wave is not trivial because the experimental value of the ratio $|R| \approx 0.47$ is close to the value of parameters $\mu \approx 0.42$ and $|f_v| \approx 0.37$. This further improvement might introduce inaccurate correction terms of the order of magnitude of neglected terms $|f_v||R| \sim |f_v|^2 \ll |f_v|$ in M_0 and $\mu|f_v||R| \sim |f_v|^3 \ll |f_v|$ in M_1 . More accurate expressions of the density of streaming sources and of the acoustic field need to be considered at least by taking into account for terms proportional to $|f_v|^2$, respectively, in Eqs. (A2) and (A4).

- ¹N. Rott, "Thermoacoustics," *Adv. Appl. Mech.* **20**, 135–175 (1980).
- ²P. H. Ceperley, "A pistonless Stirling engine—the traveling wave heat engine," *J. Acoust. Soc. Am.* **66**, 1508–1513 (1979).
- ³T. Yazaki, A. Iwata, T. Maekawa, and A. Tominaga, "Traveling wave thermoacoustic engine in a looped tube," *Phys. Rev. Lett.* **81**, 3128–3132 (1998).
- ⁴S. Backhaus and G. W. Swift, "A thermoacoustic Stirling engine," *Nature (London)* **399**, 335–338 (1999); "A thermoacoustic-Stirling heat engine: Detailed study," *J. Acoust. Soc. Am.* **107**, 3148–3166 (2000).
- ⁵T. Yazaki, T. Biwa, and A. Tominaga, "A pistonless Stirling cooler," *Appl. Phys. Lett.* **80**, 157–159 (2002).
- ⁶R. S. Reid, W. C. Wards, and G. W. Swift, "Cyclic thermodynamics with open flow," *Phys. Rev. Lett.* **80**, 4617–4620 (1998).
- ⁷R. S. Reid and G. W. Swift, "Experiments with a flow-through thermoacoustic refrigerator," *J. Acoust. Soc. Am.* **108**, 2835–2842 (2000).
- ⁸D. Gedeon, "DC gas flow in Stirling and pulse-tube cryocoolers," *Cryogenics* **9**, 385–392 (1997).
- ⁹V. Gusev, S. Job, H. Bailliet, P. Lotton, and M. Bruneau, "Acoustic streaming in annular thermoacoustic prime-movers," *J. Acoust. Soc. Am.* **108**, 934–945 (2000).
- ¹⁰G. W. Swift, "Thermoacoustic engines," *J. Acoust. Soc. Am.* **84**, 1145–1180 (1988).
- ¹¹J. R. Olsen and G. W. Swift, "Acoustic streaming in pulse tube refrigerators: Tapered pulse tubes," *Cryogenics* **37**, 769–776 (1998).
- ¹²V. Gusev, H. Bailliet, P. Lotton, and M. Bruneau, "Asymptotic theory of nonlinear acoustic waves in a thermoacoustic prime-mover," *Acust. Acta* **86**, 25–38 (2000).
- ¹³G. Penelet, E. Gaviot, V. Gusev, P. Lotton, and M. Bruneau, "Experimental investigation of transient nonlinear phenomena in an annular thermoacoustic prime-mover: Observation of a double-threshold effect," *Cryogenics* **42**, 527–532 (2002).
- ¹⁴F. P. Incropera and D. P. De Witt, *Introduction to Heat Transfer*, 3rd ed. (Wiley, New York, 1996).
- ¹⁵L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, Oxford, 1982).
- ¹⁶H. Bailliet, V. Gusev, R. Raspet, and R. A. Hiller, "Acoustic streaming in closed thermoacoustic devices," *J. Acoust. Soc. Am.* **110**, 1808–1821 (2001).
- ¹⁷F. M. White, *Fluid Mechanics*, 2nd ed. (McGraw-Hill, New York, 1986).

Analysis of axially polarized piezoelectric cylinders with arbitrary boundary conditions on flat surfaces

D. D. Ebenezer^{a)} and R. Ramesh

Naval Physical and Oceanographic Laboratory, Kochi 682 021, India

(Received 31 July 2002; revised 2 January 2003; accepted 24 January 2003)

A method to analyze axially polarized piezoelectric ceramic cylinders with arbitrary length to radius ratio and arbitrary boundary conditions on the flat surfaces is presented. The method is based on the use of functions that individually exactly satisfy the axisymmetric equations of motion and the Gauss electrostatic conditions. The axial and radial components of displacement and the potential are expressed as weighted sums of these functions. The functions form complete sets in the radial direction and arbitrary boundary conditions can, therefore, be satisfied on the flat surfaces. On the curved surface, certain uniform boundary conditions can be exactly satisfied. The weights are easily determined by using the orthogonal property of the functions. The input electrical admittance is a function of only the average boundary conditions and is very easily determined. Several special cases and numerical results are presented for electrically and mechanically excited cylinders to illustrate the method. The series expressions for displacements, potential, and stress converge rapidly and they are in good agreement with results obtained using ATILA—a finite element program. The method can be extended to analyze hollow cylinders with arbitrary boundary conditions on all surfaces. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1560171]

PACS numbers: 43.38.Ar, 43.38.Fx, 43.20.Bi [AJZ]

I. INTRODUCTION

Various methods have been used to analyze axially polarized piezoelectric cylinders of finite length and arbitrary length to radius ratio. Lloyd and Redwood¹ used a finite difference approach. Eer Nisse² studied cylinders with electrodes on all surfaces by using a variational approach. Paul and Natarajan³ also studied the free vibration of a cylinder with electrodes on all surfaces by using orthogonal functions. Hussein and Heyliger⁴ used a semianalytic discrete layer model to determine the resonance frequencies of cylinders with electrodes only on the flat surfaces. Several investigators⁵ have used finite element analysis.

Holland and Eer Nisse⁶ presented a variational treatment of elastic and piezoelectric structures of arbitrary shape. One of the structures used to illustrate the approach was a piezoelectric cylinder with electrodes only on the flat surfaces. The forced response of the structure was expressed as a weighted sum of eigenfunctions. The weights were determined by using variational methods. They showed that either short-circuit or open-circuit eigenfunctions can be used. The eigenfunctions themselves were expressed as a weighted sum of trial functions and the weights were again determined by using a variational approach. Complete sets of frequency-independent trial functions that did not satisfy the governing equations were used in all cases. They found that a very large number of trial functions was needed to yield solutions convergent to the exact modes when the functions have a zero gradient on the boundary but the actual eigenfunctions do not. Faster convergence was obtained by using an over-complete set of trial functions. In some cases, the use of trial

functions that are complete only in the norm-wise sense and not in the point-wise sense⁷ yielded convergence to the wrong values. For the special case of axially polarized cylinders with electrodes only on the flat surfaces, Holland and Eer Nisse used over-complete sets of trial functions for the axial and radial components of displacement and a point-wise complete set for potential. Their results were in good agreement with experimental results. In general, they found that fewer terms were required to accurately find resonance frequencies than to find displacements.

Ebenezer and Abraham⁸ used eigenfunctions to analyze radially polarized cylindrical membranes. They showed that the solution converged to the wrong value when the radial displacement was specified to be zero at the ends of the cylinder and the displacement was expressed as a weighted sum of eigenfunctions each of whose radial components was zero at the ends. They pointed out that it could be concluded that the solution was incorrect without comparing it with the solution obtained using some other method because the radial displacement distribution had discontinuities at the ends. Correct solutions, without discontinuities in the radial displacement, were obtained by expressing the displacement as a weighted sum of eigenfunctions plus the low-frequency solution. They have also analyzed radially⁹ and axially¹⁰ polarized thin piezoelectric cylindrical shells with internal loss.

Recently, Ebenezer and Ramesh¹¹ presented exact solutions to the governing equations of axially polarized piezoelectric cylinders that were apparently overlooked by Paul.¹² They showed that only these solutions were required to exactly satisfy certain uniform boundary conditions and approximately satisfy others in an average sense.

In this paper, axially polarized piezoelectric ceramic cylinders with arbitrary boundary conditions on the flat sur-

^{a)}Electronic mail: tsonpol@vsnl.com

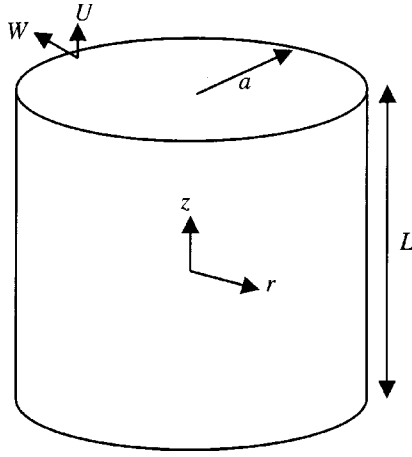


FIG. 1. An axially polarized solid piezoelectric ceramic cylinder of length L and radius a . The flat surfaces are fully electroded.

faces, arbitrary length to radius ratio, and electrodes on only the top and bottom flat surfaces are analyzed using the solutions to the governing equations presented in Refs. 11 and 12. Axial and radial components of displacement and potential are used as primary variables and are expressed as weighted sums of only certain exact axisymmetric solutions to the exact equations of motion and electrostatic conditions. In the set of solutions used, the functions describing the axial displacement are orthogonal to each other and form a point-wise complete set in the radial direction. Similarly, the functions describing radial displacement and potential, respectively, form point-wise complete sets in the radial direction. However, shear stress and radial component of electric displacement are expressed in terms of norm-wise complete sets of functions. Arbitrary boundary conditions can, therefore, be satisfied on the flat surfaces. The weights are determined by satisfying the boundary conditions. On the curved surface, only certain uniform boundary conditions can be exactly satisfied and arbitrary boundary conditions can be satisfied in an average sense. On the flat surfaces, certain uniform boundary conditions can be exactly satisfied and arbitrary boundary conditions can be satisfied in a mean square error sense. Interestingly, the input electrical admittance is expressed not as a series but as the sum of two terms only. It depends only on average boundary conditions. Special cases are presented and the solutions converge rapidly to those solutions obtained using ATILA¹³—a finite element program.

II. THEORETICAL ANALYSIS

Consider an axially polarized piezoceramic solid cylinder of finite length L and radius a as shown in Fig. 1. The top and bottom flat surfaces are completely electroded. All boundary conditions are axisymmetric. On the flat surfaces, arbitrary boundary conditions are specified. On the curved surface, the shear stress is zero and the normal displacement is specified. The displacement, potential, and stress distributions within the cylinder and the input electrical admittance are of interest. They can be determined for nonaxisymmetric

boundary conditions by using Fourier series expansions in the tangential direction.

The excitation and, therefore, the response of the cylinder are axisymmetric. The exact equilibrium equations are expressed in cylindrical coordinates (r, θ, z) as

$$\frac{\partial T_{zz}}{\partial z} + \frac{\partial T_{rz}}{\partial r} + \frac{1}{r} T_{rz} = -\rho \omega^2 U \quad (1a)$$

and

$$\frac{\partial T_{rr}}{\partial r} + \frac{\partial T_{rz}}{\partial z} + \frac{1}{r} (T_{rr} - T_{\theta\theta}) = -\rho \omega^2 W, \quad (1b)$$

where U and W are the axial and radial displacements, respectively, T_{rr} , T_{zz} , and $T_{\theta\theta}$ are normal components of stress, T_{rz} is a shear component of stress, ρ is the density, and ω is the angular frequency. The components of strain are expressed exactly as

$$[S_{rr}, S_{\theta\theta}, S_{zz}, S_{rz}] = \left[\frac{\partial W}{\partial r}, \frac{W}{r}, \frac{\partial U}{\partial z}, \frac{\partial U}{\partial r} + \frac{\partial W}{\partial z} \right] \quad (2a)$$

and the components of electric field are expressed as

$$[E_r, E_z] = - \left[\frac{\partial \Phi}{\partial r}, \frac{\partial \Phi}{\partial z} \right], \quad (2b)$$

where Φ is the electric potential. The constitutive relations for a piezoelectric material of $6mm$ crystal class are written as

$$\begin{Bmatrix} T_{rr} \\ T_{\theta\theta} \\ T_{zz} \\ T_{rz} \\ D_r \\ D_z \end{Bmatrix} = \begin{bmatrix} c_{11}^E & c_{12}^E & c_{13}^E & 0 & 0 & -e_{31} \\ c_{12}^E & c_{11}^E & c_{13}^E & 0 & 0 & -e_{31} \\ c_{13}^E & c_{13}^E & c_{33}^E & 0 & 0 & -e_{33} \\ 0 & 0 & 0 & c_{44}^E & -e_{15} & 0 \\ 0 & 0 & 0 & e_{15} & \epsilon_{11}^S & 0 \\ e_{31} & e_{31} & e_{33} & 0 & 0 & \epsilon_{33}^S \end{bmatrix} \begin{Bmatrix} S_{rr} \\ S_{\theta\theta} \\ S_{zz} \\ S_{rz} \\ E_r \\ E_z \end{Bmatrix}. \quad (3)$$

where c_{11}^E , c_{12}^E , c_{13}^E , c_{33}^E , and c_{44}^E are the elastic stiffness coefficients, e_{31} , e_{33} , and e_{15} are the piezoelectric stress coefficients, and ϵ_{11}^S and ϵ_{33}^S are dielectric permittivity coefficients. The components of electric displacement, D_z and D_r , must satisfy the Gauss condition

$$\frac{\partial(rD_r)}{r\partial r} + \frac{\partial D_z}{\partial z} = 0. \quad (4)$$

The exact axisymmetric governing equations are obtained by substituting Eqs. (2)–(4) in Eq. (1) and expressed as¹²

$$\begin{bmatrix} c_{33}^E \frac{\partial^2}{\partial z^2} + c_{44}^E \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} \right] + \rho \omega^2 U & (c_{13}^E + c_{44}^E) \left[\frac{\partial^2}{\partial r \partial z} + \frac{1}{r} \frac{\partial}{\partial z} \right] & e_{33} \frac{\partial^2}{\partial z^2} + e_{15} \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} \right] \\ (c_{13}^E + c_{44}^E) \frac{\partial^2}{\partial r \partial z} & c_{11}^E \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} - \frac{1}{r^2} \right] + c_{44}^E \frac{\partial^2}{\partial z^2} + \rho \omega^2 W & (e_{31} + e_{15}) \frac{\partial^2}{\partial r \partial z} \\ e_{33} \frac{\partial^2}{\partial z^2} + e_{15} \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} \right] & (e_{31} + e_{15}) \left[\frac{\partial^2}{\partial r \partial z} + \frac{1}{r} \frac{\partial}{\partial z} \right] & -\epsilon_{33}^S \frac{\partial^2}{\partial z^2} - \epsilon_{11}^S \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} \right] \end{bmatrix} \begin{Bmatrix} U \\ W \\ \Phi \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}. \quad (5)$$

It is easily verified that

$$[U \quad W \quad \Phi]^T = [U_1 \quad W_1 \quad \Phi_1]^T + [U_2 \quad W_2 \quad \Phi_2]^T \quad (6a)$$

is the sum of two independent exact solutions to Eq. (5) where¹¹

$$\begin{Bmatrix} U_1 \\ W_1 \\ \Phi_1 \end{Bmatrix} = \begin{Bmatrix} A \sin(K_z z) \\ B J_1(K_r r) \\ A \frac{e_{33}}{\epsilon_{33}^S} \sin(K_z z) + Dz + E \end{Bmatrix}, \quad (6b)$$

$$K_z = \omega \sqrt{\frac{\rho}{c_{33}^E + \frac{e_{33}^2}{\epsilon_{33}^S}}}, \quad (6c)$$

and

$$K_r = \omega \sqrt{\frac{\rho}{c_{11}^E}}, \quad (6d)$$

and¹²

$$\begin{Bmatrix} U_2 \\ W_2 \\ \Phi_2 \end{Bmatrix} = \begin{Bmatrix} \sum_{m=1}^M \sum_{s=1}^3 A_{ms} J_0(k_{rm} r) \sin(k_{zms} z) \\ \sum_{m=1}^M \sum_{s=1}^3 A_{ms} \psi_{ms} J_1(k_{rm} r) \cos(k_{zms} z) \\ \sum_{m=1}^M \sum_{s=1}^3 A_{ms} \chi_{ms} J_0(k_{rm} r) \sin(k_{zms} z) \end{Bmatrix}, \quad (6e)$$

where J_μ is the μ th order Bessel function of the first kind. Equation (6e) is an exact solution for arbitrary values of k_{rm} , $m=1,2,3,\dots,M$. For each value of k_{rm} , the frequency-dependent values of k_{zms} are determined by substituting Eq. (6e) in Eq. (5) and equating the determinant of the resulting equation to zero. The characteristic equation is cubic in k_{zms}^2 . It is sufficient to retain only the three positive roots of the characteristic equation. ψ_{ms} and χ_{ms} are obtained by substituting Eq. (6e) and the known values of k_{zms} in the homogeneous Eq. (5).

Here, $k_{rm}a$ are chosen to be the roots of $J_1(k_{rm}a)=0$ and are approximately equal to 0,3.83,7.02,10.17,... for $m=0,1,2,3,\dots$, respectively, and $M=\infty$. When $m=0$, $k_{rm}a$ and radial displacement are zero but axial displacement and potential are nonzero and functions of $K_z z$. The $m=0$ term is not included in Eq. (6e) because this solution is contained¹¹ in Eq. (6b). It is noted that $J_\mu(k_{rm}r)$ $\mu=0$ or 1 and m

$=0,1,2,\dots$ form a point-wise complete set of functions when $\mu=0$ and a norm-wise complete set of functions when $\mu=1$, i.e.,¹⁴

$$\int_0^a J_\mu(k_{rm}r) J_\mu(k_{rn}r) r dr = \begin{cases} \frac{a^2}{2} J_0^2(k_{rm}a) & m=n \\ 0 & m \neq n \end{cases} \quad \mu=0 \text{ or } 1, \quad m=0,1,2,\dots \quad (7)$$

The axial displacement in Eq. (6e) is antisymmetric about the center of the cylinder. A symmetric axial displacement can be obtained by using \sin in place of \cos and \cos in place of \sin . The Bessel function of the second kind has not been included in Eq. (6e) in order to satisfy the finiteness condition at the origin. If the cylinder is hollow, this function should also be included.

Other quantities of interest are now easily determined by using Eqs. (2), (3), and (6). The normal components of stress are expressed as

$$\begin{aligned} T_{rr} = & A \left(\frac{E}{13} + \frac{e_{31}e_{33}}{\epsilon_{33}^S} \right) K_z \cos(K_z z) \\ & + B \left[c_{11}^E K_r J_0(K_r r) - (c_{11}^E - c_{12}^E) \frac{1}{r} J_1(K_r r) \right] + D e_{31} \\ & + \sum_{m=1}^M \sum_{s=1}^3 A_{ms} \left\{ [c_{11}^E \psi_{ms} k_{rm} \right. \\ & + (c_{13}^E + e_{31}\chi_{ms}) k_{zms}] J_0(k_{rm} r) \\ & + (c_{12}^E - c_{11}^E) \frac{1}{r} \psi_{ms} J_1(k_{rm} r) \left. \right\} \cos(k_{zms} z), \end{aligned} \quad (8)$$

and

$$\begin{aligned} T_{zz} = & A \left(e_{33}^E + \frac{e_{33}^2}{\epsilon_{33}^S} \right) K_z \cos(K_z z) + B c_{13}^E K_r J_0(K_r r) + D e_{33} \\ & + \sum_{m=1}^M \sum_{s=1}^3 A_{ms} (c_{13}^E \psi_{ms} k_{rm} + c_{33}^E k_{zms} \\ & + e_{33}\chi_{ms} k_{zms}) J_0(k_{rm} r) \cos(k_{zms} z). \end{aligned} \quad (9)$$

The shear stress is expressed as

$$\begin{aligned} T_{rz} = & \sum_{m=1}^M \sum_{s=1}^3 -A_{ms} [c_{44}^E (k_{rm} + \psi_{ms} k_{zms}) \\ & + e_{15}\chi_{ms} k_{rm}] J_1(k_{rm} r) \sin(k_{zms} z) \end{aligned} \quad (10)$$

and is zero on $r=a$.

The components of electric displacement are expressed as

$$D_z = B e_{31} K_r J_0(K_r r) - D \varepsilon_{33}^S + \sum_{m=1}^M \sum_{s=1}^3 A_{ms} [e_{31} \psi_{ms} k_{rm} + e_{33} k_{zms} - \varepsilon_{33}^S \chi_{ms} k_{zms}] J_0(k_{rm} r) \cos(k_{zms} z). \quad (11a)$$

and

$$D_r = \sum_{m=1}^M \sum_{s=1}^3 -A_{ms} [e_{15}(k_{rm} + \psi_{ms} k_{zms}) - \varepsilon_{11}^S \chi_{ms} k_{rm}] \times J_1(k_{rm} r) \sin(k_{zms} z). \quad (11b)$$

Then, the current is expressed as

$$I = - \int_0^a j \omega D_z 2 \pi r dr = - j \omega \pi a (2 B e_{31} J_1(K_r a) - D \varepsilon_{33}^S a) \quad (12)$$

because¹⁴

$$\int_0^a J_0(k_{rm} r) r dr = 0. \quad (13)$$

The integral in Eq. (12) can be evaluated at arbitrary z —as should be the case from a consideration of the Gauss zero-divergence condition in Eq. (4). It is noted that the current depends only on average boundary conditions, i.e., it is independent of A_{ms} . The input electrical admittance,

$$Y = \frac{I}{\Phi_0}, \quad (14)$$

is now easily obtained by using Eq. (12) and the specified applied voltage, Φ_0 .

Some observations are now made regarding the type of functions used. In the radial direction, U is expressed in terms of a point-wise complete set of functions: $J_0(k_{rm} r)$, $m=0,1,2,\dots$. The expression for Φ is over-complete as there are three terms in Eq. (6b) that are independent of r . As noted in Refs. 6 and 11, a potential gradient can exist even when the displacements are zero. Therefore, it is necessary to include the terms containing D and E in the expression for Φ . Similarly, the expression for T_{zz} is over-complete as the terms containing A and D in Eq. (9) are independent of r . The expression for W is also point-wise complete because even though all the terms containing A_{ms} are zero at $r=a$, the term containing B is not zero at the same point—except at the frequencies at which $J_1(K_r a)=0$. The expression for T_{rr} is over-complete in the r direction. However, it will not occur in boundary conditions on the flat surfaces. T_{rz} and D_r are expressed in terms of a norm-wise complete set of functions: $J_1(k_{rm} r)$. All of them are zero at $r=a$. It is also noted that $\partial U/\partial r$ and $\partial \Phi/\partial r$ are zero at $r=0$ and $r=a$.

U , Φ , and T_{zz} have been expressed in terms of complete sets of functions and boundary conditions on the flat surfaces involving them can, therefore, be satisfied in a mean-square-error sense. Expressions for D_r and T_{rz} are zero on the curved surface. Therefore, the curved surface is bounded by vacuum and the shear stress must be zero on it. After finding

the solution for a specified set of boundary conditions, T_{rz} and D_r must be examined for any discontinuity at $r=a$. If there is a discontinuity, the solution will be incorrect because the boundary condition only appears to be satisfied because of the discontinuity.

Special cases

Several special cases are considered now and the origin of the coordinate system is fixed at the center of the cylinder for convenience.

For all the cases, the potential at the bottom electrode is zero, i.e.,

$$\Phi = 0 \quad \text{on} \quad 0 \leq r \leq a, \quad z = -L/2, \quad (15a)$$

and the potential on the top electrode is expressed as

$$\Phi = \Phi_0 \quad \text{on} \quad 0 \leq r \leq a, \quad z = +L/2, \quad (15b)$$

where Φ_0 is one for the applied potential condition, zero for the short-circuit condition, and is to be determined for the open-circuit condition.

The electrical boundary conditions are satisfied in a mean-square sense. After substituting Eq. (6) in the electrical boundary conditions in Eq. (15), multiplying both sides by $J_0(k_m r)r$, integrating over r , and using the orthogonal property in Eq. (7), it is seen that

$$-A \frac{e_{33}}{\varepsilon_{33}^S} \sin(K_z L/2) - D L/2 + E = 0 \quad (16a)$$

and

$$A \frac{e_{33}}{\varepsilon_{33}^S} \sin(K_z L/2) + D L/2 + E = \Phi_0 \quad (16b)$$

when $n=0$ and

$$\sum_{s=1}^3 A_{ms} \chi_{ms} \sin(k_{zms} L/2) = 0 \quad (16c)$$

when $n=m=1,2,\dots$

Case 1. Applied potential $\Phi_0=1$, zero shear stress on all surfaces, zero axial stress on the flat surfaces, specified uniform displacement and zero normal component of electric displacement on the curved surface, i.e.,

$$T_{rz} = 0 \quad \text{on} \quad \begin{cases} 0 \leq r \leq a, & z = \pm L/2, \\ r = a, & |z| \leq L/2, \end{cases} \quad (17a)$$

$$T_{zz} = 0 \quad \text{on} \quad 0 \leq r \leq a, \quad z = \pm L/2, \quad (17c)$$

$$W = W_0 \quad \text{or} \quad r = a, \quad |z| \leq L/2, \quad (17d)$$

and

$$D_r = 0 \quad \text{on} \quad r = a, \quad |z| \leq L/2. \quad (17e)$$

The coefficients A , B , D , E and A_{ms} in Eq. (6) are now determined by using the boundary conditions in Eqs. (15) and (17).

Four boundary conditions are exactly satisfied. The zero shear stress boundary condition on the flat surfaces in Eq. (17a) is satisfied exactly by equating the coefficients of $J_1(k_{rm}r)$ in Eq. (10) to zero, i.e.,

$$\sum_{s=1}^3 A_{ms} [c_{44}^E(k_{rm} + \psi_{ms}k_{zms}) + e_{15}\chi_{ms}k_{rm}] \sin(k_{zms}L/2) = 0. \quad (18)$$

It is seen from Eqs. (10) and (11b) that the boundary conditions in Eqs. (17b) and (17e), respectively, are exactly satisfied because $J_1(k_{rm}a) = 0$. The specified uniform radial displacement condition in Eq. (17d) is also satisfied exactly. After substituting Eq. (6) in the uniform radial displacement boundary condition in Eq. (17d), multiplying the resulting equations by $J_0(k_{rm}r)r$, integrating over r , and using the orthogonal property in Eq. (8), it is seen that

$$B = \frac{W_0}{J_1(K_r a)}. \quad (19)$$

The zero axial stress boundary condition in Eq. (17c) is satisfied in a mean-square-error sense. It is seen by using Eq. (17c), the expression for axial stress in Eq. (9), and¹⁴

$$\int_0^a J_0(K_r r) J_0(k_{rm} r) r dr = \Re, \quad (20a)$$

where

$$\Re = \begin{cases} 0, & K_r = k_{rm}, \quad n \neq m \\ \frac{a^2}{2} J_0^2(k_{rm} a), & K_r = k_{rm} \\ \frac{K_r^2 a}{K_r^2 - k_{rm}^2} J_0(k_{rm} a) J_1(K_r a), & K_r \neq k_{rm}, \quad n = 1, 2, 3, \dots \end{cases} \quad (20b)$$

that

$$A \left(c_{33}^E + \frac{e_{33}^2}{\epsilon_{33}^S} \right) K_z \cos(K_z L/2) + D e_{33} = - \frac{2B c_{13}^E J_1(K_r a)}{a} \quad (21a)$$

and

$$\begin{aligned} a^2 J_0^2(k_{rm} a) / 2 \sum_{s=1}^3 A_{ms} [c_{33}^E k_{zms} + c_{13}^E \psi_{ms} k_{rm} \\ + e_{31} \chi_{ms} k_{zms}] \cos(k_{zms} L/2) \\ = B c_{13}^E K_r \Re, \quad m = 1, 2, 3, \dots \end{aligned} \quad (21b)$$

The coefficients A , D , and E are now easily determined by solving Eqs. (16a), (16b), and (21a) and expressed as

$$A = \frac{\frac{\Phi_0}{2} e_{33} + W_0 c_{13}^E \frac{L}{a}}{\frac{e_{33}^2}{\epsilon_{33}^S} \sin\left(\frac{K_z L}{2}\right) - \left(c_{33}^E + \frac{e_{33}^2}{\epsilon_{33}^S}\right) \frac{K_z L}{2} \cos\left(\frac{K_z L}{2}\right)}, \quad (22a)$$

$$D = \frac{\Phi_0}{L} - \frac{2A e_{33}}{L \epsilon_{33}^S} \sin\left(\frac{K_z L}{2}\right), \quad (22b)$$

and

$$E = \frac{\Phi_0}{2}. \quad (23)$$

The coefficients A_{ms} are now determined by solving the 3×3 inhomogeneous Eqs. (16c), (18), and (21b) for $m = 1, 2, \dots$. Then all functions of interest such as stress, electric displacement, IEA, etc., are computed using Eqs. (8)–(14). The number of terms in the series required for convergence increases with frequency. It is seen from Eq. (21b) that k_{rm} must be much greater than K_r for the series to converge where M is the number of terms in the series.

The more general case of specified nonuniform axial and shear stress on the flat surfaces can also be analyzed in a similar manner. This condition will exist when the flat surfaces are in contact with another solid as, for example, in sandwich transducers.

Case 2. Open circuit condition. All other boundary conditions are the same as that in case 1.

Φ_0 is not known and is eliminated from the expressions for the coefficients after equating the current to zero in Eq. (12). It is then seen that the open circuit potential can be expressed as

$$\Phi_0 = \frac{2W_0 L}{a} \frac{e_{31}/\epsilon_{33}^S - \Theta}{c_{33}^E + e_{33}^2/\epsilon_{33}^S}$$

where

$$\Theta = (c_{13}^E + e_{31} e_{33} / \epsilon_{33}^S) (e_{33} / \epsilon_{33}^S) \tan(K_z L/2) / (K_z L/2). \quad (24)$$

Case 3. Applied potential, $\Phi_0 = 1$, zero shear stress on all surfaces, specified axial displacement on the flat surfaces varying linearly with r , constant radial displacement on the curved surface and zero normal component of electric displacement on the curved surface, i.e., Eqs. (17a), (17b), and

$$U = \pm U_0 \left(1 - \frac{r}{a} \right) \quad \text{on } 0 \leq r \leq a, \quad z = \pm L/2, \quad (25)$$

and Eqs. (17d) and (17e). The electrical and shear stress boundary conditions for this case are the same as those for case 1. It therefore follows that Eqs. (16) and (18) are applicable for this case also.

After substituting Eq. (6) in the specified axial displacement condition in Eq. (25), multiplying both sides by $J_0(k_{rm}r)r$, $n = 0, 1, 2, \dots$, integrating over r , and using the orthogonal property in Eq. (7), it is seen that

$$A = \frac{U_0}{\sin(k_z L/2)} \quad (26a)$$

and

$$\begin{aligned} (a^2/2) J_0^2(k_{rm} a) \sum_{s=1}^3 A_{ms} \chi_{ms} \sin(k_{zns} L/2) \\ = - \int_0^a U_0 (r^2/a) J_0(k_{rm} r) dr, \quad m = 1, 2, \dots, \end{aligned} \quad (26b)$$

where the integral in the above equation is done numerically. The specified radial displacement condition is the same as

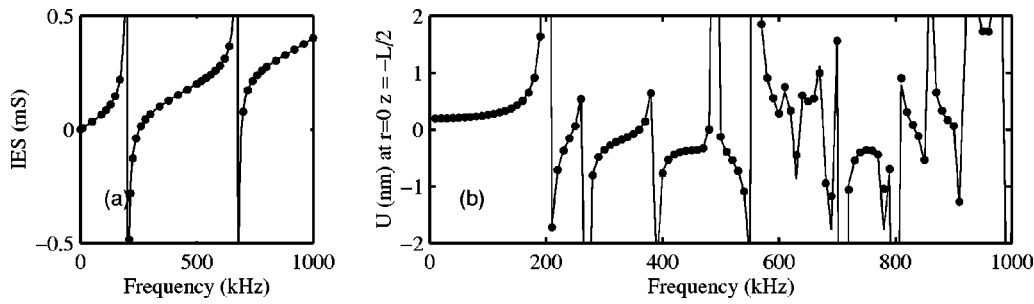


FIG. 2. (a) Input electrical susceptance and (b) axial displacement for case 1. Line—analytical; dots—ATILA.

Eq. (17d) and the expression for B in Eq. (19) is, therefore, valid for this case also.

The values of D and E are now obtained by using Eq. (26a) and solving Eqs. (16a), (16b) and the values of A_{ms} are obtained by solving the inhomogeneous Eqs. (18), (21b), and (26b). It is seen that

$$D = \frac{\Phi_0}{L} + \frac{2e_{33}}{Le_{33}^5} U_0 \quad (26c)$$

and

$$E = \frac{\Phi_0}{2}. \quad (26d)$$

When the specified radial displacement on the curved surface is zero, only B becomes zero. In contrast to case 2, A_{ms} does not become zero.

The number of terms needed for convergence in this case depends on the spatial variation of the specified normal displacement on the flat surfaces. Only a few terms are needed to approximate linear distributions of the type used in Eq. (25). The number of terms is independent of frequency because the specified displacement is independent of frequency.

III. NUMERICAL RESULTS AND DISCUSSION

Numerical results are presented to illustrate displacement, potential, stress, and electric displacement within the cylinder and to show that they are in good agreement with results obtained using ATILA.¹³ They are also used to show that the input electrical admittances obtained using the present approach and ATILA are in better agreement than field variables. In ATILA, second-order axisymmetric, rectangular, equisized elements were used.

All results shown here are for a cylinder of length 10 mm and radius 5 mm. The material properties used as input to ATILA are those given for Navy Type I piezoceramics by Wilson.¹⁵ The coefficients needed for the present approach were computed using the coefficients used in ATILA. ATILA results were obtained using $I \times J$ elements, i.e., I elements in the axial direction and J elements in the radial direction. All results obtained using the present approach are shown by lines and those obtained using ATILA are shown by dots. Only a few nodal values computed using ATILA have been shown in the figures for clarity.

Results are shown in Fig. 2 for case 1. The input electrical susceptance (IES) is shown in Fig. 2(a) and the axial

displacement, U , at $r=0$ and $z=-L/2$, is shown in Fig. 2(b). The resolution is 10 kHz. In Fig. 2(a), resonance occurs when D becomes infinity, i.e., at the nonzero frequencies at which the denominator of A in Eq. (22a) is zero. The analytical IES is not expressed in series form and is equal to 0.403 mS at 1 MHz. The ATILA IES is 0.402 mS at 1 MHz when 25×10 , 40×25 , 40×40 , and 50×25 elements are used. In Fig. 2(b), resonances occur when the determinant of the 3×3 matrix used to find A_{ms} becomes zero. The agreement between the present approach and ATILA is good except at frequencies that are very close to resonance frequencies in Fig. 2(b).

Field variables, for case 1 and W_0 is 0.1 nm, are shown at 100 kHz and 1 MHz in Figs. 3 and 4, respectively. There is no resonance near 100 kHz and the agreement between the present approach and ATILA is very good. There is a resonance in Fig. 2(b) at about 990 kHz and the agreement in Fig. 4 is not as good as the agreement in Fig. 3. In Fig. 4, the agreement at $r=0$ is not as good as that for $r>0$. This may occur because the weight function in Eq. (7) is zero at $r=0$. In Figs. 3 and 4, there is no discontinuity in any of the variables. In particular, each term in the series for T_{rz} and D_r is zero at $r=a$. However, Figs. 3(f) and 4(d), as well as Figs. 3(h) and 4(e), show that there is no discontinuity in the analytical solution for T_{rz} and D_r , respectively, at $r=a$. This and the agreement with ATILA show that the solutions obtained using the present approach have converged to the correct values.

Φ and T_{zz} are both zero at $z=-L/2$ and so they have been shown at $z=-L/5$ in Figs. 3(c) and (e), respectively. Similarly, W is specified on $r=a$ and is shown at $r=0.4a$ in Fig. 3(j), and T_{rz} is zero on $r=a$ and is shown at $r=0.8a$ in Fig. 3(n). In Fig. 4(d), T_{rz} is shown at $z=-L/5$ as it is zero at $z=-L/2$.

Figure 4 was obtained using $M=20$ in the present approach and 50×25 elements in ATILA. After finding the coefficients, the functions that are specified at the boundaries were computed using the truncated series and compared with the specified boundary conditions. In the present approach, four boundary conditions are satisfied exactly and two, on the flat surfaces, are satisfied only in a mean-square sense. The error, at any particular r , in satisfying the latter boundary conditions on the flat surfaces is very small but larger at $r=0$ and $r=a$ than at other values of r . This may be due to the constraints the form of the solution places on the functions, mentioned at the end of the description of the analytical model. As seen in Fig. 4(f), the error in satisfying the zero

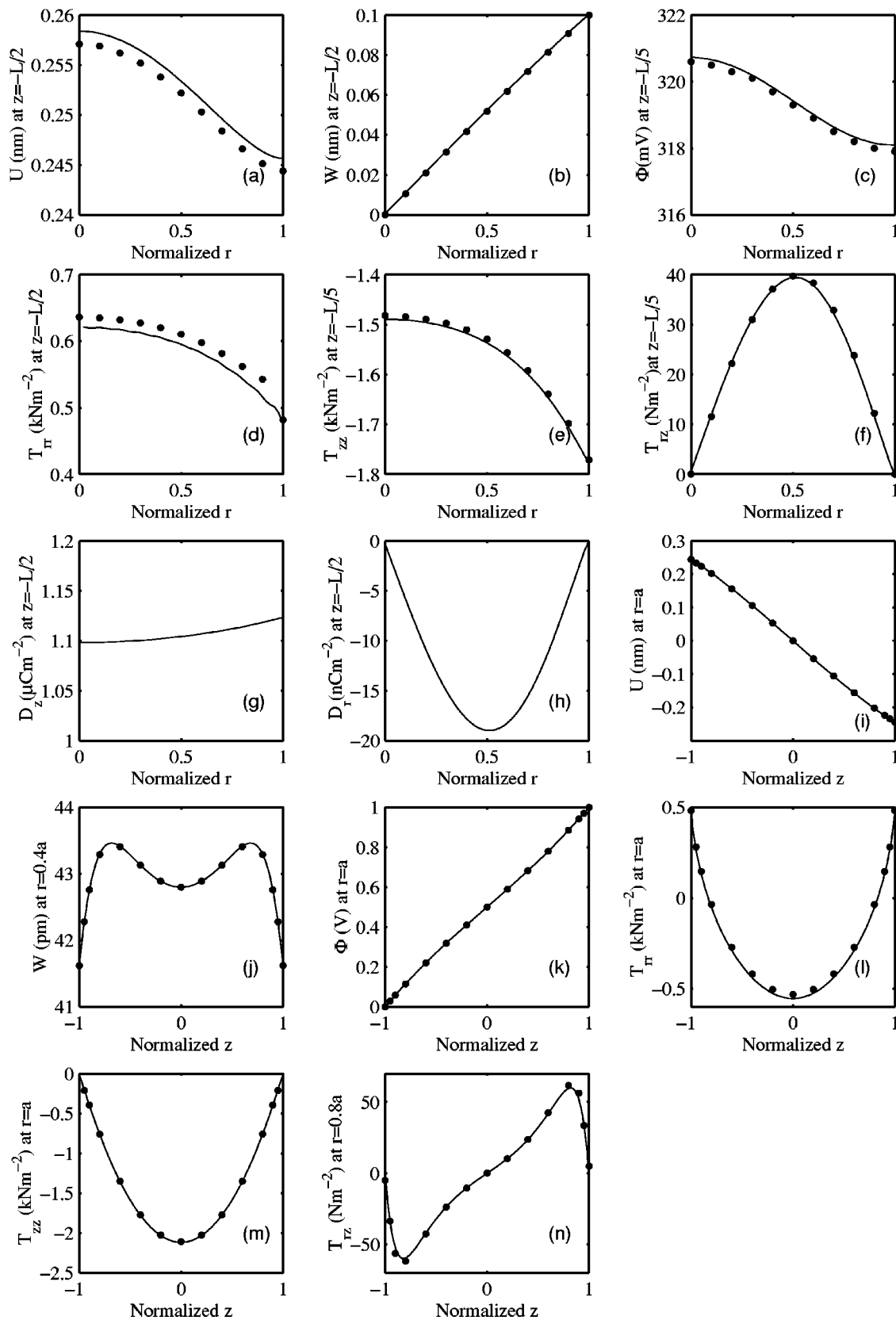


FIG. 3. Displacement, potential, stress, and electric displacement at 100 kHz for case 1. Line—analytical; dots—ATILA.

normal stress boundary condition is less than that in ATILA. The maximum error in ATILA is less than 1% of the maximum T_{zz} in the cylinder. The boundary conditions on Φ , a primary variable, are satisfied exactly in ATILA and approxi-

mately in the present approach. However, in the present approach, the error is less than 10^{-15} V at all points on the flat surfaces.

Increasing the number of terms in the series does not

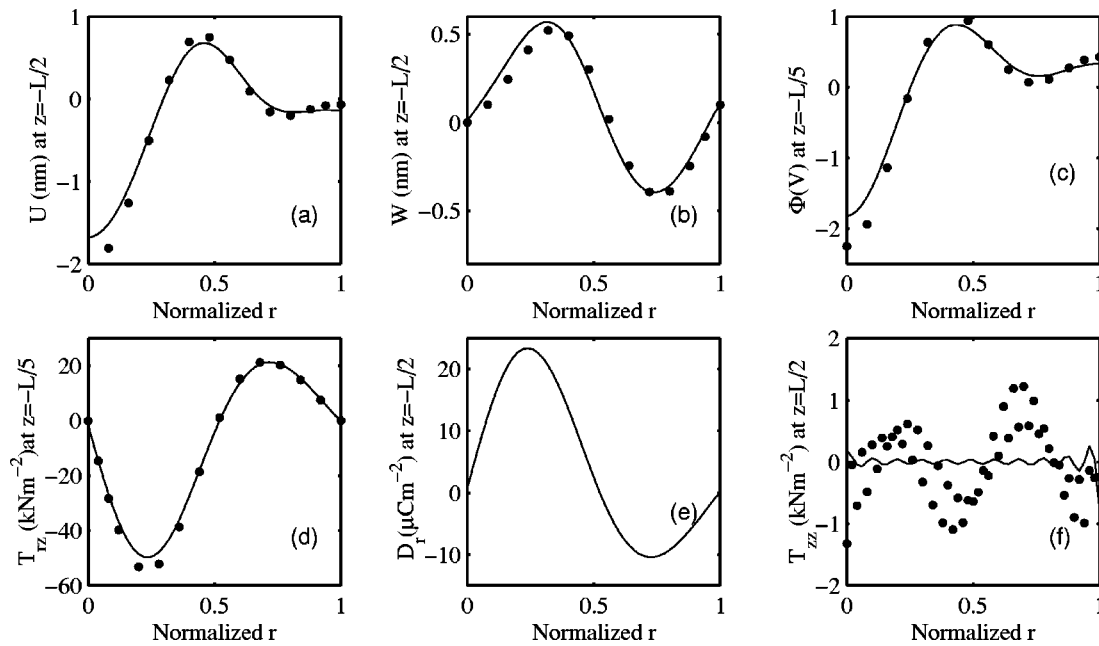


FIG. 4. Displacement, potential, stress, and electric displacement at 1 MHz for case 1. Line—analytical; dots—ATILA.

have a significant effect on the values computed using the present approach. For example, in Fig. 4, the values of U at $r=0$ and $z=-L/2$ for $M=5, 10, 15$, and 20 are $-1.666, -1.687, -1.684$, and -1.685 nm, respectively. At the same location, when $25 \times 10, 40 \times 25, 40 \times 40$, and 50×25 elements are used in ATILA, the axial displacement is $-2.60, -2.02, -2.02$, and -2.01 nm, respectively. Therefore, results obtained using both methods have converged. As seen in Fig. 4 agreement between the present results and ATILA is worst at this point.

The open circuit voltage is shown in Fig. 5 for case 2. Resonances are present and there is good agreement over the entire frequency band.

Results for case 3 are shown in Fig. 6 with $U_0=W_0=0.1$ nm. Here, values of A_{ms} are not zero but the agreement between the present approach and ATILA is excellent. $M=10$ was used in the present method. 40×20 elements were used in ATILA—fewer than in Fig. 4. Figures 6(d) and (e) are of particular interest as there is no discontinuity in T_{rz} and D_r , respectively, at $r=a$. T_{rz} is shown at $r=0.95a$ as it is zero at $r=a$. It varies sharply near the flat surfaces but there is good agreement between the analytical and numerical results.

IV. CONCLUSIONS

A method to analyze axially polarized cylinders with arbitrary length to radius ratio and arbitrary boundary conditions on the flat surfaces is presented. In the method norm-wise and point-wise complete sets of functions as well as over-complete sets of functions are used. However, no case of convergence to wrong values or slow convergence is encountered. No discontinuity in the analytical results is seen near $r=a$, the line on which all the functions of the norm-wise complete set are zero. The leading terms in the complete sets of functions were presented by Ebenezer and Ramesh¹¹ and the others were presented by Paul.¹² Numeri-

cal results obtained using the present approach are found to be in good agreement with those obtained using ATILA.

Special cases with zero stress and specified nonuniform normal displacement, respectively, on the flat surfaces are considered. Other cases where a combination of these functions is specified can also be satisfied because normal and shear stress and normal displacement are expressed in terms of complete sets of functions in the radial direction. It, therefore, follows that arbitrary boundary conditions, including continuity conditions that must be satisfied when another solid or fluid is in contact with the cylinder, can be satisfied on the flat surfaces.

When Bessel functions of the second kind are included in the solutions, the method can be used to analyze hollow cylinders with arbitrary boundary conditions on flat surfaces. Arbitrary boundary conditions on all surfaces can be satisfied by expressing the displacements and potential as the sum of the solutions used here and the solutions to the governing equations that form a complete set in the axial directions.

The approach presented here can be extended to include arbitrary boundary conditions on all surfaces and used to model tonpilz transducers, 1-3 piezocomposites, cylindrical

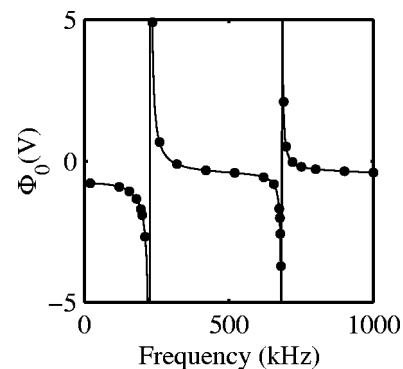


FIG. 5. Open-circuit potential for case 2. Line—analytical; dots—ATILA.

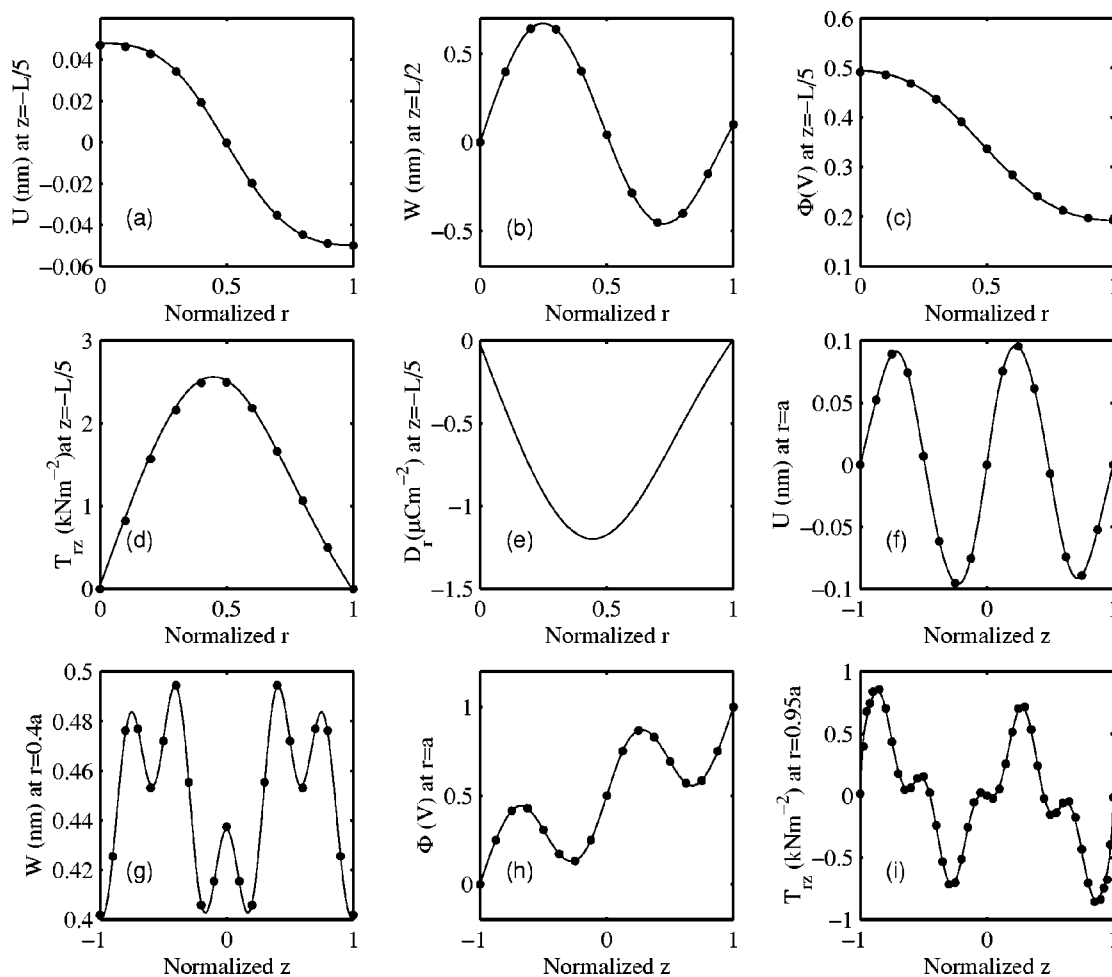


FIG. 6. Displacement and potential at 1 MHz for case 3. Line—analytical; dots—ATILA.

panels used in smart structures, etc. In each of these applications, continuity conditions must be satisfied at the interfaces and it is possible to do so because displacements and stresses are expressed in terms of complete sets of functions.

ACKNOWLEDGMENT

Permission from Director NPOL to publish this paper is gratefully acknowledged.

- ¹P. Lloyd and M. Redwood, "Finite difference method for the investigation of the vibrations of solids and the evaluation of the equivalent-circuit characteristics of piezoelectric resonators. I and II," *J. Acoust. Soc. Am.* **39**, 346–361 (1966).
- ²E. P. Eer Nisse, "Variational method for electroelastic vibration analysis," *IEEE Trans. Sonics Ultrason.* **SU-14**, 153–160 (1967).
- ³H. S. Paul and K. Natarajan, "Axisymmetric free vibrations of piezoelectric finite cylindrical bone," *J. Acoust. Soc. Am.* **96**, 213–220 (1994).
- ⁴M. Hussein and P. R. Heyliger, "Discrete layer analysis of axisymmetric vibrations of laminated piezoelectric cylinders," *J. Sound Vib.* **192**, 995–1013 (1996).
- ⁵R. Lerch, "Simulation of piezoelectric devices by two and three dimensional finite elements," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **37**, 233–247 (1990).
- ⁶R. Holland and E. P. Eer Nisse, *Design of Resonant Piezoelectric Devices*, Research Monograph No. 56 (MIT, Cambridge, 1969); "Variational evaluation of admittances of multielectroded three dimensional piezoelectric structures," *IEEE Trans. Sonics Ultrason.* **SU-15**, 119–132 (1968).

- ⁷Sets of functions that are complete in the ordinary Fourier or norm-wise sense and are not all zero at the same point are said to be point-wise complete.
- ⁸D. D. Ebenezer and P. Abraham, "Eigenfunction analysis of radially polarized piezoelectric cylindrical shells of finite length," *J. Acoust. Soc. Am.* **102**, 1549–1558 (1997).
- ⁹D. D. Ebenezer and P. Abraham, "Closed-form analysis of thin radially polarized piezoelectric ceramic cylindrical shells with loss," *Curr. Sci.* **83**, 981–988 (2002).
- ¹⁰D. D. Ebenezer and P. Abraham, "Analysis of axially polarized piezoelectric ceramic cylindrical shells of finite length with internal losses," *J. Acoust. Soc. Am.* **112**, 1953–1960 (2002).
- ¹¹D. D. Ebenezer and R. Ramesh, "Exact analysis of axially polarized piezoceramic cylinders with certain uniform boundary conditions," *Curr. Sci.* (submitted).
- ¹²H. S. Paul, "Vibrations of circular cylindrical shells of piezoelectric silver iodide crystals," *J. Acoust. Soc. Am.* **40**, 1077–1080 (1966).
- ¹³J. 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).
- ¹⁴*Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1965), Eqs. 11.3.20, 11.3.29, and 11.4.5.
- ¹⁵O. B. Wilson, *Introduction to the Theory and Design of Sonar Transducers* (Peninsula, Los Altos, CA, 1988).

On the dynamic stiffness of preloaded vibration isolators in the audible frequency range: Modeling and experiments

Leif Kari

MWL/Department of Vehicle Engineering, Royal Institute of Technology, 10044 Stockholm, Sweden

(Received 19 May 1999; accepted for publication 12 November 2002)

The nonlinear, preload-dependent dynamic stiffness of a cylindrical vibration isolator is examined via measurements and modeling within an audible frequency range covering 50 to 1000 Hz at various preloads. The stiffness is found to depend strongly on frequency—resulting in peaks and troughs, and on preload—particularly above 500 Hz. The problems of simultaneously modeling the rubber prestrain dependence and its audible short-term response are removed by adopting a nearly incompressible material model, being elastic in dilatation while displaying viscoelasticity in deviation. The latter exhibits a time strain separable relaxation tensor with a single function embodying its time dependence. This function is based on a continuous fractional order derivative model, the main advantage being the minimum number of parameters required to successfully model the rubber properties over a broad structure-borne sound frequency domain, while embodying a continuous distribution of relaxation time. The weak formulations corresponding to the stiffness problem are solved by an updated Lagrangian nonlinear finite-element procedure. The model and measurement results agree strikingly well with static and dynamic measurements throughout the whole frequency domain for the examined preloads. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1557214]

PACS numbers: 43.40.At, 43.40.Yq [PJR]

I. INTRODUCTION

Vibrational sources, such as machines, are frequently mounted upon vibration isolators to reduce the transmitted energy to receiving structures. Traditionally, a vibration isolator provides insulation designed to meet low-frequency requirements, while its structure-borne sound property remains arbitrary. However, increasing interest in noise abatement has heightened the need for effective isolation within the audible frequency domain, requiring suitable models for the isolators' audible dynamic stiffness.

The aim of the paper is to include the geometrical and material nonlinearities, aligned with finite isolator preloads, while properly modeling the influences of geometry and constraining metal plates in the audible frequency domain. This seemingly innocuous nonlinear extension is not an easy task; albeit extensively studied over the years, it remains a prodigious source of challenging open problems.

Particularly suitable in an engineering environment are the single integral visco-elastic constitutive theories: for example, Pipkin and Rivlin model,¹ finite linear visco-elastic model,^{2–4} Chang, Bloch, and Tschoegl model,⁵ modified finite linear visco-elastic model,⁶ and Zdunek model⁷ are all based on a solid type of material damping, while Bernstein, Kearsley, and Zappas model,⁸ generalized solid liquid model,⁹ Sullivan and Mazich model,¹⁰ Bush, Tanner, and Phan-Thien model¹¹ are all of a liquid type. Recently, Morman¹² extends to nonisothermal conditions while Govindjee and Simo^{13–15} include damage effects.

Suitable constitutive relations including material damping are found in linear visco-elastic models. The classical example for structure-borne sound analysis is the linear structural damping model with a frequency-independent loss factor.¹⁶ Recently, a number of promising linear visco-elastic

models have been proposed: the fractional derivative model,¹⁷ augmented thermal field model,^{18–21} augmented Hooke's law,²² generalized Maxwell model,²³ and mini-oscillator model.²⁴ The last four models are all based on the concept of hidden variables,²⁵ recently also applied to the fractional derivative example.²⁶ The augmented thermal field model, originally restricted to one-dimensional problems, is enlarged to encompass three-dimensional problems by an anelastic displacement model²¹ or by the augmented Hookes law. The generalized Maxwell model is based on his classical model.

All of the recently developed linear models are mathematically well founded, in contrast to the structural damping model with a frequency-independent loss factor. This model, as pointed out by Crandall,²⁷ exhibits noncausal behavior, which is physically unrealistic. Unfortunately, the previously mentioned models suffer from some limitations, mainly concerning the applicable frequency range. The minimum material parameter number required to successfully describe the polymer material relaxation or creep properties, over a broad structure-borne sound frequency range is small, only for the mini-oscillator and fractional derivative model. In addition, constitutive equations with fractional derivatives have sound bases in polymer visco-elasticity molecular theory.²⁸ The fractional derivative model is enlarged to complex derivatives,²⁹ and to fractional integrals.³⁰ The application of fractional derivative models in visco-elasticity is recently reviewed.^{31,32}

A nearly incompressible material model is presented, obeying decoupled spherical and deviatoric responses, being elastic in dilatation while displaying visco-elasticity in deviation. The latter exhibits a time strain separable relaxation tensor with a single function embodying its time dependence.

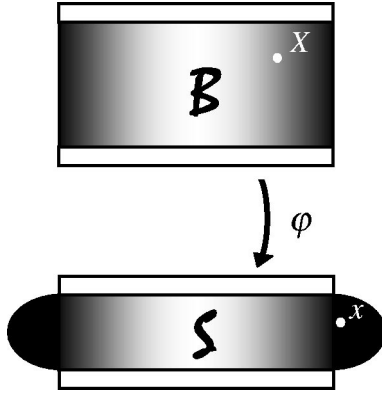


FIG. 1. A configuration φ of a vibration isolator from the reference configuration \mathcal{B} to the current in \mathcal{S} . Also shown are material and spatial points; X and x .

This function is based on an extended fractional order derivative model. The properties of this model, embodying a continuous distribution of fractional derivatives, are discussed. The dynamic axial stiffness within an audible frequency domain is solved by a nonlinear finite-element procedure, being particularly suitable to the proposed material model. The method is applied to a cylindrical vibration isolator with the results compared to those of measurements.

II. MODEL

A. Notations

Tensors are denoted by boldface and their components by lightface letters. Tr, T, dev, det, and div abbreviate trace, transpose, (symmetrized) deviation, determinant, and divergence. The operators \Re and \Im denote real and imaginary part, respectively. Moreover, the operators ∇ and \cdot denote (spatial) covariant derivative and double contraction.³³ The tensor products \otimes , \boxtimes , and \boxdot are defined through $(a \otimes b):c = (b:c)a$, $(a \boxtimes b):c = acb^T$, and $(a \boxdot b):c = ac^T b^T$, where a , b , and c are arbitrary second-order tensors. Finally, \mathbb{R} , \mathbb{Z} , and \mathbb{N} denote sets of real, integer, and natural numbers.

B. Kinematic preliminaries

Consider the vibration isolator in Fig. 1, a simple body consisting of continuously distributed material occupying a reference configuration \mathcal{B} in its natural state; stress-free and undeformed, and refer to \mathcal{S} as an ambient space in which the evolution of the body takes place, where \mathcal{B} and $\mathcal{S} \subset \mathbb{R}^3$ are sufficiently smooth, oriented open manifolds endowed with metrics \mathbf{I} and \mathbf{i} , respectively. A configuration space \mathcal{C} is a set of all admissible configurations; $\varphi: \mathcal{B} \rightarrow \mathcal{S}$. The isolator undergoes a regular motion; a curve in $\mathcal{C}: t \in \mathbb{R} \rightarrow \varphi_t \in \mathcal{C}$ and $x = \varphi_t(X) = \varphi(X, t)$, where $\varphi_t^{-1}: \mathcal{S} \rightarrow \mathcal{B}$ exists, $X \in \mathcal{B}$ and $x \in \mathcal{S}$ are material and spatial points. Therefore, ripping, cracking, breaking, pinching, or interpenetration of material is not considered while still allowing for self-contact under large preloads.

The Finger deformation tensor is $\mathbf{b} = \mathbf{F}\mathbf{F}^T$, where \mathbf{F} is the usual deformation gradient to φ . The deformation gradient is decomposable into an isochoric deformation gradient $\bar{\mathbf{F}} = J^{-1/3}\mathbf{F}$ and a volume strain measure

$$J = \sqrt{\frac{\det(\mathbf{i})}{\det(\mathbf{I})}} \frac{\partial(\varphi^1, \varphi^2, \varphi^3)}{\partial(X^1, X^2, X^3)}, \quad (1)$$

being the Jacobian of φ . The isochoric Finger deformation tensor is defined as $\bar{\mathbf{b}} = \bar{\mathbf{F}}\bar{\mathbf{F}}^T$.

By introducing a relative deformation gradient

$$\mathbf{f}_s(t) = \mathbf{F}(t)\mathbf{F}^{-1}(s), \quad (2)$$

the relative Green deformation tensor is given by

$$\mathbf{c}_s = \mathbf{f}_s^T \mathbf{f}_s. \quad (3)$$

In particular, $\mathbf{c}_s|_{t=s} = \mathbf{f}_s|_{t=s} = \mathbf{i}$. Likewise, the isochoric relative Green deformation tensor is given by

$$\bar{\mathbf{c}}_s = \bar{\mathbf{f}}_s^T \bar{\mathbf{f}}_s, \quad (4)$$

where the corresponding deformation gradient $\bar{\mathbf{f}}_s = j_s^{-1/3}\mathbf{f}_s$ and j_s is the associated Jacobian.

The principal invariants of \mathbf{b} read

$$I_1 = \text{tr } \mathbf{b}, \quad (5)$$

$$I_2 = \det \mathbf{b} \text{tr}(\mathbf{b}^{-1}), \quad (6)$$

and

$$I_3 = \det \mathbf{b} = J^2, \quad (7)$$

while those of $\bar{\mathbf{b}}$ read

$$\bar{I}_1 = \text{tr } \bar{\mathbf{b}} = J^{-2/3}I_1, \quad (8)$$

$$\bar{I}_2 = \text{tr}(\bar{\mathbf{b}}^{-1}) = J^{-4/3}I_2, \quad (9)$$

and

$$I_3 = \det \bar{\mathbf{b}} = 1, \quad (10)$$

where $\text{tr}(\cdot) = \mathbf{i}:(\cdot)$.

C. Constitutive assumptions

The rubber material is assumed to be isotropic, homogeneous, nearly incompressible, and time invariant, while being confined to isothermal rubber region conditions. In particular, the analysis excludes low-temperature crystallization and glass hardening, aging and dynamic amplitude dependence. The latter is motivated by the vanishing low dynamic strains commonly encountered in structure-borne sound,³⁴ while primarily focusing on unfilled and lightly filled carbon black rubber applications. The rubber is, in addition, modeled as a simple material³⁵ with delayed hyperelastic stress responses, embodying uncoupled spherical and deviatoric stresses; being elastic in dilatation while displaying viscoelasticity in deviation. Although subsequently confined to constitutive rubber relations, isolator metal parts are, nevertheless, readily joined by linear visco-elastic, elastic, or rigid body models.

The constitutive relation for the Kirchhoff stress $\boldsymbol{\tau}$, related to the Cauchy stress $\boldsymbol{\sigma}$ as $\boldsymbol{\tau} = J\boldsymbol{\sigma}$, is given by

$$\boldsymbol{\tau}(t) = s(J, \bar{\mathbf{b}}) + \int_{r=0}^{\infty} \mathbf{s}(\bar{\mathbf{c}}_t(t-r) - \mathbf{i}; \bar{\mathbf{b}}(t)), \quad (11)$$

where \mathbf{s} is a short-term response functional with $\mathbf{i}:\mathbf{s}=0$, $\mathbf{s}_{r=0}^\infty(\mathbf{0};\bar{\mathbf{b}})=\mathbf{0}$, and s an equilibrium stress response tensor. The latter response tensor is identified as the delayed hyperelastic stress response $\tau_\infty=s$, derivable from a spatial equilibrium Helmholtz' free energy $\psi=\psi(\mathbf{b})$ per unit undeformed volume³⁶ as

$$\tau_\infty = 2 \frac{\partial \psi}{\partial \mathbf{b}} \mathbf{b}, \quad (12)$$

which here, after variable change, equals

$$\tau_\infty = J \frac{\partial \psi}{\partial J} \mathbf{i} + 2 \frac{\partial \psi}{\partial \bar{I}_1} \text{dev}(\bar{\mathbf{b}}) - 2 \frac{\partial \psi}{\partial \bar{I}_2} \text{dev}(\bar{\mathbf{b}}^{-1}), \quad (13)$$

where $\text{dev}(\cdot) = (1 - \mathbf{i} \otimes \mathbf{i}/3) : (\cdot)$ and the spatial fourth-order unit tensor $\mathbb{1} = (\mathbf{i} \otimes \mathbf{i} + \mathbf{i} \otimes \mathbf{i})/2$.

The free-energy function ψ is assumed to be additively decomposable into

$$\psi(J, \bar{I}_1, \bar{I}_2) = \frac{\kappa_\infty}{2} (J-1)^2 + \sum_{n \in \mathbb{N}} \sum_{m \in \mathbb{N}} C_{nm} (\bar{I}_1 - 3)^n (\bar{I}_2 - 3)^m, \quad (14)$$

the first term giving the spherical part while the double sum the deviatoric part of the Kirchhoff stress, where $C_{nm} \in \mathbb{R}$, $C_{00}=0$; $\mu_\infty = 2(C_{10} + C_{01})$ and κ_∞ are the classical equilibrium shear and bulk moduli, respectively. The double sum is identified as the usual Rivlin type of exactly incompressible material representation, endowed, however, with the formal replacements; $I_1 \leftarrow \bar{I}_1$ and $I_2 \leftarrow \bar{I}_2$. In particular, the equilibrium stress is reduced to

$$\tau_\infty = \kappa_\infty J(J-1) \mathbf{i} + (\mu_\infty + 4C_{20}(\bar{I}_1 - 3) + 6C_{30}(\bar{I}_1 - 3)^2) \text{dev}(\bar{\mathbf{b}}), \quad (15)$$

for Yeoh materials,³⁷ reported to agree with carbon black filled rubber behavior.

To linearize the constitutive relation at a given prestrain, the configuration may additively be decomposed into $\varphi_\epsilon = \hat{\varphi} + \epsilon \Delta \varphi + \mathcal{O}(\epsilon^2)$, where $\epsilon \Delta \varphi: \mathcal{S} \rightarrow \mathcal{S}$ is an infinitesimal configuration—a point mapping with an associable infinitesimal displacement $\epsilon \Delta \mathbf{u}$ —superimposed on a finite configuration $\hat{\varphi}$; $\epsilon \in \mathbb{R}$, $|\epsilon| \ll 1$. The infinitesimal stress $\epsilon \Delta \boldsymbol{\tau}$ due to this superimposed infinitesimal deformation reads

$$\begin{aligned} \Delta \boldsymbol{\tau} = & \kappa_\infty \dot{J}(2\dot{J} - 1) \text{tr}(\Delta \boldsymbol{\epsilon}) \mathbf{i} + \Delta \mathbf{w} \text{dev}(\dot{\boldsymbol{\tau}}_\infty) - \text{dev}(\dot{\boldsymbol{\tau}}_\infty) \Delta \mathbf{w} \\ & + \mathbb{k} : \text{dev}(\Delta \boldsymbol{\epsilon}) + 2 \int_{-\infty}^t \dot{\mathbb{C}}(t-s) : \frac{\partial \text{dev}(\Delta \boldsymbol{\epsilon}(s))}{\partial s} ds, \end{aligned} \quad (16)$$

with error in $\mathcal{O}(\epsilon^2)$, where the infinitesimal strain $\epsilon \Delta \boldsymbol{\epsilon}$ and rotation $\epsilon \Delta \mathbf{w}$ are $\Delta \boldsymbol{\epsilon} = (\nabla \Delta \mathbf{u} + (\nabla \Delta \mathbf{u})^T)/2$ and $\Delta \mathbf{w} = (\nabla \Delta \mathbf{u} - (\nabla \Delta \mathbf{u})^T)/2$; $\dot{\mathbb{C}}$ is an isotropic fourth-order relaxation tensor at $\hat{\varphi}$ while \mathbb{k} is a deviatoric Jaumann elasticity tensor at $\hat{\varphi}$. Concisely, this elasticity reads⁷

$$\begin{aligned} \mathbb{k} = & 2 \frac{\partial \psi}{\partial \bar{I}_1} \left(\dot{\bar{\mathbf{b}}} \otimes \mathbf{i} - \frac{2}{3} \dot{\bar{\mathbf{b}}} \otimes \mathbf{i} + \mathbf{i} \otimes \dot{\bar{\mathbf{b}}} - \frac{2}{3} \mathbf{i} \otimes \dot{\bar{\mathbf{b}}} \right) \\ & + 2 \frac{\partial \psi}{\partial \bar{I}_2} \left(\dot{\bar{\mathbf{b}}}^{-1} \otimes \mathbf{i} - \frac{2}{3} \dot{\bar{\mathbf{b}}}^{-1} \otimes \mathbf{i} + \mathbf{i} \otimes \dot{\bar{\mathbf{b}}}^{-1} - \frac{2}{3} \mathbf{i} \otimes \dot{\bar{\mathbf{b}}}^{-1} \right) \\ & + 4 \left(\frac{\partial^2 \psi}{\partial \bar{I}_1^2} \text{dev}(\dot{\bar{\mathbf{b}}}) \otimes \text{dev}(\dot{\bar{\mathbf{b}}}) - \frac{\partial^2 \psi}{\partial \bar{I}_1 \partial \bar{I}_2} \text{dev}(\dot{\bar{\mathbf{b}}}^{-1}) \otimes \text{dev}(\dot{\bar{\mathbf{b}}}) \right. \\ & \left. - \frac{\partial^2 \psi}{\partial \bar{I}_1 \partial \bar{I}_2} \text{dev}(\dot{\bar{\mathbf{b}}}) \otimes \text{dev}(\dot{\bar{\mathbf{b}}}^{-1}) \right. \\ & \left. + \frac{\partial^2 \psi}{\partial \bar{I}_2^2} \text{dev}(\dot{\bar{\mathbf{b}}}^{-1}) \otimes \text{dev}(\dot{\bar{\mathbf{b}}}^{-1}) \right), \end{aligned} \quad (17)$$

at $\varphi_n = \hat{\varphi}_n$, which is reduced to

$$\begin{aligned} \mathbb{k} = & (\mu_\infty + 4C_{20}(\bar{I}_1 - 3) + 6C_{30}(\bar{I}_1 - 3)^2) \\ & \times \left(\dot{\bar{\mathbf{b}}} \otimes \mathbf{i} - \frac{2}{3} \dot{\bar{\mathbf{b}}} \otimes \mathbf{i} + \mathbf{i} \otimes \dot{\bar{\mathbf{b}}} - \frac{2}{3} \mathbf{i} \otimes \dot{\bar{\mathbf{b}}} \right) \\ & + 8(C_{20} + 3C_{30}(\bar{I}_1 - 3)) \text{dev}(\dot{\bar{\mathbf{b}}}) \otimes \text{dev}(\dot{\bar{\mathbf{b}}}), \end{aligned} \quad (18)$$

for Yeoh materials. Assume further a time and prestrain separable relaxation tensor:³⁸ $\dot{\mathbb{C}}(t) = g(t) \mathbb{k}$, where g is a single relaxation function embodying the time dependence while \mathbb{k} its prestrain dependence. The relaxation function $g = (\mu/\mu_\infty - h)/2$, where μ is the familiar infinitesimal shear relaxation function and h the step function. Let the bulk modulus $\kappa = \kappa_\infty = a\mu_\infty$ where the constant $a \gg 1$, typically $\sim 10^2$ to 10^5 , in accordance with the near-incompressibility assumption. The incremental relation (16) becomes

$$\begin{aligned} \Delta \boldsymbol{\tau} = & a\mu_\infty \dot{J}(2\dot{J} - 1) \text{tr}(\Delta \boldsymbol{\epsilon}) \mathbf{i} + \Delta \mathbf{w} \text{dev}(\dot{\boldsymbol{\tau}}_\infty) - \text{dev}(\dot{\boldsymbol{\tau}}_\infty) \Delta \mathbf{w} \\ & + \mathbb{k} : \left(\text{dev}(\Delta \boldsymbol{\epsilon}) + 2 \int_{-\infty}^t g(t-s) \frac{\partial \text{dev}(\Delta \boldsymbol{\epsilon}(s))}{\partial s} ds \right). \end{aligned} \quad (19)$$

Clearly, the response of a prestrained, homogeneous, and isotropic visco-elastic material to a subsequent infinitesimal deformation, disagrees in general with that of a homogeneous, isotropic visco-elastic material, corresponding to the classical infinitesimal theory. The apparently anisotropic nonhomogeneous behavior of the incremental response (19) originates from the rubber structure's prestrain dependence: Rubber displays a preferred chain alignment in the extension direction, eventually inducing crystallinity at still higher stretches, whereas dilatation merely changes the amorphous' gross density—with all the effects showing, at nonuniform prestrain, a spatial point dependence; that is, apparently being nonhomogeneous.

Fourier transformation $(\sim) = \int_{-\infty}^{\infty} (\cdot) \exp(-i\omega t) dt$, of the constitutive relaxation relation (19), with i denoting the imaginary unit and ω angular frequency, results in

$$\Delta \tilde{\tau} = a \mu_{\infty} \dot{J} (2\dot{J} - 1) \text{tr}(\Delta \tilde{\epsilon}) \mathbf{i} + \Delta \tilde{\mathbf{w}} \text{dev}(\dot{\tilde{\tau}}_{\infty}) - \text{dev}(\dot{\tilde{\tau}}_{\infty}) \Delta \tilde{\mathbf{w}} + \frac{\hat{\mu}}{\mu_{\infty}} \mathbf{k} : \text{dev}(\Delta \tilde{\epsilon}), \quad (20)$$

where the shear modulus is

$$\hat{\mu} = (1 + 2i\omega \tilde{g}) \mu_{\infty}. \quad (21)$$

The corresponding infinitesimal constitutive relation

$$\Delta \tilde{\sigma} = a \mu_{\infty} \text{div}(\Delta \tilde{\mathbf{u}}) \mathbf{i} + 2 \hat{\mu} \text{dev}(\nabla \Delta \tilde{\mathbf{u}}), \quad (22)$$

is redeemed at vanishing prestrains; $\dot{\tilde{\tau}}_{\infty} \rightarrow \mathbf{0}$, $\dot{\tilde{I}}_1 = \dot{\tilde{I}}_2 \rightarrow 3$, $\dot{J} \rightarrow 1$, and $\mathbf{k} : \text{dev}(\Delta \tilde{\epsilon}) \rightarrow 2 \mu_{\infty} \text{dev}(\nabla \Delta \tilde{\mathbf{u}})$, where $\text{div}(\cdot) = \mathbf{i} : \nabla(\cdot)$. Not surprising, since every well-posed large prestrain constitutive theory reveals its legitimate small motion behavior.

A typical dynamic stiffness problem consists of an equation of motion with mixed boundary conditions; a configuration $\varphi_d \in \mathcal{C}$ given on the boundary $\partial \mathcal{B}_d$ and a dead-loading \mathbf{t} given on the boundary $\partial \mathcal{B}_t$, where $\partial \mathcal{B}_d \cap \partial \mathcal{B}_t = \emptyset$ and $\partial \mathcal{B}_d \cup \partial \mathcal{B}_t = \partial \mathcal{B}$. To obtain expressions convenient for numerical computations, the problem may be translated into a weak form suitable for a direct method, such as a Galerkin procedure.³⁹ In particular, the governed relations may be transformed into a weak two-field formulation of displacement and pressure character, suitable for an updated Lagrangian nonlinear finite-element procedure. This procedure is particularly suitable for the proposed material model, and is subsequently used.

D. Audible short term response

A classical linear relaxation function is the generalized Maxwell function

$$g(t) = \frac{1}{2} \sum_{n \in \mathbb{Z}_+} \Delta_n e^{-t/\tau_n} h(t), \quad (23)$$

where τ_n and Δ_n are individual relaxation time and intensity, respectively, resulting in a total relaxation intensity $\Delta = \sum_{n \in \mathbb{Z}_+} \Delta_n$ such that $\lim_{t \rightarrow 0} g(t) = \mu_{\infty} (1 + \Delta)$. The relaxation function is in fact identical to that of the augmented Hooke's law. However, the term number required to successfully describe the rubber relaxation properties over a broad structure-borne sound-frequency range is often large, resulting in a cumbersome number or a frequency range reduction. A major reduction with a preserved broad frequency range is obtained by the fractional derivative model.¹⁷ In addition, constitutive relations with fractional derivatives have sound bases in polymer visco-elasticity molecular theory.²⁸ The relaxation function of a fractional standard linear solid reads⁴⁰

$$g(t) = \frac{\Delta}{2} E_{\alpha} \left(-\Delta \left\{ \frac{\mu_{\infty} t}{\mu_v} \right\}^{\alpha} \right) h(t), \quad (24)$$

where the equilibrium shear modulus, $\mu_{\infty} = \lim_{t \rightarrow \infty} \mu(t) \geq 0$, the relaxation intensity $\Delta \geq 0$, the fraction derivative order $0 < \alpha \leq 1$ and $\mu_v \geq 0$ are material constants. The Mittag-Leffler function is

$$E_{\alpha}(x) = \sum_{n=0}^{\infty} \frac{x^n}{\Gamma(1+n\alpha)}, \quad (25)$$

where the gamma function $\Gamma(z) = \int_0^{\infty} s^{z-1} e^{-s} ds$, $z > 0$. This rubber material model results in an admissible behavior with a finite instantaneous relaxation function $\lim_{t \rightarrow 0} g(t) = \Delta/2$. The model embodies a rubber, a transition, and a glassy region, as does its ordinary counterpart, known as the standard linear solid, with the relaxation function $g(t) = \Delta \exp(-t \Delta \mu_{\infty} / \mu_v) h(t)/2$ —being its weak limit as $\alpha \rightarrow 1$. Restricted to frequencies within the rubber region, the relaxation intensity is undetermined and may be set to $\Delta \rightarrow \infty$; thus, reducing to a fractional Kelvin–Voigt example with a relaxation function

$$g(t) = \frac{1}{2} \left(\frac{\mu_v}{\mu_{\infty}} \right)^{\alpha} I_{\alpha}(t), \quad (26)$$

where the Abel operator kernel

$$I_{\alpha}(t) = \frac{h(t)}{t^{\alpha} \Gamma(1-\alpha)}. \quad (27)$$

This function may be extended into a generalized fractional Kelvin–Voigt example with a relaxation function

$$g(t) = \frac{1}{2} \sum_{n \in \mathbb{Z}_+} \left(\frac{\mu_{vn}}{\mu_{\infty}} \right)^{\alpha_n} I_{\alpha_n}(t), \quad (28)$$

where the equilibrium shear modulus $\mu_{\infty} \geq 0$, the fraction derivative order $0 < \alpha_n \leq 1$ and $\mu_{vn} \geq 0$ are material constants, with $n \in \mathbb{Z}_+$. In addition, continuous fractional-order derivatives are possible, such as if the terms of (28) are so numerous that the summation over n is replaced by an integration over α , or if the experimentally determined exponents are so uncertain that the material constants are better, replaced by piecewise continuous functions of α . Caputo^{41,42} and Caputo and Mainardi⁴³ propose continuous fractional-order derivatives in visco-elastic relations; also discussed by Koeller,⁴⁰ while being generalized to a general input–output model by Wainwright.⁴⁴ To this end, consider the continuous fractional-order Kelvin–Voigt relaxation function

$$g(t) = \frac{1}{2} \int_{0+}^1 w(\alpha) \left(\frac{\mu_v}{\mu_{\infty}} \right)^{\alpha} I_{\alpha}(t) d\alpha, \quad (29)$$

where w is a weighting function, normalized as $\int_{0+}^1 w(\alpha) d\alpha = 1$. Applying this into the incremental relation (19), explicitly using fractional derivatives, yields

$$\Delta \tau = a \mu_{\infty} \dot{J} (2\dot{J} - 1) \text{tr}(\Delta \epsilon) \mathbf{i} + \Delta \mathbf{w} \text{dev}(\dot{\tau}_{\infty}) - \text{dev}(\dot{\tau}_{\infty}) \Delta \mathbf{w} + \mathbf{k} : \left(\text{dev}(\Delta \epsilon) + \int_{0+}^1 w(\alpha) \times \left(\frac{\mu_v}{\mu_{\infty}} \right)^{\alpha} {}_{-\infty} D_t^{\alpha} \{ \text{dev}(\Delta \epsilon) \} d\alpha \right), \quad (30)$$

where the fractional derivative of order $\beta \geq 0$ is defined through analytical continuation of a Riemann–Liouville integral⁴⁵ as

$${}_c D_t^\beta f(t) = \frac{1}{\Gamma(1-\beta + [\beta])} \frac{d^{[\beta]+1}}{dt^{[\beta]+1}} \int_c^\infty \frac{f(s)}{(t-s)^{\beta-[\beta]}} ds, \quad (31)$$

where f is a function of t , $[\cdot]$ is integer part of (\cdot) ; c and t are integration terminals with the specializations $c=0$ and $c=-\infty$, giving the Riemann and Liouville definitions, respectively. The fractional derivative operator is nonlocal in temporal sense and exhibits local characteristic only for $\beta \in \mathbb{N}$; thus, producing the same result for $\beta \in \mathbb{N}$ as an ordinary differentiation. Other definitions, such as a distributional fractional derivative on convolution form are also possible.⁴⁶

A typical polymer fractional derivative lies in an interval from α_1 to α_2 . Torvik and Bagley⁴⁷ find that $\alpha_1=1/2$ and $\alpha_2=2/3$ by performing broad frequency band experiments on approximately 130 high damping materials, while Curro *et al.*⁴⁸ report $\alpha_1=0.1$ and $\alpha_2=0.3$ from numerous long-term relaxation and retardation experiments on elastomers, such as natural rubber. In this connection, consider the weighting function being interpretable as a probability analysis density function. Inasmuch as a relaxation molecular process is of stochastic nature, a distribution of fractional derivatives rather than a single one might be an appropriate macroscopic property model. A uniform derivative distribution in α_1 to α_2 , $\Delta\alpha = \alpha_2 - \alpha_1$; $w(\alpha) = \{h(\alpha - \alpha_1) - h(\alpha - \alpha_2)\} / \Delta\alpha$, is a suitable weighting function, with a mean value $\langle\alpha\rangle = (\alpha_2 + \alpha_1)/2$, tending to $w(\alpha) = \delta(\alpha - \langle\alpha\rangle)$ as $\Delta\alpha \rightarrow 0$, where δ is a Dirac function.

Temporal Fourier transformation of the constitutive relaxation relation (30) having a uniform derivative distribution, with

$$\begin{aligned} \int_{-\infty}^{\infty} {}_{-\infty} D_t^\alpha f(t) e^{-i\omega t} dt &= (i\omega + 0^+)^\alpha \tilde{f} \\ &\stackrel{\text{def}}{=} |\omega|^\alpha e^{i\text{sign}(\omega)(\pi\alpha/2)} \tilde{f} = (i\omega)^\alpha \tilde{f}, \end{aligned} \quad (32)$$

results in

$$\begin{aligned} \Delta \tilde{\tau} &= a \mu_\infty \dot{J} (2\dot{J} - 1) \text{tr}(\Delta \tilde{\epsilon}) \mathbf{i} + \Delta \tilde{\mathbf{w}} \text{dev}(\dot{\tilde{\tau}}_\infty) - \text{dev}(\dot{\tilde{\tau}}_\infty) \Delta \tilde{\mathbf{w}} \\ &+ \left\{ 1 + \frac{\left(\frac{\mu_\nu}{\mu_\infty} i\omega\right)^{\alpha_2} - \left(\frac{\mu_\nu}{\mu_\infty} i\omega\right)^{\alpha_1}}{\Delta\alpha \log_e\left(\frac{\mu_\nu}{\mu_\infty} i\omega\right)} \right\} \mathbf{k} : \text{dev}(\Delta \tilde{\epsilon}). \end{aligned} \quad (33)$$

The sought constitutive relaxation relation (33) embodies, as special cases, both the fractional Kelvin–Voigt model

$$\begin{aligned} \Delta \tilde{\tau} &= a \mu_\infty \dot{J} (2\dot{J} - 1) \text{tr}(\Delta \tilde{\epsilon}) \mathbf{i} + \Delta \tilde{\mathbf{w}} \text{dev}(\dot{\tilde{\tau}}_\infty) - \text{dev}(\dot{\tilde{\tau}}_\infty) \Delta \tilde{\mathbf{w}} \\ &+ \left\{ 1 + \left(\frac{\mu_\nu}{\mu_\infty} i\omega\right)^\alpha \right\} \mathbf{k} : \text{dev}(\Delta \tilde{\epsilon}), \end{aligned} \quad (34)$$

as $\Delta\alpha \rightarrow 0$, and the ordinary Kelvin–Voigt model

$$\begin{aligned} \Delta \tilde{\tau} &= a \mu_\infty \dot{J} (2\dot{J} - 1) \text{tr}(\Delta \tilde{\epsilon}) \mathbf{i} + \Delta \tilde{\mathbf{w}} \text{dev}(\dot{\tilde{\tau}}_\infty) - \text{dev}(\dot{\tilde{\tau}}_\infty) \Delta \tilde{\mathbf{w}} \\ &+ \left(1 + \frac{\mu_\nu}{\mu_\infty} i\omega \right) \mathbf{k} : \text{dev}(\Delta \tilde{\epsilon}), \end{aligned} \quad (35)$$

as $\alpha_1 = \alpha_2 = 1$, thus giving a richer structure to the model. In addition, the generalized fractional Kelvin–Voigt model

$$\begin{aligned} \Delta \tilde{\tau} &= a \mu_\infty \dot{J} (2\dot{J} - 1) \text{tr}(\Delta \tilde{\epsilon}) \mathbf{i} + \Delta \tilde{\mathbf{w}} \text{dev}(\dot{\tilde{\tau}}_\infty) - \text{dev}(\dot{\tilde{\tau}}_\infty) \Delta \tilde{\mathbf{w}} \\ &+ \left\{ 1 + \sum_{n \in \mathbb{Z}_+} \left(\frac{\mu_{\nu_n}}{\mu_\infty} i\omega \right)^{\alpha_n} \right\} \mathbf{k} : \text{dev}(\Delta \tilde{\epsilon}), \end{aligned} \quad (36)$$

is attained using $w(\alpha) = \sum_{n \in \mathbb{Z}_+} \delta(\alpha - \alpha_n)$.

Traditionally, the short-term properties are specified through the relaxation functions or complex moduli; but, as the classic relaxation functions commonly involve exponentially decaying functions, yet another property form exists: The relaxation spectrum γ_r given by

$$g(t) = \int_{0^-}^{\infty} \gamma_r(\tau) e^{-t/\tau} d\{r(\tau)\} h(t), \quad (37)$$

where r is a continuous and monotonically increasing function of the relaxation time τ . Functions widely employed are $r = \text{identity}$, \log_e , or \log_{10} ; with an optional and identity function spectra inter-relation $\gamma_r = \gamma_{\text{identity}} d\tau/dr$. Spectra have a wide application area: In studies of visco-elastic properties, crack growth and inter-relation between the macroscopic properties and molecular modes, being interpretable as the characteristic molecular time response range, which is extremely broad, thus making a logarithmic r suitable.

The classical spectra are discrete: the generalized Maxwell (23) predicts

$$\gamma_r(\tau) = \frac{1}{2} \sum_{n \in \mathbb{Z}_+} \Delta_n \delta(\tau - \tau_n) \frac{d\tau}{dr}, \quad (38)$$

while the standard linear solid

$$\gamma_r(\tau) = \frac{1}{2} \delta\left(\tau - \frac{\mu_\nu}{\Delta\mu_\infty}\right) \frac{d\tau}{dr}, \quad (39)$$

being consistent with certain molecular theories.⁴⁹ However, various experimental results suggest a $t^{-b}h(t)$ proportional stress relaxation⁵⁰ instead of an exponential, where b is between 0 and 1, yielding a continuous spectrum. Indeed, the fractional (26), generalized fractional (28), and uniformly continuous fractional (29) Kelvin–Voigt examples embody an Abel operator kernel relaxation function $\propto t^{-\alpha}h(t)$, with continuous spectra

$$\gamma_r(\tau) = \frac{1}{2\pi\tau} \left(\frac{\mu_\nu}{\mu_\infty}\right)^\alpha \frac{\sin(\pi\alpha)}{\tau^\alpha} \frac{d\tau}{dr}, \quad (40)$$

$$\gamma_r(\tau) = \frac{1}{2\pi\tau} \sum_{n \in \mathbb{Z}_+} \left(\frac{\mu_{\nu_n}}{\mu_\infty}\right)^{\alpha_n} \frac{\sin(\pi\alpha_n)}{\tau^{\alpha_n}} \frac{d\tau}{dr}, \quad (41)$$

and

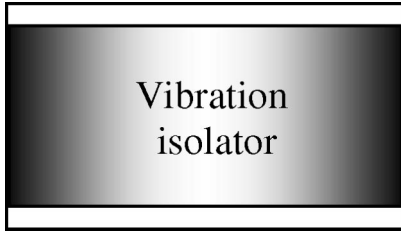


FIG. 2. Test object.

$$\begin{aligned} \gamma_r(\tau) = & \left(2\pi\tau\Delta\alpha \left\{ \log_e^2 \left(\frac{\mu_\nu}{\mu_\infty\tau} \right) + \pi^2 \right\} \right)^{-1} \left(\log_e \left(\frac{\mu_\nu}{\mu_\infty\tau} \right) \right. \\ & \times \left\{ \left(\frac{\mu_\nu}{\mu_\infty\tau} \right)^{\alpha_2} \sin(\pi\alpha_2) - \left(\frac{\mu_\nu}{\mu_\infty\tau} \right)^{\alpha_1} \sin(\pi\alpha_1) \right\} \\ & - \pi \left\{ \left(\frac{\mu_\nu}{\mu_\infty\tau} \right)^{\alpha_2} \cos(\pi\alpha_2) \right. \\ & \left. \left. - \left(\frac{\mu_\nu}{\mu_\infty\tau} \right)^{\alpha_1} \cos(\pi\alpha_1) \right\} \right) \frac{d\tau}{dr}, \end{aligned} \quad (42)$$

respectively, obtained from relaxation function g as⁵¹

$$\gamma_r(\tau) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi\tau} \mathcal{J} \left\{ \left(-\frac{1}{\tau} + i\epsilon \right) \check{g} \left(-\frac{1}{\tau} + i\epsilon \right) \right\}, \quad (43)$$

where $\check{g}(i\omega) \equiv \check{g}(\omega)$. Although the spectra are singular at vanishing time, as is their ordinary counterpart, the stress, nevertheless depends on the entire strain history (nonvanishing spectra at positive time), thus, conforming to a material with memory. Furthermore, a whole discrete damping mechanism spectra can be included in a single continuous relaxation time distribution, minimizing required material parameter number, where the continuous fractional Kelvin–Voigt example, as compared to the discrete derivative example, displays an enhanced versatility to agree with measurement.

III. RESULT AND DISCUSSION

To examine the dynamic stiffness model in the audible frequency domain, a real vibration isolator has been analyzed, presenting numerical as well as measurement results. The governed equations are solved by a general finite-element code ANIA,⁵² LINUX[®] 1997 version implemented on a PC—Pentium Pro[®]; developed for research, written in FORTRAN 77 with all calculations performed in double precision. Graphically, the results are presented by means of MATLAB[®]. Similar codes are commercially available as ABAQUS[®] and MARC[®].

Albeit subsequently restricted to the axial stiffness of a specific isolator geometry and material, whereas the model is not, the study clearly illuminates the nature of finite prestrains, particularly in audible dynamic stiffness applications, where thoroughly linear theories are inferior to the above presented.

A. Test object

The isolator in Fig. 2 is a compression molded solid cylinder of length 50.0 mm and radius 50.0 mm, cured to

TABLE I. Principal ingredients of the natural rubber compound.

Ingredient	Type	Concentration (phr)
Natural rubber	SMR CV50	100
Carbon black	N772	10
Antidegradant system	Antiozonant	1
	Antioxidant	1
	Wax	1
Activators	Stearic acid	1
	Zinc oxide	5
Vulcanizing agent	Sulfur	2
Accelerator	CBS	3
Processing oils	Paraffinic	1
	Aromatic	5

two circular steel plates of thickness 2.6 mm and radius 50.0 mm. Safe mounting is facilitated with additional plates of thickness 19.0 mm and radius 50.0 mm, having threaded center holes. Total plate thickness is 21.6 mm. The principal ingredients of the vulcanized natural rubber are given in Table I. The ground-state density is 1050 kg/m³, nominal hardness 40° IRH, and maximum long-term axial static preload is limited to 3000 N.

B. Measurement of static, dynamic, and material properties

The tests are split into two parts: a static procedure, measuring the time-delayed force-deflection characteristic of the isolator with a corresponding dynamic test of its frequency-dependent stiffness. A suitable description of the dynamic properties of the isolator provides the (blocked dynamic axial) transfer and driving point stiffness. The former is defined as the axial force (on the blocked side of the isolator) over the excitation displacement (in the frequency domain) while the latter is the excitation force over the same displacement. Positive force and displacement direction is defined as the outward normal to the isolator plate surface.

The tests are carried out at room temperature, (20.8 ± 0.3) °C. Prior to the dynamic and static testing, the isolator is preconditioned: The test object is subjected to a few deformation cycles at a slow rate in the axial direction. The extreme values slightly exceed the maximum static preload. Tests are started after about 30 min to allow for stress relaxation.

1. Static measurement method

The static properties are measured by a servo hydraulic displacement-controlled 100 kN Instron 8032 uniaxial tester, shown in Fig. 3. The test object is mounted between a 20-kN strain gauge force transducer and a servo hydraulic piston, with static deflection measured by a built-in linear variable differential transformer displacement transducer. The output signal from the displacement transducer is measured by a Fluke model 27 multimeter, while a digital HP 34401A multimeter displays the amplified force signal. The measurement is directed by an Instron 2160 Series controller using a 100-kN Instron 2518-111 force gauge in addition to the displacement transducer. To determine the delayed elastic force, the test object is rested at each displacement increment till the force signal is approximately a constant (some minute).

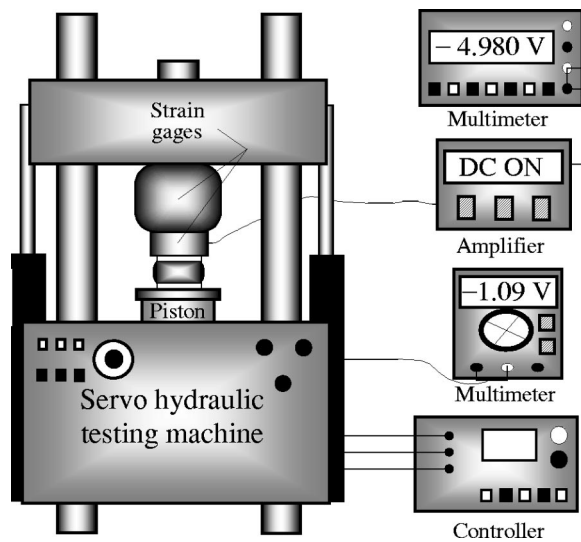


FIG. 3. Static measurement setup.

Prior to the experiment, the displacement transducer is calibrated using a Metron 10-mm dial indicator. The instruments needed are tabulated in Table II.

2. Dynamic measurement method

An indirect method is utilized in Fig. 4(a), with the test object mounted between a block and the moving table of an electro-dynamic vibration generator membrane. The motion is measured by six piezoelectric accelerometers: two symmetrically positioned, isolator close on the moving table and four symmetrically positioned, isolator close on the blocking mass. Newton's second law provides an estimate of the blocking force needed for the transfer stiffness estimation; namely, the force equals the mass of the blocking mass multiplied by its acceleration. The acceleration signals are conditioned in charge amplifiers without explicit time integration in order to increase the high-frequency accuracy. A four-channel frequency analyzer collects the data and supplies the signal to the generator via an amplifier. The measurements are processed by a personal computer. The instruments required are given in Table III. The applied dynamic measurement method is similar to that of Ref. 53, to which a reader may refer for details, such as error analysis.

For sufficiently high frequencies $f \gg f_0$, Newton's second law provides an estimate of the blocking force $\tilde{F}_b \approx \omega^2 M_b \tilde{U}_b$ needed for the transfer stiffness estimate

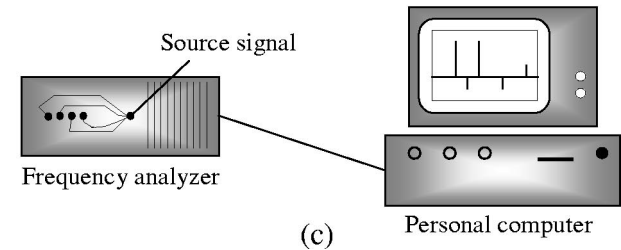
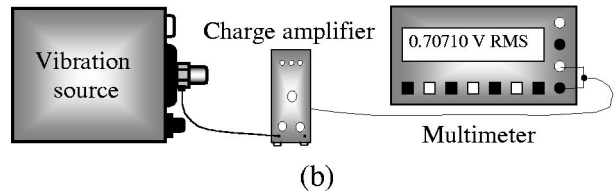
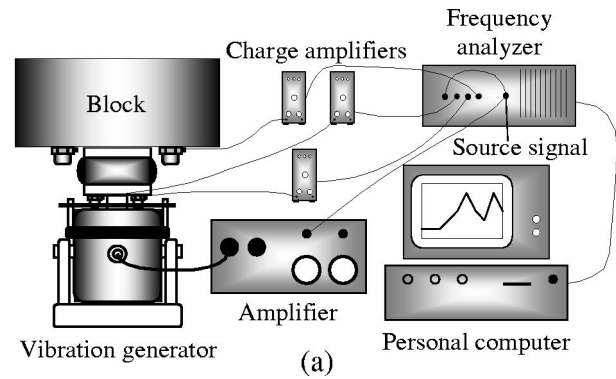


FIG. 4. (a) Dynamic measurement setup; (b) sensitivity calibration; and (c) cross channel calibration.

$$\tilde{k}_{12} \approx \omega^2 M_b \frac{\tilde{U}_b}{\tilde{U}_{mt}}, \quad (44)$$

where M_b is mass of blocking mass, \tilde{U}_b and \tilde{U}_{mt} are axial displacements of blocking mass and moving table, respectively, and where f_0 is rigid body resonance frequency for a mass-spring system consisting of blocking mass and the studied vibration isolator. Let a frequency response function

$$\tilde{H}(\tilde{X}, \tilde{Y}) = \frac{\tilde{G}(\tilde{X}, \tilde{Y})}{\tilde{G}(\tilde{X}, \tilde{X})}, \quad (45)$$

where $\tilde{G}(\tilde{X}, \tilde{Y})$ is the cross-spectrum density, defined as the temporal Fourier transform of a cross-correlation function

TABLE II. Static measurement instruments in Fig. 3.

Instrument	Type
Servo hydraulic testing machine	Instron 8032
Controller	Instron 2160 series
Strain gage	Instron 2518-111
Strain gage	Load indicator AB-20-KA5
Amplifier	HBM MVD2555
Multimeter	HP 34401A
Multimeter	Fluke model 27
Dial indicator	Metron 10 mm

TABLE III. Dynamic measurement instruments in Fig. 4.

Instrument	Type	Number
Vibration generator	LDS V409	1
Amplifier	Labgruppen SS1400	1
Accelerometer	B&K 4393	2
Accelerometer	B&K 4367	3
Accelerometer	Rion PV-84	2
Charge amplifier	B&K 2635	3
Frequency analyzer	Tektronix 2630	1
Computer	Aquila PC486	1

$$G(X, Y) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} X(\tau) Y(\tau + t) d\tau. \quad (46)$$

To reduce the influence from measurement noise, source correlation technique applies, with the output signal S from the frequency analyzer as a source signal. As a result, the stiffness (44) reads

$$\tilde{k}_{12} \approx \frac{\omega^2 M_b}{2} \times \frac{\tilde{H}(\tilde{S}, \tilde{A}_{b1}) + \tilde{H}(\tilde{S}, \tilde{A}_{b2}) + \tilde{H}(\tilde{S}, \tilde{A}_{b3}) + \tilde{H}(\tilde{S}, \tilde{A}_{b4})}{\tilde{H}(\tilde{S}, \tilde{A}_{m1}) + \tilde{H}(\tilde{S}, \tilde{A}_{m2})}, \quad (47)$$

where the axial accelerations \tilde{A}_{b1} , \tilde{A}_{b2} , \tilde{A}_{b3} , and \tilde{A}_{b4} are measured at the blocking mass and \tilde{A}_{m1} and \tilde{A}_{m2} at the moving table.

Two different blocking masses are utilized; $M_b = 214$ and 340 kg also being the isolator preloads. For comparison, stiffness results are included for the same test object but without preload using similar test arrangements from Ref. 54, to which the reader may refer for details.

The frequency domain is 50 to 1000 Hz, using a stepped sine excitation signal, initiating at 50 Hz, with a constant frequency step of 2 Hz. The signal is recorded within a 5-Hz bandwidth, averaged 30 times, and delayed 1000 ms between each recording.

Each measurement chain—embodying an accelerometer, an accelerometer coaxial cable, a charge amplifier, and an RG58 coaxial cable—is sensitivity calibrated in Fig. 4(b), with the accelerometer mounted on the shaker table of a calibrated vibration source, B&K 4291. By tuning into the actual accelerometer mass, the built-in generator is adjusted to give a sinusoidal signal with a peak value of 10.0 m/s² at 79.6 Hz. The output signal from the charge amplifier is measured by a digital HP 34401A multimeter. For transfer stiffness measurements it is sufficient to measure the ratio of the block displacement to the moving table displacement; then, it is only necessary to determine the relative sensitivity between the measurement chains. Therefore, the calibration discounts any systematic deviation from the stated vibration level of the calibrated vibration source.

Errors due to systematic phase and amplitude mismatch across the channels on the frequency analyzer are reduced in Fig. 4(c) by measuring, and gathering, the frequency response functions between them. The output signal from the frequency analyzer is used as a source signal. The frequency range, the frequency step, and the time delay are the same as those for stiffness measurements. The number of averages is, however, increased to 50 to minimize random error influence.

3. Measurement results

The measured axial static force due to axial deformation in the domain 0 to −12 mm with steps of approximately −1 mm is shown in Fig. 5(a). The extreme values correspond to a global axial strain (=axial plate displacement over rubber cylinder length) of about −24% and a compression force

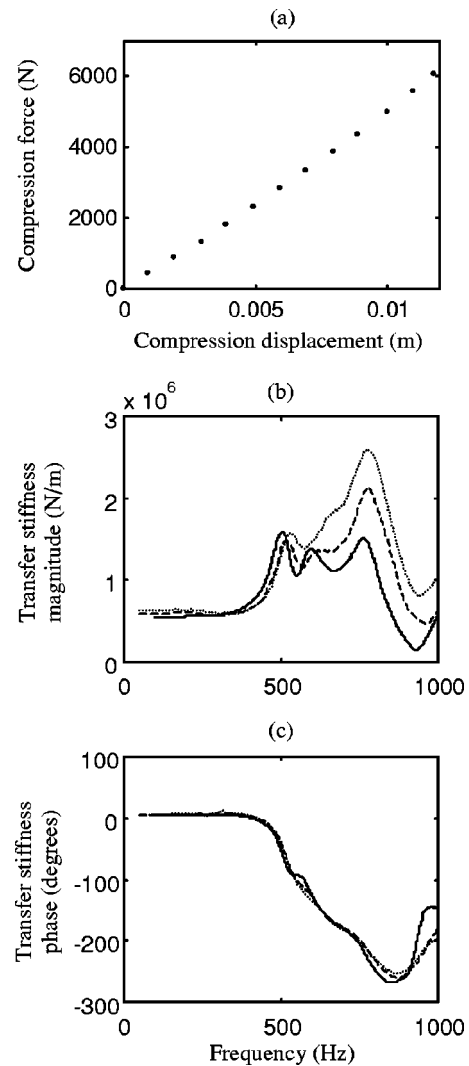


FIG. 5. Measurement results. (a) Axial compression force versus displacement; (b) magnitude and (c) unwrapped phase of transfer stiffness. 0 N (solid thin), 2100 N (dashed), and 3350 N (dotted).

surpassing 6000 N which amply covers the allowed maximum long-term axial static preload. Nevertheless, the intuitively expected compression stiffening is already evident at moderate deformations with pervading emphasis at larger compressions, not predictable in thoroughly linear theories nor accounting for material and geometric effects associated with finite strains.

The measured axial dynamic transfer stiffness in the whole frequency domain 50 to 1000 Hz is shown in Figs. 5(b) and (c) from the vanishing preload up to 3350 N. Figure 5(b) shows the magnitude with the unwrapped phase in Fig. 5(c).

Confined to a vanishing preload: The stiffness magnitude curve in Fig. 5(b) is plateau-like in the low-frequency region. Then, the curve rises steeply to a first peak at 500 Hz, drops to a trough, rises to a second peak, drops and rises to a third peak, and, finally, drops into a third trough and rises again. With respect to the magnitude curve and to the behavior of the phase curve in Fig. 5(c), the three peaks most likely correspond to antiresonances, while the first and the third troughs agree with resonances. At an antiresonance for

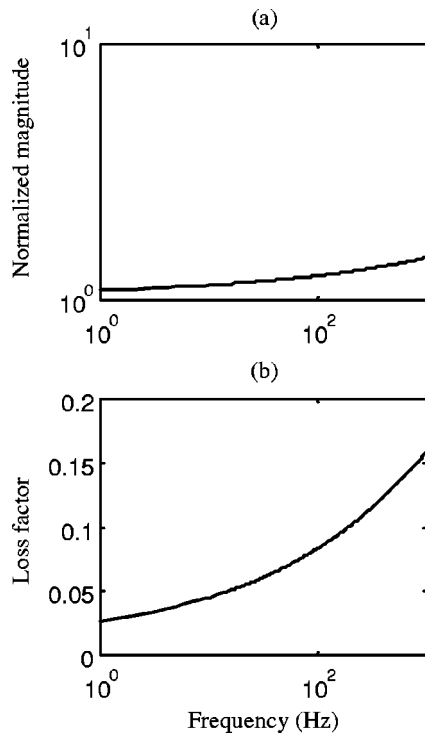


FIG. 6. (a) Normalized modulus magnitude $|\hat{\mu}|/\mu_\infty$ and (b) loss factor $\Im\hat{\mu}/\Re\hat{\mu}$ versus frequency.

elastic materials (that is, with no material damping) the stiffness shows a magnitude peak and a phase jump of -180° , while displaying a magnitude trough and a phase jump of $+180^\circ$ at a resonance. Introducing damping, the resonances and antiresonances are blunted; in general, the magnitude troughs and peaks are rounded, while the sudden phase jumps disappear, showing a “slower” phase shift. Therefore, it may become difficult to distinguish individual resonances and antiresonances, particularly for high damping material (such as rubber) at closely spaced resonances and antiresonances.

Regarding preload dependence, the low-frequency magnitude plateau shows an increase with preload as expected. Above ~ 500 Hz the stiffness magnitude increase with preload is pronounced, particularly for the largest preload.

Although heavily filled carbon black rubber eventually displays a dynamic strain dependence at larger amplitudes, the studied isolator stiffness is independent within a 100-fold dynamic strain alteration, throughout the whole frequency domain, using a similar isolator.⁵³

C. Material model parameters and calculated static and dynamic properties

A small test sample subjected to a controllable deformation is, traditionally, the method to estimate the material parameters. Here, the parameters are estimated from the static and dynamic measurements for the isolator. Although the mechanical fields within the isolator are complex, the axial motion of the isolator plates is readily generated and measured. By using the actual isolator, the effects of variations in component processing and in compound ingredients are eliminated. Here, the parameters to (33) are estimated from the dynamic measurements with vanishing isolator preloads,

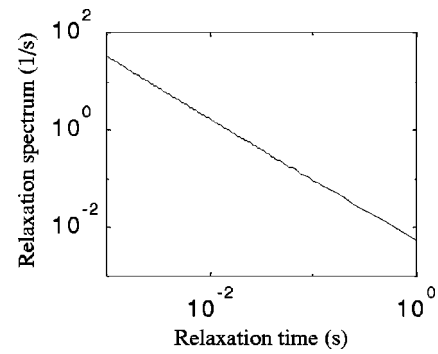


FIG. 7. Relaxation spectrum γ_{identity} versus relaxation time τ .

resulting in: $\mu_\infty = 5.94 \times 10^5$ N/m², $\mu_y = 13.0$ Ns/m², $\alpha_1 = 0.080$, $\alpha_2 = 0.625$, and $a = 2.22 \times 10^3$, together with $\rho = 1.05 \times 10^3$ kg/m³. The parameter values are from Ref. 54—estimating the parameters at vanishing preload while using a waveguide model—to which the reader may refer for details. The estimated parameters are realistic: the equilibrium Poisson’s ratio is 0.4998, the density equals the stated value, while the equilibrium bulk and shear moduli slightly exceed the stated values.

The normalized shear modulus in the frequency domain 1 to 1000 Hz is in Fig. 6, showing a slight magnitude and loss factor increase with frequency, as expected. The bulk modulus is a constant; 1.32×10^9 N/m².

The continuous relaxation spectrum γ_{identity} in the relaxation time τ domain 0.001 to 1 s is shown in Fig. 7, showing a rapid decrease with time.

The static measurements are used for the remaining nonlinear material parameter estimations. In particular, the Yeoh model (15) results in: $C_{20} = -4.50 \times 10^4$ N/m², $C_{30} = 1.50 \times 10^4$ N/m², other $C_{nm} = 0$, together with $C_{10} = \mu_\infty/2$. The parameters were fitted manually to yield as similar force–deflection curve as possible to that of measurement.

The isochoric part of the equilibrium Helmholtz free energy ψ versus the isochoric principal stretches $\bar{\lambda}_1$ and $\bar{\lambda}_2$ —given by a multiplicative split of the usual principal stretches λ_n into $\bar{\lambda}_n = J^{1/3} \lambda_n$ where $\bar{\lambda}_1 \bar{\lambda}_2 \bar{\lambda}_3 = 1$ and $n = 1, 2, 3$ —is shown in Fig. 8, for an extended stretch domain; clearly displaying an admissible isochoric free-energy behavior: Namely, there is one and only one stationary point, a

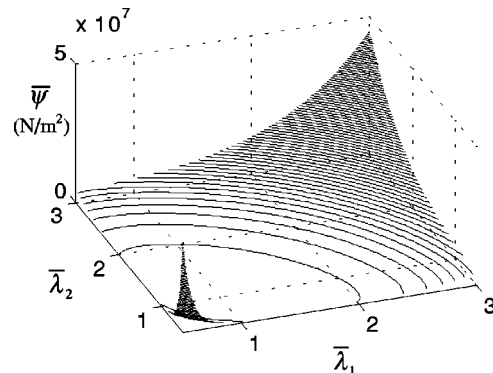


FIG. 8. Isochoric equilibrium Helmholtz free energy versus isochoric principal stretches.

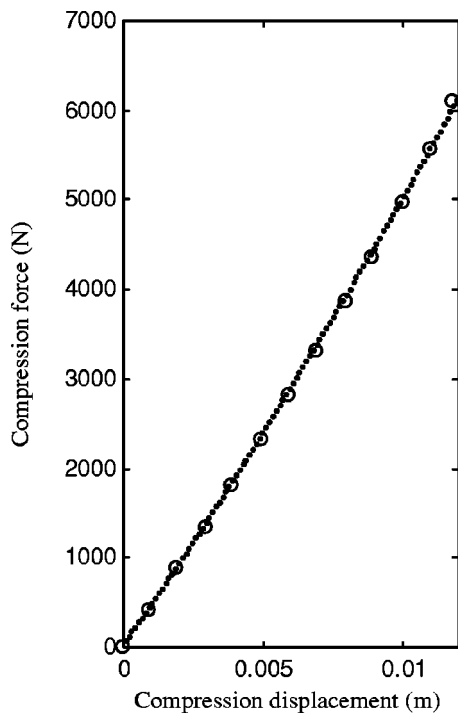


FIG. 9. Calculated (dots) and measured (circles) axial compression force versus displacement.

global minimum located in the undistorted state $\bar{\lambda}_1 = \bar{\lambda}_2 = 1$. Thus, estimated Yeoh parameters result in an admissible behavior.

The model's ability to agree with measurements for a variety of preloads, with the fixed parameters, is subsequently investigated. To further investigate the generality of the model, including both its static and dynamic behavior, other deformation modes may be applied, such as a bending. This is left for future studies, since the focus in this study is on the audible axial stiffness.

The dynamic stiffness and the time-delayed response are calculated employing 3000 (50×60) axisymmetric hybrid elements, within a mesh progressively refined towards the stress singularities at the rubber cylinder corners. Spatial integration of extrapolated Gauss point stresses yields a superior force estimate; where Gauss points closest to the central cylinder cross section are sufficient for the static axial force calculation, while the corresponding dynamic computation requires points drawn almost at the endplates. These estimates coincide at vanishing frequency due to declining inertia and viscoelastic contributions. Integration path shift from rubber cylinder ends into mesh incorporated massless—almost rigid—elastic endplates, avoids the stress singularities yielding an enhanced dynamic force estimation. As a result, the *a priori* superior Lagrangian nine-node (biquadratic displacement–bilinear pressure) and the four-node (bilinear displacement–constant pressure) hybrid elements, present negligible global differences. In addition, no stress-corrupting checkerboard mode³⁹ is observed for the latter elements.

The axial static force versus axial deformation 0.0 to -12.0 mm is shown in Fig. 9, showing negligible differences

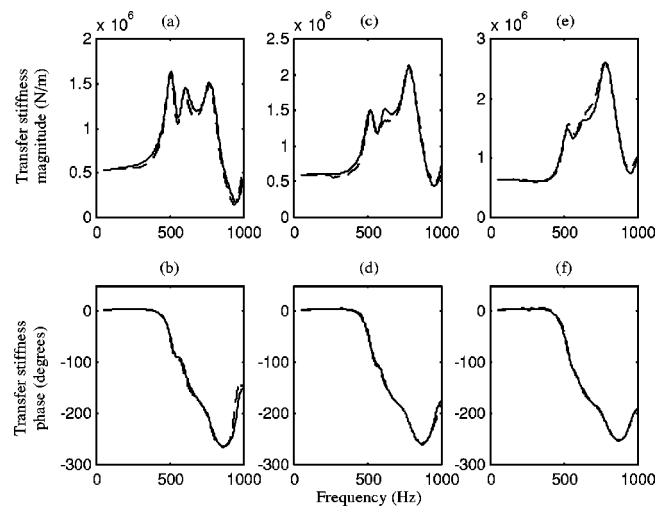


FIG. 10. Calculated (solid) and measured (dashed) transfer stiffness. (a) and (b) 0 N; (c) and (d) 2100 N; (e) and (f) 3350 N preload.

between Yeoh parameter fitted calculations and measurements.

The calculated stiffness in Figs. 10 and 11 is for a variety of preloads. Figure 11 displays the driving point stiffness while Fig. 10 shows the transfer stiffness. For comparison, the stiffness measurement results are in Fig. 10, plotted in dashed lines while the calculations are in solid.

The calculated transfer stiffness agrees very well with measurements for the examined preloads. The low-frequency plateau, the peaks and the troughs in Figs. 10(a), (c), and (e), almost coincide. In Figs. 10(b), (d), and (f), the calculated phase agrees with the measured phase with only small deviations.

In the considered frequency range, the driving point stiffness $\tilde{k}_{11} = \tilde{k}_{11}|_{\text{no plates}} - \omega^2 m_{\text{mp}}$, where m_{mp} is the mass of the rigid plate. Three isolator plate arrangements are in Fig. 11; plates not included are plotted in solid lines, 2.6-mm plates included are in dotted lines while (2.6+19.0) mm plates included are in dash-dotted lines.

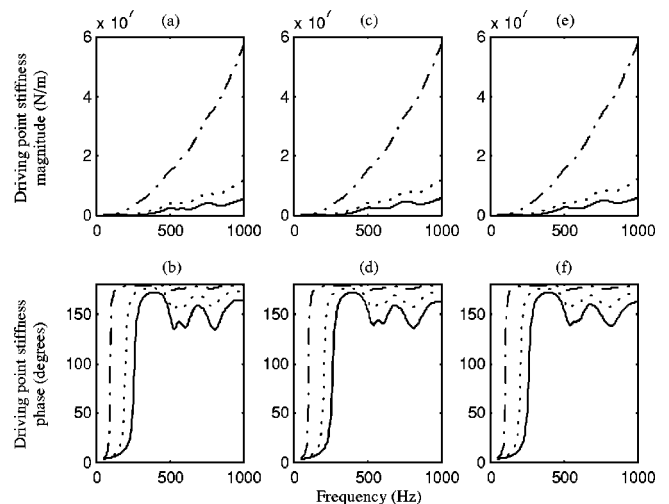


FIG. 11. Calculated driving point stiffness. Bonded steel plates not included (solid), 2.6 mm (dotted) and (2.6+19.0) mm bonded steel plates included (dash-dotted). (a) and (b) 0 N; (c) and (d) 2100 N; (e) and (f) 3350 N preload.

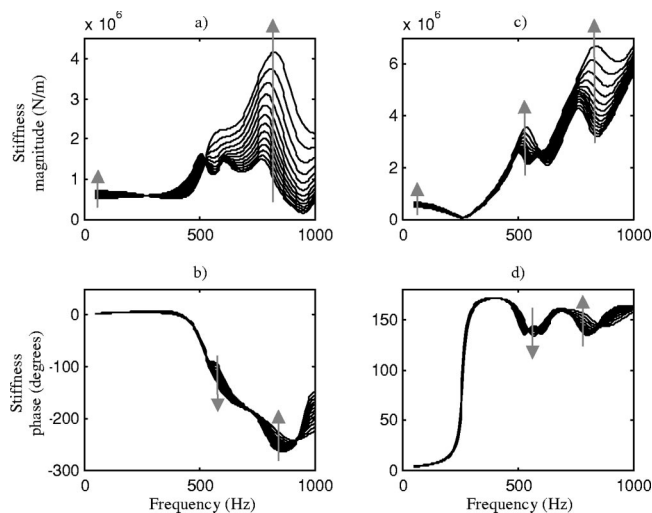


FIG. 12. Calculated (a) and (b) transfer stiffness; (c) and (d) driving point stiffness. From 0.0 to 12.0-mm precompression displacement with step of 1.0 mm. Bonded steel plates not included. Direction for increasing precompression is marked.

The magnitude curve of the driving point stiffness in Fig. 11(a) [12(c)] without preload and plates (that is, $m_{mp} \equiv 0$), drops directly into a deep trough at 250 Hz, rising steeply to peak at 500 Hz, then it drops again to a second trough at 540 Hz, rises to peak at 580 Hz, dropping to a third trough at 630 Hz, peaking again at 760 Hz. Finally, the curve drops to a fourth trough at 840 Hz and rises. With reference to the magnitude curve and the behavior of the corresponding phase curve in Fig. 11(b), the three peaks most likely correspond to antiresonances, while the four troughs agree with resonances. In Fig. 11(a), though more definite in Fig. 11(b), the first resonance is shifted to 190 and 90 Hz, respectively, for 2.6-mm and (2.6+19.0) mm plates.

As the preload increases, the magnitude curves without plates show a gentle growth in Figs. 11(c) and (e), while the first resonance frequency is roughly the same. Embodying the plates, the magnitude curves are dominated by the plate mass contributions.

Finally, the calculated stiffness in the extended deformation domain; 0.0 to -12.0 mm with -1.0-mm displacement increments, for 50 to 1000 Hz is shown in Fig. 12. Figures 12(a) and (b) show the transfer stiffness, while Figs. 12(c) and (d) display the driving point stiffness.

Clearly, the left side of the low-frequency plateau for transfer stiffness magnitude in Fig. 12(a) increases with preload—in line with the static predictions—while its right side declines. The frequency for the first transfer stiffness antiresonance slightly increases with the preload. Above ~500 Hz the transfer stiffness magnitude increase with preload is pronounced. The calculated high preload behavior of the transfer stiffness is validated in a special test rig using the same isolator, with the test facility described in Ref. 53.

The magnitude of the driving point stiffness in Fig. 12(c) without plates, increases with preload throughout the whole frequency domain, while the first resonance frequency is essentially constant.

Regarding the transmitted energy flow through vibration isolators, the driving point stiffness imaginary part entirely

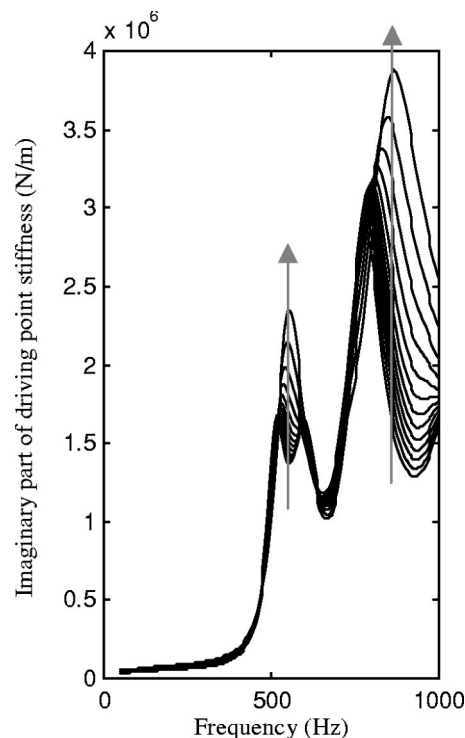


FIG. 13. Imaginary part of calculated driving point stiffness. From 0.0 to 12.0-mm precompression displacement with step of 1.0 mm. Direction for increasing precompression is marked.

determines this flow in contrast to the transfer stiffness, where both the real and imaginary part must be taken into account. The former imaginary part is, however, independent on the mass of the rigid plates and varies moderately with the preload, being particularly notable in the high-frequency region in Fig. 13 and is of the same order as the transfer stiffness magnitude.

1. Other material models

To further examine the short-term influence on stiffness, two popular models in rubber viscoelasticity are applied to the nearly incompressible model. The parameters are adjusted to yield equal ground-state density, bulk modulus, equilibrium shear modulus, and the loss factor of shear modulus (at 100 Hz), as in the current model; resulting in e.g., an equal static response. The transfer stiffness in the frequency domain 50 to 1000 Hz at 2100 N preload due to frequency-independent shear modulus, $\hat{\mu}/\mu_\infty = 1 + 8.35 \times 10^{-2} i$, and ordinary Kelvin-Voigt model (35), $\hat{\mu}/\mu_\infty = 1 + 8.35 \times 10^{-4} f i$, is shown in Fig. 14(a). Also shown are the measurement results and previously determined transfer stiffness, with the shear modulus in Fig. 6. The corresponding driving point stiffness is in Fig. 14(b). Obviously, the popular models fail to model the stiffness adequately. In particular, the first shows a low material damping for high frequencies while the second example is correspondingly high, in addition to poorly predicted resonance and antiresonance frequencies. Although it is possible to adjust the examples to give better stiffness agreements for $f < 400$ Hz, the equilib-

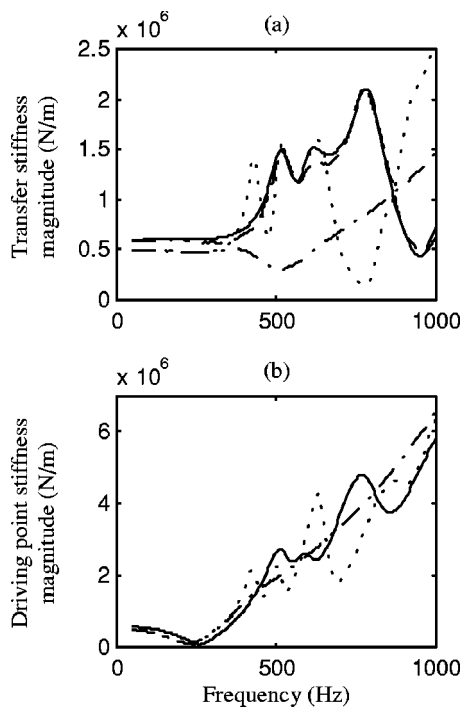


FIG. 14. Calculated (a) transfer stiffness and (b) driving point stiffness magnitude. Measurement (dashed); present (solid thick), frequency independent (dotted), and Kelvin-Voigt model (dash-dotted). 2100 N preload. Bonded steel plates not included.

rium shear moduli would be unrealistically large (and erroneous). The large equilibrium shear moduli would also result in an incorrect static stiffness prediction.

IV. CONCLUSIONS

A nonlinear preload dynamic stiffness model for vibration isolators within an audible frequency domain of particular interest for noise abatement is presented. The problem of simultaneously modeling the rubber prestrain dependence and its audible short-term response is addressed. A nearly incompressible material model is adopted, obeying decoupled spherical and deviatoric responses, being elastic in dilatation while displaying viscoelasticity in deviation. The latter exhibits a time strain separable relaxation tensor with a single function embodying its time dependence. An updated Lagrangian nonlinear finite-element procedure solves the weak formulations corresponding to the stiffness problem. The method is applied to a cylindrical vibration isolator. The results agree strikingly well with static and dynamic measurements within an audible frequency domain at various preloads. In particular, the calculated low-frequency plateau, the peaks and troughs of the transfer stiffness magnitude almost coincide with measurements. Above 500 Hz the transfer stiffness magnitude increase with preload is pronounced. The driving point stiffness magnitude is dominated by the isolator plate mass contributions. Above 500 Hz the (plate mass independent) imaginary part of the driving point stiffness is of same order as the transfer stiffness magnitude.

To derive other components of the dynamic stiffness is an interesting continuation of this work.

ACKNOWLEDGMENTS

Trelleborg Industrial AVS donation of vibration isolators and Adam Zdunek's excellent ANIA code support and good advice are gratefully acknowledged.

- ¹A. C. Pipkin and R. S. Rivlin, *Arch. Ration. Mech. Anal.* **8**, 297 (1961).
- ²B. D. Coleman and W. Noll, *Rev. Mod. Phys.* **33**, 239 (1961).
- ³B. D. Coleman and W. Noll, *Rev. Mod. Phys.* **36**, 1103 (1964).
- ⁴G. Lianis, in *Proceedings of Fourth International Congress on Rheology*, Providence, Rhode Island, 1963 (Interscience, New York, 1965), pp. 109–119.
- ⁵W. V. Chang, R. Bloch, and N. W. Tschoegl, *Rheol. Acta* **15**, 367 (1976).
- ⁶K. N. Morman, *Rheol. Acta* **27**, 3 (1988).
- ⁷A. Zdunek, *Comput. Methods Appl. Mech. Eng.* **105**, 63 (1993).
- ⁸B. Bernstein, E. A. Kearsley, and L. J. Zappas, *Trans. Soc. Rheol.* **7**, 391 (1963).
- ⁹J. L. Sullivan, *J. Rheol.* **31**, 271 (1987).
- ¹⁰J. L. Sullivan and K. A. Mazich, *Rubber Chem. Technol.* **62**, 68 (1989).
- ¹¹M. B. Bush, R. I. Tanner, and N. Phan-Thien, *J. Non-Newtonian Fluid Mech.* **18**, 143 (1985).
- ¹²K. N. Morman, in *Proceedings, Mechanics of Materials and Structures*, Charlottesville, Virginia, 1993 (1994), Vol. 35, pp. 297–318.
- ¹³S. Govindjee and J. Simo, *J. Mech. Phys. Solids* **39**, 87 (1991).
- ¹⁴S. Govindjee and J. Simo, *J. Mech. Phys. Solids* **40**, 213 (1992).
- ¹⁵S. Govindjee and J. Simo, *Int. J. Solids Struct.* **29**, 1737 (1992).
- ¹⁶D. J. Ewins, *Modal Testing: Theory and Practice* (Research Studies, Baldock, 1991).
- ¹⁷R. L. Bagley and P. J. Torvik, *AIAA J.* **21**, 741 (1983).
- ¹⁸G. A. Lesieutre, Ph.D. thesis, Aerospace Engineering, University of California, Los Angeles (1989).
- ¹⁹G. A. Lesieutre and D. L. Mingori, *J. Guid. Control Dyn.* **13**, 1040 (1990).
- ²⁰G. A. Lesieutre, *Int. J. Solids Struct.* **29**, 1567 (1992).
- ²¹G. A. Lesieutre and E. Biauchini, in *Proceedings of the 34th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Material Conference* (La Jolla, California, 1993), pp. 2101–2119.
- ²²K. Dovstam, *Int. J. Solids Struct.* **32**, 2835 (1995).
- ²³Y. C. Yiu, in Ref. 21, pp. 2110–2119.
- ²⁴D. J. McTavish and P. C. Hughes, *J. Vibr. Acoust.* **115**, 103 (1993).
- ²⁵M. A. Biot, *J. Appl. Phys.* **25**, 1385 (1954).
- ²⁶M. Enelund, Ph.D. thesis, Division of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden (1996).
- ²⁷S. H. Crandall, in *Air Space and Instruments, Draper Anniversary Volume*, edited by S. Lees (McGraw-Hill, New York, 1963), pp. 183–193.
- ²⁸R. J. Gaylord, G. H. Weiss, and E. A. Dimarzio, *Macromolecules* **19**, 927 (1986).
- ²⁹N. Makris and M. C. Constantinou, *ASCE J. Eng. Mech.* **119**, 1453 (1993).
- ³⁰W. G. Glockle and T. F. Nonnenmacher, *Macromolecules* **24**, 6426 (1991).
- ³¹Y. A. Rossikhin and M. V. Shitikova, *Appl. Mech. Rev.* **50**, 15 (1997).
- ³²N. Shimizu and W. Zhang, *JSME Int. J., Ser. C* **42**, 825 (1999).
- ³³J. E. Marsden and T. J. R. Hughes, *Mathematical Foundations of Elasticity* (Dover, New York, 1994).
- ³⁴A. R. Payne and R. E. Whittaker, *Rubber Chem. Technol.* **44**, 440 (1971).
- ³⁵C. Truesdell and W. Noll, *The Non-Linear Field Theories of Mechanics* (Springer, Berlin, 1992).
- ³⁶M. F. Beatty, *Appl. Mech. Rev.* **40**, 1699 (1987).
- ³⁷O. H. Yeoh, *Rubber Chem. Technol.* **63**, 792 (1990).
- ³⁸K. N. Morman and J. C. Nagtegaal, *Int. J. Numer. Methods Eng.* **19**, 1079 (1983).
- ³⁹T. I. R. Hughes, *The Finite Element Method* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
- ⁴⁰R. C. Koeller, *J. Appl. Mech.* **51**, 299 (1984).
- ⁴¹M. Caputo, *Geophys. J. R. Astron. Soc.* **13**, 529 (1967).
- ⁴²M. Caputo, *Elasticità e Dissipazione* (Bologna, 1969).
- ⁴³M. Caputo and F. A. Mainardi, *Pure Appl. Geophys.* **91**, 134 (1971).
- ⁴⁴W. L. Wainwright, *Lect. Notes Math.* **457**, 298 (1975).
- ⁴⁵B. Ross, *Mathematics Mag.* **50**, 115 (1977).
- ⁴⁶I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, London, 1964), Vol. I.
- ⁴⁷P. J. Torvik and R. L. Bagley, *ASME, DE* **5**, 125 (1987).
- ⁴⁸J. G. Curro, D. S. Pearson, and E. Helfand, *Macromolecules* **18**, 1157 (1985).

- ⁴⁹J. D. Ferry, *Viscoelastic Properties of Polymers* (Wiley, New York, 1980).
- ⁵⁰P. G. Nutting, J. Franklin Inst. **191**, 679 (1921).
- ⁵¹B. Gross, *Mathematical Structure of the Theories of Viscoelasticity* (Hermann, Paris, 1953).
- ⁵²A. Zdunek, Ph.D. thesis, Department of Solid Mechanics, Royal Institute of Technology, Stockholm, Sweden (1992).
- ⁵³L. Kari, Noise Control Eng. J. **49**, 88 (2001).
- ⁵⁴L. Kari, J. Sound Vib. **244**, 211 (2001).

A substructure approach for the midfrequency vibration of stochastic systems

Abhijit Sarkar^{a)}

Department of Mechanical Engineering, McGill University, 817 Sherbrooke Street, Montreal, Quebec, Canada

Roger Ghanem^{b)}

Department of Civil Engineering, The Johns Hopkins University, Baltimore, Maryland 21218

(Received 23 January 2002; accepted for publication 14 January 2003)

A novel substructure coupling technique based on the proper orthogonal decomposition method is presented for the midfrequency range vibration of linear dynamical systems with parameter uncertainty. For a given frequency band, the methodology permits the derivation of an adaptive basis for each subsystem and the construction of a reduced-order model of the global structure. The formulation is directed toward the efficient probabilistic characterization of model-based predictions in the framework of a stochastic finite element method. The efficiency of the substructure method has been contrasted both from the viewpoint of adopting free-free and fixed-fixed substructure proper orthogonal modes in order to arrive at a reduced subsystem model. The distinction as well as similarity of the present methodology with the component mode synthesis is also pointed out. The proper orthogonal modes are obtained from both frequency- and time-domain approaches, and their suitability is discussed in relation to the behavior of a specific system. The substructure approach elegantly integrates with a version of stochastic finite elements based on orthogonal decompositions and projections of stochastic processes. © 2003 Acoustical Society of America.

[DOI: 10.1121/1.1558374]

PACS numbers: 43.40.Cw, 43.40.Qi, 43.58.Ta [RLW]

I. INTRODUCTION

The analysis of linear dynamical systems often involves representing the original continuous differential operator, relating to the equilibrium and evolution of the system, by a finite-dimensional reduced-order model which can efficiently reproduce the salient features in the behavior of the original system. The traditional modal analysis in conjunction with the finite element method (FEM) has been established as an efficient tool for linear dynamic analysis in the low frequency range as only a few structural natural modes can effectively capture the system behavior. For high-frequency vibration, the statistical energy analysis (SEA) has been proven to be an efficient tool as a very large number of resonant modes contribute to the total response. In the case of SEA, the sensitivity of the high frequency response to small perturbations in the dynamical parameters is accounted for by modeling the natural frequencies of the subsystems as random points along the frequency axis. In the case of interconnected subsystems, the applicability of the SEA is therefore restricted to those cases where each subsystem displays uniformly high modal density.

In the medium frequency range, the short wavelength vibration patterns necessitate a fine spatial resolution yielding large-scale matrices. Moreover, and in contrast to low frequency range vibrations, a large number of structural modes, mostly in the medium-frequency band, contribute to the total response. The calculation of these midfrequency

modes with adequate precision is often beset with difficulties due to a number of reasons. On the one hand, computational procedures for evaluating eigenmodes beyond the first few are sensitive to numerical errors, with the effect of these errors being very significant for the higher modes. This particular difficulty is more pronounced the larger the scale of the computational problem. Moreover, the significance of damping is pronounced in the midfrequency range, with the predicted response being very sensitive to its form and magnitude, thus justifying the use of more accurate nonproportional damping models which yield complex mode shapes. Along similar lines, certain model-reduction and adaptation techniques require recourse to complex mode shapes in order to minimize the bleeding from outside the frequency band of interest. In these situations, and in particular for very large-scale systems, the traditional modal superposition technique becomes impractical. While using a direct numerical integration technique to determine the response due to broadband deterministic or random excitations, a simulation with long duration and a very small integration time step becomes necessary which is often associated with computational difficulties yielding significant numerical noise in the response predictions. On the other hand, the applicability of SEA becomes questionable as different subsystems exhibit high and low frequency behavior simultaneously, resulting in non-uniform modal density of the global structure in the midfrequency band. Furthermore, the presence of imperfections, discontinuities, incomplete and imprecise knowledge of subsystem coupling parameters and secondary systems, significantly influences the short wavelength vibration features in the midfrequency range which raises questions regarding the

^{a)}Electronic mail: asarkar@mecheng.mcgill.ca

^{b)}Electronic mail: ghanem@jhu.edu

warranted confidence in the predicted behavior even with detailed deterministic finite element method (FEM) models. Probabilistic models are usually adopted to characterize the uncertainty in the dynamical parameters of the system, thus permitting the quantification of the confidence in the simulated responses. This approach, typically, increases significantly the computational requisites for the analysis.

A number of studies have reported recently in the literature on the efficient analysis of midfrequency range vibration. These papers are based on the ideas of adopting a frequency adaptive basis¹⁻⁷ or a combination of modal analysis and SEA.^{8,9} Most of the aforementioned methodologies are formulated in the context of global systems which can be expensive in terms of numerical computation in handling large-scale finite element models. A substructure approach can be efficiently used to overcome this limitation, especially when the dynamical system itself is composed of built-up and interconnected subsystems.

The traditional substructuring coupling methods¹⁰⁻¹² such as the popular *Guyan and Irons* condensation method and the *Craig and Bampton* component mode synthesis (CMS) procedure are frequently used for the following two reasons: (a) it is advantageous to reduce the size of the eigenvalue problem at the outset in order to obtain lower order natural frequencies of the global system and (b) in the context of a large-scale system, the entire structure is frequently divided into several parts, each of which is analyzed separately and assembled subsequently to reconstruct the complete model. These methods necessitate a careful definition and treatment of the interconnecting edges and surfaces of the substructures so as to ensure the compatibility of different system components. In the CMS procedure, the boundary degrees of freedom are represented by the static condensation procedure while the internal variables are expressed by vibration mode shapes of the subsystem obtained by fixing its boundaries to the adjacent substructures. A unified formulation of the other variants of the CMS procedure based on fixed, free, loaded or any other hybrid combination of these basic options for subsystem interfaces has also been developed.¹¹

In some applications, the structural dynamics models are coupled with fluid dynamic systems. The distinguishing feature of the fluid systems, in contrast to the structural dynamic model, is the presence of nonsymmetric matrices (although the presence of gyroscopic effects can also lead to nonsymmetric structural matrices). In such cases, it becomes increasingly difficult to perform the complex eigenvalue analysis of the fluid systems, especially when the size of the system matrices is large, leading to difficulties in the modal analysis procedure.

The aforementioned substructuring approaches are based on the modal analysis of the component substructures ideally suited for low frequency vibration analysis. In the medium frequency ranges, however, as mentioned previously, it is necessary to finely discretize the subsystems spatially in order to capture small amplitude wavelength vibration features. Clearly, very large-scale computational models ensue. In performing the modal analysis of these components, the higher order subsystem modes are polluted by numerical errors thus

undermining the efficacy of the substructuring approach. In parallel, sensitivity of the response due to perturbations in the parameters as well as unmodeled system dynamics demands a probabilistic analysis for reliable response prediction thus further increasing the computational requisite for the analysis. Recent efforts have been reported in the literature for developing substructuring techniques adapted to the medium-frequency range.¹³ These techniques are based on the energy operator concept,^{3,4} and have been restricted to deterministic systems.

Recently,^{14,15} a new methodology has been explored for the probabilistic analysis of midfrequency dynamics. The suitability of this method for the numerical treatment of midfrequency structural dynamics has been validated in the context of the stochastic finite element method. This paper continues this investigation to develop a novel scheme for substructuring in the midfrequency vibration analysis of stochastic systems. The proper orthogonal decomposition method is used as the vehicle for model reduction. In particular, the proper orthogonal modes of free-free and fixed-fixed subsystems are used to obtain the reduced-order subsystem model which accounts for the parameter uncertainty of each subsystem independently. The reduced-order subsystem models are then combined to create the global reduced system, in analogy with the traditional FEM assembly. The methodology avoids any operation involving the full-scale global matrices. Moreover, the approach presents a manageable methodology for the quantification and propagation of uncertainty in these systems.

Section II will present a review of probabilistic concepts including the Hilbert space formalism used to effect the orthogonal decompositions and projections used in the stochastic characterizations. Following that, a review is presented of model reduction techniques in the medium-frequency range with an emphasis on the proper orthogonal decomposition method (PODM). This is followed by the introduction and the analysis of a novel substructuring method that capitalizes on the PODM properties and the stochastic projections.

II. REVIEW OF PROBABILISTIC CONCEPT

In this section, a brief review is presented of the probabilistic approach to be followed in the remainder of the paper. This approach will be used subsequently for the system parameter uncertainty analysis. This particular approach capitalizes on the possibility of perceiving random variables as functions on the space of elementary events. As functions, various techniques from deterministic approximation theory are applicable to random variables. The main question to be addressed in this context concerns methods to characterize the solution to a physical problem where some parameters of the model have been modeled as stochastic processes. The answer to this question lies in the realization that in the deterministic finite element method, as well as most other numerical analysis techniques, a solution to a boundary-value problem is quantified once its projection on a basis in an appropriate function space has been evaluated. In the stochastic case, the solution is a stochastic process, and is similarly identified with its projection on a set of appropriately chosen basis functions. The random character of these basis

functions, as well as any other random quantity, will be identified through an argument, θ , which can be construed as representing an additional dimension to include in any approximation scheme. The procedure of Monte Carlo simulation can be viewed as a collocation along this θ dimension. Other approximations along this dimension are possible, and can be associated with different choices of basis functions in the corresponding space of random variables. This theoretical development is consistent with the identification of the space of second-order random variables as a Hilbert space with the inner product on it defined as the operation of statistical correlation.¹⁶ As a reminder, second-order random variables are those random variables with finite variance, they are mathematically similar to deterministic functions with finite energy.

The mathematical model to be adopted can be summarized as follows: The data, modeled as random variables or processes, span the Hilbert space \mathcal{H}_G . Assuming the data to be well defined in a probabilistic sense provides a full characterization of this space, in which a set of basis functions, $\{\xi\}$, must be identified. This is accomplished in Sec. II A on Karhunen–Loeve expansions. The state of the system, again modeled as a random process, resides in the Hilbert space \mathcal{H}_L . A set of basis functions, $\{\Psi\}$, is also identified in this space, which in general is different from the basis $\{\xi\}$ since this latter one spans only a subset of the space of second-order random variables, namely those that characterize the data. Identifying a basis for the space \mathcal{H}_L is accomplished in Sec. II B.

A. Karhunen–Loeve expansion

The Karhunen–Loeve expansion¹⁶ of a stochastic process, $\alpha(\mathbf{x}, \theta)$, is based on the spectral expansion of its covariance function $R_{\alpha\alpha}(\mathbf{x}, \mathbf{y})$. Here, \mathbf{x} and \mathbf{y} are used to denote spatial coordinates. The covariance function being symmetrical and positive definite, by definition, has all its eigenfunctions mutually orthogonal, and they form a complete set spanning the function space to which $\alpha(\mathbf{x}, \theta)$ belongs. It can be shown that if this deterministic set is used to represent the process $\alpha(\mathbf{x}, \theta)$, then the random coefficients used in the expansion are also orthogonal (i.e., uncorrelated). The expansion then takes the following form:

$$\alpha(\mathbf{x}, \theta) = \bar{\alpha}(\mathbf{x}) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i(\theta) \phi_i(\mathbf{x}), \quad (1)$$

where $\bar{\alpha}(\mathbf{x})$ denotes the mean of the stochastic process, and $\{\xi_i(\theta)\}$ forms a set of orthogonal random variables. Furthermore, $\{\phi_i(\mathbf{x})\}$ are the eigenfunctions and $\{\lambda_i\}$ are the eigenvalues, of the covariance kernel, and can be evaluated as the solution to the following integral equation:

$$\int_{\mathcal{D}} R_{\alpha\alpha}(\mathbf{x}, \mathbf{y}) \phi_i(\mathbf{y}) d\mathbf{y} = \lambda_i \phi_i(\mathbf{x}), \quad (2)$$

where \mathcal{D} denotes the spatial domain over which the process $\alpha(\mathbf{x}, \theta)$ is defined.

The most important aspect of this spectral representation is that the spatial random fluctuations have been decomposed into a set of deterministic shapes multiplying random ampli-

tudes. If the random process being expanded, $\alpha(\mathbf{x}, \theta)$, is Gaussian, then the random variables $\{\xi_i\}$ form an orthonormal Gaussian vector. The Karhunen–Loeve expansion is mean-square convergent irrespective of the probabilistic structure of the process being expanded, provided it has a finite variance.¹⁶ The closer a process is to white noise, the more terms are required in its expansion, while at the other limit, a random variable can be represented by a single term.

B. Polynomial chaos expansion

The covariance function of the solution process is not known a priori, and hence the Karhunen–Loeve expansion cannot be used to represent it. Thus, although the solution process spans a subset of the Hilbert space of random variables, that space cannot be delineated a priori, and therefore an adapted basis set cannot be identified. In this case, a generic basis, which is complete in the space of all second-order random variables, will be identified and used in the approximation process. Since the solution process is a function of the material properties, nodal solution variables, denoted here by $u(\theta)$, can be formally expressed as some nonlinear functional of the set $\{\xi_i(\theta)\}$ used to represent the material stochasticity. It has been shown that this functional dependence can be expanded in terms of polynomials in Gaussian random variables, referred to as polynomial chaoses.^{17–19} Namely,

$$u(\theta) = a_0 \Gamma_0 + \sum_{i_1=1}^{\infty} u_{i_1} \Gamma_1(\xi_{i_1}(\theta)) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} u_{i_1 i_2} \Gamma_2(\xi_{i_1}(\theta), \xi_{i_2}(\theta)) + \cdots \quad (3)$$

In Eq. (3), the symbol $\Gamma_n(\xi_{i_1}, \dots, \xi_{i_n})$ denotes the polynomial chaos of order n in the variables $(\xi_{i_1}, \dots, \xi_{i_n})$. These are generalizations of the multidimensional Hermite polynomials to the case where the independent variables are themselves measurable functions (in this case they are random variables). Introducing a one-to-one mapping to a set with ordered indices denoted by $\{\Psi_j(\theta)\}$ and truncating the polynomial chaos expansion after the P th term, Eq. (3) can be rewritten as

$$u(\theta) = \sum_{j=0}^P u_j \Psi_j(\theta). \quad (4)$$

These polynomials are orthogonal in the sense that their inner product $\langle \Psi_j \Psi_k \rangle$, which is defined as the statistical average of their product, is equal to zero for $j \neq k$. Moreover, they can be shown to form a complete basis in the space of second-order random variables. A complete probabilistic characterization of the process $u(\theta)$ is obtained once the deterministic coefficients u_j have been calculated. A given truncated series can be refined along the random dimension either by adding more random variables to the set $\{\xi_i\}$ or by increasing the maximum order of polynomials included in the polynomial chaos expansion. The first refinement takes into account higher frequency random fluctuations of the underlying stochastic process, while the second refinement cap-

tures strong nonlinear dependence of the solution on this underlying process.

III. MIDFREQUENCY STRUCTURAL DYNAMICS OF UNCERTAIN SYSTEMS

For a linear time invariant distributed parameter system, the frequency domain finite element approximation of the solution in an n -dimensional subspace is given by

$$u_n(\mathbf{x}, \omega) = \sum_{i=1}^n q_i(\omega) N_i(\mathbf{x}), \quad (5)$$

with $N_i(\mathbf{x})$ being the finite element shape functions and q_i the i th nodal response quantity. Consequently, the governing equation of motion can be expressed as

$$\mathbf{A}_n \mathbf{q} = \mathbf{F}_n, \quad (6)$$

where the dynamic stiffness matrix, \mathbf{A}_n , is given by

$$\mathbf{A}_n = -\omega^2 \mathbf{M}_n + i\omega \mathbf{D}_n + \mathbf{K}_n, \quad (7)$$

with \mathbf{M}_n , \mathbf{D}_n , and \mathbf{K}_n being the mass, damping, stiffness matrices and force vector, respectively.

A. Deterministic proper orthogonal decomposition method

The dimensionality, spatial and temporal complexities of midfrequency dynamics have been successfully addressed using the proper orthogonal decomposition.^{20–25} The mathematical underpinnings of the PODM are identical to the Karhunen–Loeve expansion presented earlier. The method optimally extracts spatial information and identifies the dimensionality of a system from a set of time series data of the response gathered from numerical or physical experiments. For a time series $u(\mathbf{x}, t)$, the proper orthogonal modes, or *coherent structures*, $\{\Phi\}$ are obtained by solving the integral eigenvalue problem,

$$\int_{\Omega} \langle u(\mathbf{x}, t) u(\mathbf{y}, t) \rangle > \Phi_i(\mathbf{y}) d\mathbf{y} = \lambda_i \Phi_i(\mathbf{x}). \quad (8)$$

Here Ω is the spatial extent of the problem, $\langle \cdot \rangle$ is the temporal averaging operator, and the eigenfunctions $\{\Phi_i(\mathbf{x})\}$ represent the dominant coherent fluctuations [or the proper orthogonal modes (POM)] and λ_i represents the amount of energy captured by the corresponding i th mode. The correlation function being symmetric and positive-definite, the POMs form a complete orthonormal basis which can represent a FEM response field. They are also optimal in the mean-square sense as they capture more energy, in a finite subset, than any other set of basis functions. The optimal dimension N of the system can be determined such that $\sum_{i=1}^N \lambda_i / \sum_{i=1}^n \lambda_i \geq 99\%$ (n being total degrees of freedom), assuming the sufficiency of the POMs to capture 99% of the energy of the signal. The PODM has been successfully demonstrated for the construction of computationally efficient reduced-order models for midfrequency vibration^{14,15} and nonlinear dynamics.^{26,27} In the context of a finite element discretization, Eq. (8) can be expressed in matrix form as

$$\mathbf{C}_{u_n u_n} \Phi_i = \lambda_i \Phi_i, \quad (9)$$

where $\mathbf{C}_{u_n u_n}$ is obtained from Eq. (8) through appropriate shape function integration and assembly of elemental matrices.²⁸ Alternatively, $\mathbf{C}_{u_n u_n}$ can be obtained from a frequency-based formulation in the following form:

$$\mathbf{C}_{u_n u_n} = \int_B \Re[\mathbf{A}_n^{*-1} \mathbf{A}_n^{-1}] d\omega. \quad (10)$$

Here, \mathbf{A}_n is the dynamic stiffness matrix given in Eq. (7). Implicit in Eq. (10) is the assumption that the subsystem is forced to vibrate with an incoherent band-limited stationary vector white noise with unit strength to excite the entire spectrum of natural frequencies in the midfrequency band of interest.

Once the dominant eigen-subspace has been identified, an $n \times N$ transformation matrix, Σ , can be constructed whose columns are the N eigenvectors corresponding to the N highest eigenvalues in Eq. (9). The reduced system, representing the projection of the original operator on the dominant eigenspace, can be represented similarly to Eq. (6), with reduced dynamic stiffness matrix \mathbf{A}_N , force vector \mathbf{F}_N , and transformed coordinate \mathbf{q} given by

$$\mathbf{A}_N(\omega) = \Sigma^T \mathbf{A}_n(\omega) \Sigma, \quad (11)$$

$$\mathbf{F}_N = \Sigma^T \mathbf{F}_n, \quad (12)$$

$$\mathbf{q} = \Sigma^T \mathbf{U}. \quad (13)$$

B. Probabilistic reduction of uncertain dynamical system: Karhunen–Loeve/polynomial chaos formalism

The above-presented midfrequency vibration analysis procedure, demonstrated in the context of deterministic systems, can be exploited for the probabilistic uncertainty analysis of these systems. The stochastic finite element method based on the orthogonal decompositions and projections²⁹ readily integrates with the model reduction strategy using PODM in the midfrequency range. It has been demonstrated^{14,15} that the dominant POMs obtained from using the mean system properties in Eq. (10) form an efficient basis for the midfrequency vibration of stochastic systems. Thus, expanding the reduced dynamic system matrix, \mathbf{A}_N , in its Karhunen–Loeve expansion,²⁹ the following expression for the equilibrium equation is obtained for the uncertain dynamical system:

$$\left[\sum_{k=0}^{nkl} \xi_k \mathbf{A}_N^k \right] \mathbf{U} = \mathbf{F}_N, \quad (14)$$

where the quantity \mathbf{A}_N^k is the reduced dynamic stiffness matrix associated with the k th eigenvector of the covariance matrix of the random material property, and is given by

$$\mathbf{A}_N^k = \Sigma^T \mathbf{A}_n^k(\omega) \Sigma. \quad (15)$$

Here nkl represents the number of terms retained in the Karhunen–Loeve expansion, and $\{\xi_i\}$ are the random variables obtained from the Karhunen–Loeve expansion, and are Gaussian whenever the material properties are modeled as Gaussian processes. Consequently, the solution process is approximated by its finite polynomial chaos decomposition as

$$\mathbf{U} = \sum_{j=0}^{npc} \Psi_j(\theta) \mathbf{U}_j, \quad (16)$$

where npc is the number of terms retained in the polynomial chaos expansion. Substituting the above-given expansion into the reduced model equation and projecting the error onto each of the polynomial chaoses used in the approximation results in

$$\sum_{j=0}^{npc} \sum_{i=0}^{nkl} \langle \xi_i \Psi_j \Psi_k \rangle \mathbf{A}_N^i \mathbf{U}_j = \langle \Psi_k \mathbf{F}_N \rangle, \quad k=0, \dots, npc, \quad (17)$$

where the deterministic coefficients \mathbf{U}_j are computed as the solution of the aforementioned deterministic matrix equation.

It should be noted that for stochastic systems, the frequency response curve, at a fixed value of frequency, is a complex valued random variable whenever the system parameters are modeled as random fields. A pronounced non-stationary trend is introduced in the frequency response curve due to the presence of natural frequencies with the standard deviation peaking up significantly near the resonant points even with small random perturbations of the system parameters. Analytical and computational models for the extremes of the response can be adversely affected by this sensitivity.

IV. SUBSTRUCTURE METHOD

The midfrequency dynamic analysis procedure described in the previous sections has been formulated in the context of the entire structural systems. For a large-scale structural assembly, it becomes often convenient and necessary to divide the total system into a number of subsystems and to construct the reduced-order model of each subsystem separately. In this section, a novel substructuring approach based on the proper orthogonal decomposition technique is formulated for the construction of reduced-order models in the midfrequency range.

The dynamic equilibrium equation of a typical substructure (identified with the superscript s), having n_s degrees of freedom, can be written as

$$\begin{bmatrix} \mathbf{A}_{BB}^s(\omega) & \mathbf{A}_{BI}^s(\omega) \\ \mathbf{A}_{IB}^s(\omega) & \mathbf{A}_{II}^s(\omega) \end{bmatrix} \begin{Bmatrix} \mathbf{U}_B^s \\ \mathbf{U}_I^s \end{Bmatrix} = \begin{Bmatrix} \mathbf{F}_B^s \\ \mathbf{F}_I^s \end{Bmatrix}. \quad (18)$$

The partitions B and I relate to the boundary dofs n_B and interior dofs n_I of the substructure with $n_s = n_B + n_I$. This convention is analogous to the Craig and Bampton CMS formalism. This partition conveniently allows one to perform

the assembly of all the substructure within a finite elements framework.

A. Substructure approach using fixed-fixed subsystem proper orthogonal modes

Analogous to the CMS methodology, while keeping the boundary dofs unaltered, the interior dofs are transformed according to

$$\mathbf{U}_{I[n_I \times 1]}^s = \mathbf{\Sigma}_{[n_I \times m]}^s \mathbf{q}_{[m \times 1]}^s, \quad (19)$$

where n_I and m denote the number of internal dofs and dominant proper orthogonal modes, respectively, and the bracketed subscripts refer to the dimensions of the respective matrices. This transformation draws on information obtained from the measured (or simulated) motion of the system in the frequency band of interest, and can therefore be construed as having been adapted to that frequency band. The procedure to obtain $\mathbf{\Sigma}^s$ is discussed in the subsequent sections. The resulting transformation matrix for the total subsystem is of the form,

$$\mathbf{R} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\bar{\mathbf{A}}_{II}^{s-1} \bar{\mathbf{A}}_{IB}^s & \mathbf{\Sigma}^s \end{bmatrix}. \quad (20)$$

Here $\bar{\mathbf{A}}_{II}^s$ and $\bar{\mathbf{A}}_{IB}^s$ are submatrices of the dynamic stiffness matrix calculated at the central frequency ω_B of the frequency band of interest. These are in contrast to the CMS methodology where the response of the internal dofs due to boundary displacements is calculated using static condensation which neglects inertial effects^{10,13} and leads to the term $\mathbf{K}_{II}^{s-1} \mathbf{K}_{IB}^s$, involving the static stiffness matrices. In the mid-frequency range, it is important to account for the frequency-dependent inertial terms as inertial effects play a significant role here. It will be verified in a subsequent section that the approximation involved in calculating the submatrices $\bar{\mathbf{A}}_{II}^s$ and $\bar{\mathbf{A}}_{IB}^s$ only at the central frequency ω_B , can effectively reproduce the original system behavior for a sufficiently broad midfrequency band.

The equilibrium equation of the reduced system in the transformed coordinates is given by

$$\begin{aligned} \mathbf{R}^T \begin{bmatrix} \mathbf{A}_{BB}^s(\omega) & \mathbf{A}_{BI}^s(\omega) \\ \mathbf{A}_{IB}^s(\omega) & \mathbf{A}_{II}^s(\omega) \end{bmatrix} \mathbf{R} \begin{Bmatrix} \mathbf{U}_B^s \\ \mathbf{q}^s \end{Bmatrix} \\ = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\bar{\mathbf{A}}_{II}^{s-1} \bar{\mathbf{A}}_{IB}^s & \mathbf{\Sigma}^s \end{bmatrix}^T \begin{Bmatrix} \mathbf{F}_B^s \\ \mathbf{F}_I^s \end{Bmatrix}. \end{aligned} \quad (21)$$

Equation (21) expands to

$$\begin{bmatrix} \mathbf{A}_{BB}^s - \mathbf{A}_{BI}^s \bar{\mathbf{A}}_{II}^{s-1} \bar{\mathbf{A}}_{IB}^s - \bar{\mathbf{A}}_{BI}^s \bar{\mathbf{A}}_{II}^{s-1} \mathbf{A}_{IB}^s + \bar{\mathbf{A}}_{BI}^s \bar{\mathbf{A}}_{II}^{s-1} \mathbf{A}_{II}^s \bar{\mathbf{A}}_{II}^{s-1} \bar{\mathbf{A}}_{IB}^s \\ \mathbf{\Sigma}^{sT} (\mathbf{A}_{IB}^s - \mathbf{A}_{II}^s \bar{\mathbf{A}}_{II}^{s-1} \bar{\mathbf{A}}_{IB}^s) \end{bmatrix} \begin{Bmatrix} \mathbf{U}_B^s \\ \mathbf{q}^s \end{Bmatrix} = \begin{Bmatrix} \mathbf{F}_B^s - \bar{\mathbf{A}}_{BI}^s \bar{\mathbf{A}}_{II}^{s-1} \mathbf{F}_I^s \\ \mathbf{\Sigma}^{sT} \mathbf{F}_I^s \end{Bmatrix}. \quad (22)$$

Note that the reduced model has only $n_B + m$ dofs with $m \ll n_I$ compared to the original substructure dofs n_s . It is highlighted again that the approach presented here differs

from the CMS methodology mainly in two ways: (a) the proper orthogonal modes rather than the structural normal modes are used to reduce the interior dofs, (b) the inertial

effects are accounted for in order to obtain the response of the interior dofs due to the boundary motions in contrast to the CMS where a static condensation is adopted to achieve the same objective.

B. Substructure approach using free-free subsystem proper orthogonal modes

For a free-free substructuring approach, the transformation matrix of the subsystem is obtained as

$$\mathbf{R} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}^s \end{bmatrix}. \quad (23)$$

Consequently, the equilibrium equation of the reduced system in the transformed coordinates is given by

$$\begin{bmatrix} \mathbf{A}_{BB}^s & \mathbf{A}_{BI}^s \mathbf{\Sigma}^s \\ \mathbf{\Sigma}^{sT} \mathbf{A}_{IB}^s & \mathbf{\Sigma}^{sT} \mathbf{A}_{II}^s \mathbf{\Sigma}^s \end{bmatrix} \begin{Bmatrix} \mathbf{U}_B^s \\ \mathbf{q}^s \end{Bmatrix} = \begin{Bmatrix} \mathbf{F}_B^s \\ \mathbf{\Sigma}^{sT} \mathbf{F}_I^s \end{Bmatrix}. \quad (24)$$

The decision on using one or the other of these two boundary conditions will depend on the degree of coupling of a typical subsystem with the adjacent boundaries and other substructures as well as on the expediency from the viewpoint of numerical implementation. In the absence of any knowledge regarding the subsystem couplings, the relative merits of the fixed-fixed or free-free substructure methodologies are discussed in a subsequent section in the context of a numerical example.

V. SUBSYSTEM EIGENVALUE ANALYSIS FOR PROPER ORTHOGONAL DECOMPOSITIONS

In this section, the eigenvalue analysis for the proper orthogonal modes will be discussed both in the frequency and time domain. The relative merits of the time and frequency-domain approaches will be pointed out in the context of a system having a large number of dofs.

A. Frequency domain approach

The transformation matrix $\mathbf{\Sigma}^s$ is constructed from the dominant eigenvectors of the following eigenvalue problem:

$$\mathbf{C}_{\mathbf{U}^s \mathbf{U}^s} \Phi^s = \lambda^s \Phi^s. \quad (25)$$

Here $\mathbf{C}_{\mathbf{U}^s \mathbf{U}^s}$ is the time-averaged spatial correlation matrix obtained assuming fixed-fixed or free-free boundary conditions of the substructure. The correlation matrix is obtained from the frequency response of the subsystem sampled at M discrete frequencies as

$$\mathbf{C}_{\mathbf{U}^s \mathbf{U}^s} = \frac{1}{M} \sum_{j=1}^M \mathbf{U}_j^s(\omega_j) \mathbf{U}_j^{s*T}(\omega_j), \quad (26)$$

where $\mathbf{U}_j^s(\omega_j)$ denotes the frequency transform of the response vector of subsystem s at frequency ω_j . Using the frequency transfer function, and considering the response to a band-limited incoherent vector white noise excitation, an expression for the covariance matrix is obtained of the form,

$$\mathbf{C}_{\mathbf{U}^s \mathbf{U}^s} = \int_B \Re[\mathbf{A}^{s*} \mathbf{A}^{s-1}] d\omega. \quad (27)$$

Note that the methodology necessitates the inversion of the dynamic stiffness matrix at every frequency point. Evidently, the substructure approach is efficient when the number of substructure dofs is significantly smaller than that of the original system. The calculation of the correlation matrix from Eq. (27) is exact in the framework of linear random vibration theory. However, the computational burden arises in the frequency domain method while performing the inversion of the dynamic stiffness matrix at every discrete frequency when the dofs of the subsystem are considerably large. The time domain approach offers an appealing alternative that alleviates this computational burden.

B. Time domain approach

In the time domain, the expression of the correlation function is given by

$$\mathbf{C}_{\mathbf{u}^s \mathbf{u}^s} = \frac{1}{M} \sum_{j=1}^M \mathbf{u}_j^s(t_j) \mathbf{u}_j^{s*T}(t_j). \quad (28)$$

The response $\mathbf{u}^s(t)$ at M discrete time instants can be obtained by exciting the substructure by a sample or realization of a band-limited incoherent Gaussian white noise vector, $\mathbf{w}(t)$, and solving the corresponding equilibrium equation,

$$\mathbf{M}^s \frac{d^2 \mathbf{u}^s}{dt^2} + \mathbf{C}^s \frac{d \mathbf{u}^s}{dt} + \mathbf{K}^s \mathbf{u}^s = \mathbf{w}(t) \quad (29)$$

with the appropriate boundary conditions associated with either the fixed-fixed or the free-free substructure. For example, if a Houbolt integration scheme is adopted,²⁹ the solution vector $\mathbf{u}_{t+\Delta t}^s$ is obtained as the solution of the following algebraic matrix equation:

$$\begin{aligned} & \left(\frac{2}{\Delta t^2} \mathbf{M}^s + \frac{11}{6\Delta t} \mathbf{C}^s + \mathbf{K}^s \right) \mathbf{u}_{t+\Delta t}^s \\ &= \mathbf{w}_{t+\Delta t} + \left(\frac{5}{\Delta t^2} \mathbf{M}^s + \frac{3}{\Delta t} \mathbf{C}^s \right) \mathbf{u}_t^s - \left(\frac{4}{\Delta t^2} \mathbf{M}^s + \frac{3}{2\Delta t} \mathbf{C}^s \right) \\ & \times \mathbf{u}_{t-\Delta t}^s + \left(\frac{1}{\Delta t^2} \mathbf{M}^s + \frac{1}{3\Delta t} \mathbf{C}^s \right) \mathbf{u}_{t-2\Delta t}^s. \end{aligned} \quad (30)$$

The time-domain method is computationally less expensive compared to the frequency domain approach especially for subsystems with a large number of dofs as one has to invert the operator

$$\left[\frac{2}{\Delta t^2} \mathbf{M}^s + \frac{11}{6\Delta t} \mathbf{C}^s + \mathbf{K}^s \right]$$

only once. Note that a more efficient numerical scheme, specifically designed for the high-frequency response analysis,³⁰ can also be used to enhance the computational performance. Moreover, when direct inversion of a subsystem dynamic stiffness matrix in the frequency domain suffers from numerical instabilities,³¹ the time domain simulation can be conveniently used to obtain the correlation matrix for PODM. Despite the fact that an approximation is introduced in the time domain approach while constructing the spatial correlation matrix from a sample rather than an ensemble of random inputs, the approximation leads to acceptable results as will be verified through the numerical examples.

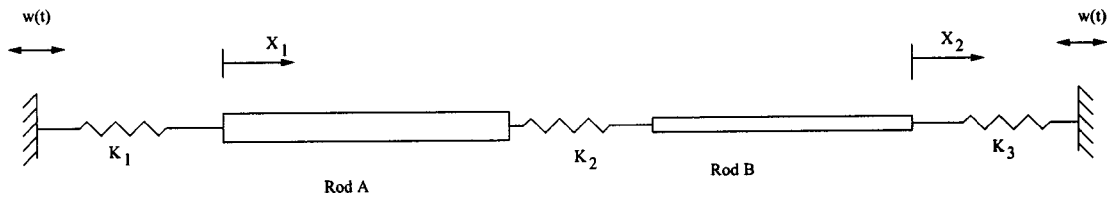


FIG. 1. Schematic diagram for the coupled rod.

VI. A REDUCED SUBSYSTEM EIGENVALUE ANALYSIS FOR PROPER ORTHOGONAL DECOMPOSITIONS: SNAPSHOT APPROACH

In the case where the data required to construct the correlation matrix consist of moderate length time histories or frequency samples with high spatial resolution ($n_s \gg M$), the snapshot method provides an alternative procedure for reducing the size of the original eigenvalue problem.²⁵ In this approach, the eigenvectors of the correlation matrix are expressed as a mixture of the snapshots of the response given by

$$\Phi_s = \sum_{j=1}^M \alpha_j \mathbf{u}^s(t_j). \quad (31)$$

Thus, the original eigenproblem in Eq. (25), of size n_s , is transformed into a reduced eigenproblem of size M . The resulting eigenproblem has the following form:

$$\mathbf{C}_t^s \alpha = \bar{\lambda}^s \alpha, \quad (32)$$

where the jk element of the reduced covariance matrix is given by

$$\mathbf{C}_{ijk}^s = \frac{1}{M} \mathbf{u}^{sT}(t_j) \mathbf{u}^s(t_k). \quad (33)$$

Evidently, the cost of numerical computation reduces significantly for $n_s \gg M$. The snapshot approach can also be formulated in the frequency domain in an analogous manner.

VII. GLOBAL SYSTEM SYNTHESIS FOR UNCERTAINTY ANALYSIS

In this section, a reduced model for the overall system is synthesized from models of the various subsystems which have been reduced in accordance with the proper orthogonal decomposition method. These models are integrated into a stochastic finite element framework as described in a previous section, by relying on the Karhunen–Loeve and polynomial chaos decomposition along with the orthogonal projection along chaos subspaces. The proper orthogonal modes obtained from the mean system properties of each substructure can be effectively used to construct the reduced subsystem model having random parameter variations in a given midfrequency band. After performing the assembly of all substructure system equations given in Eq. (22) or Eq. (24) using the traditional finite element formalism by matching the boundary dofs, the global stochastic system equation is written as indicated in Eq. (14). The random variables, $\{\xi_k\}$, are now associated with the random fluctuations in the global reduced matrix emerging from each substructure reduced dynamic stiffness. A characterization of the solution of this ran-

dom system equation can be readily obtained using an equation similar to Eq. (17) which can be solved for the coordinates on the chaos basis of the solution process. As is evident at this stage, the substructure-based reduced model has the following advantages over the reduced model obtained from the analysis of the complete system: (a) the substructure approach mitigates the computational burden associated with very large-scale systems of algebraic equations, (b) the reduced model of a given subsystem can be constructed independently and irrespective of the coupling strength and boundary condition of the given subsystem (in contrast to the case of obtaining the reduced model of the global system without substructure approach), (c) the methodology permits the combination of the reduced model constructed using traditional substructuring methodologies (based on modal analysis for the subsystems having low frequency behavior) and those obtained by proper orthogonal decomposition (obtained from substructure approach for the subsystem displaying the midfrequency behavior), or other model reduction methodologies.

VIII. NUMERICAL RESULTS

For the numerical investigation of the mathematical formulations presented in the previous sections, we consider a spring-supported coupled rod assembly interconnected by an intermediate spring as shown in Fig. 1. This simple system provides insight into the formulations without the undue intrinsic complexity associated with the dynamics of a more complex system. The differential equation governing the motion of a typical rod is given by

$$\frac{\partial}{\partial x} \left[EA(x) \frac{\partial u}{\partial x} + C_1(x) \frac{\partial^2 u}{\partial t \partial x} \right] = m(x) \frac{\partial^2 u}{\partial t^2} + C_2(x) \frac{\partial u}{\partial t}, \quad (34)$$

with appropriate initial and boundary conditions. Here EA , m , C_1 , and C_2 are the stiffness, mass, strain-rate dependent, and velocity-dependent damping per unit length of the rod system.

The following properties of the rods are considered for the numerical investigations: for rod A, the mean $EA_A = 405\,284.712$ N; $m_A = 200.0$ kg; $C_{1A} = 100.0$ Ns and $C_{2A} = 100.0$ Ns/m², $L_A = 100$ m; for rod B, the mean $EA_B = EA_A/10$; $m_B = m_A$; $C_{1B} = C_{1A}/10$, $C_{2B} = C_{2A}/10$, and $L_B = L_A$. The number of the finite elements for rods A and B are taken to be 100 and 200, respectively. This case simulates a situation of one subsystem exhibiting low frequency behavior (rod A) with the other (rod B) showing medium frequency features.

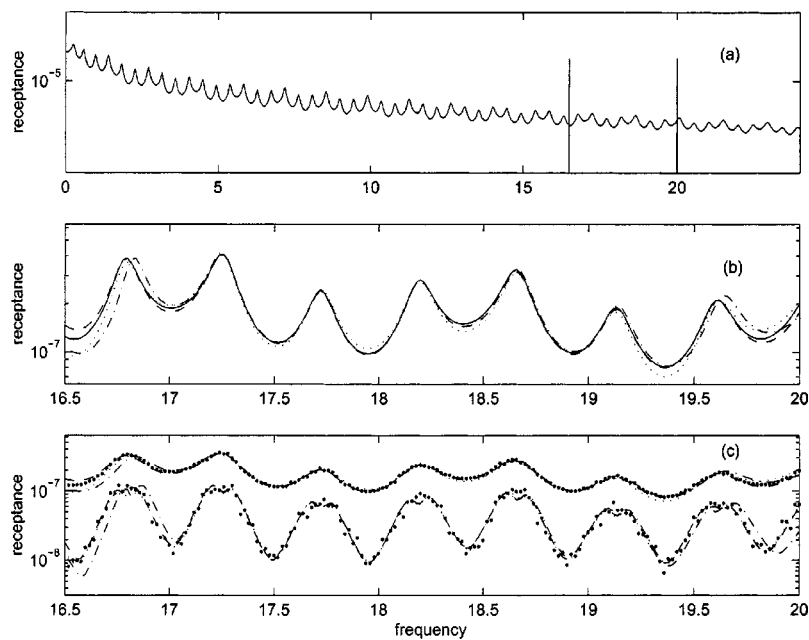


FIG. 2. Receptance function for the case of weak subsystem couplings: (a) reference solution of deterministic system; (b) (—) reference solution, reduced models: (---) without substructure approach, (---) fixed-fixed substructure approach, (···) free-free substructure approach; (c) mean and standard deviation of receptance function of stochastic system: (·) Monte Carlo simulation results on full system; reduced stochastic finite element results: (---) without substructure approach, (---) fixed-fixed substructure approach and (···) free-free substructure approach.

The random uncertainty is considered only in rod B as the higher frequency range dynamics are sensitive to system parameter variations. The stiffness of rod B is modeled by a Gaussian random process with a coefficient of variation of 5%. The autocovariance function of the process is chosen to be of the form, $R(x,y) = e^{-|x-y|/b}$ where b denotes the correlation length, assumed to be equal to half of the length of the rod. Only two terms in the Karhunen–Loeve expansion are used. In the polynomial chaos expansion, the first six terms corresponding to the complete second-order expansion are considered. It is worthwhile pointing out that the coefficient of variation should be interpreted relative to the structural damping and frequency band of interest. Small varia-

tions in the system parameter can significantly alter the system performance when the driving force excites the medium and high frequency dynamics of the system. This is not surprising as the short wavelength vibration relating to the medium and high frequency responses are sensitive to parameter variation, in contrast to the long wavelength oscillation involving low frequency response which exhibit minor influence to such perturbation in the system parameter.

Figures 2 and 3 present the result for the case of weak couplings with $K_1 = K_2 = K_3 = 1000$ N. The reference cross-receptance function between X_1 and X_2 (see Fig. 1) is plotted in Fig. 2(a). For the deterministic case, the receptance functions from the reduced models constructed with and without

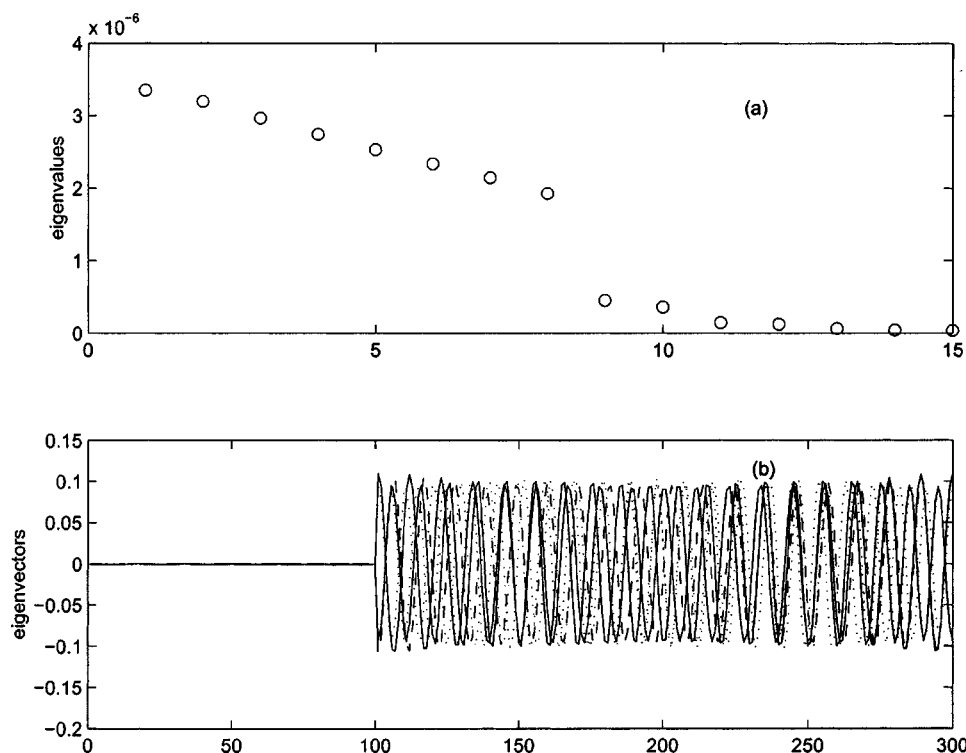


FIG. 3. Eigensubspace of correlation matrix of full system without substructure approach for the case of weak subsystem couplings: (a) distribution of eigenvalues, (b) first few dominant eigenvectors.

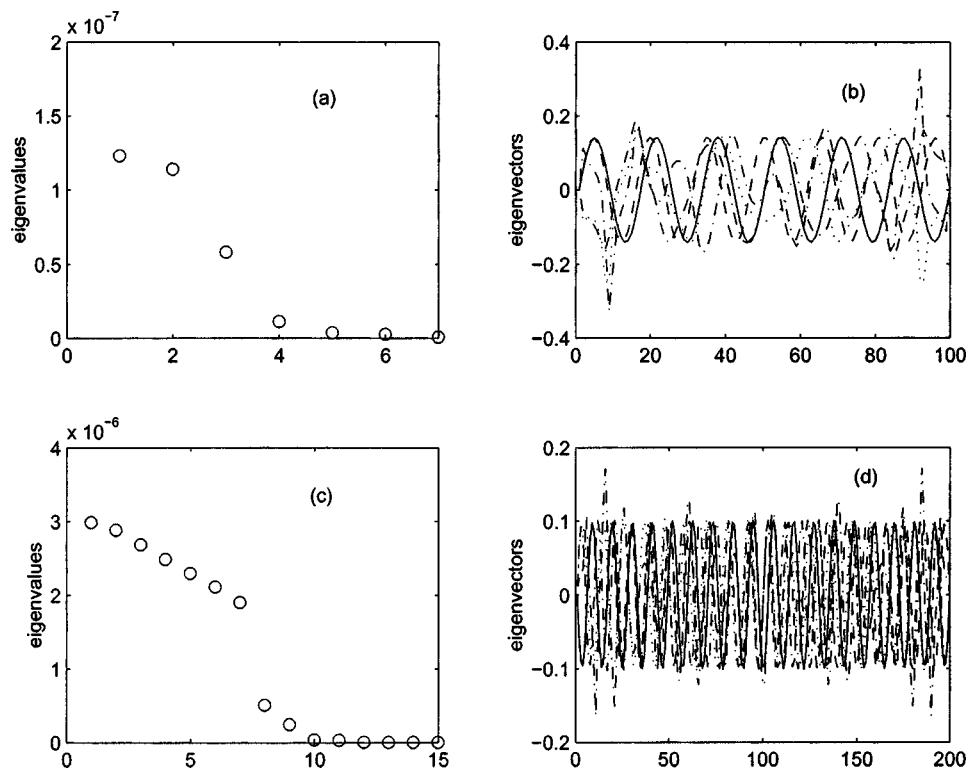


FIG. 4. Eigensubspace of correlation matrix of fixed-fixed substructure approach. For rod A: (a) distribution of eigenvalues, (b) first few dominant eigenvectors; for rod B: (c) distribution of eigenvalues, (d) first few dominant eigenvectors.

the substructure approach are plotted in Fig. 2(b) for the frequency band 16.5–20 rad/s. The receptance functions constructed with and without substructuring show good agreement with the reference solution. In case of substructuring, the result from the fixed-fixed substructure approach appears to be superior to that obtained using the free-free substructure method although the difference is not very significant.

The results for the stochastic system are plotted in Fig. 2(c). The line plots refer to the predictions using the Karhunen-Loeve and chaos projections, while the superimposed point plots show the results associated with a Monte Carlo solution, used to verify the chaos methodology. The top and bottom plots represent the mean and standard deviation, respectively, of the receptance function. The mean and standard deviation of the receptance functions obtained with and without the substructure approach are also plotted which again show good agreement with the Monte Carlo simulation results. These were obtained using the chaos decomposition approach.

Figure 3 shows the dominant eigenspace of the correlation functions obtained by analyzing the full system without substructure approach. The spatial correlation functions are constructed using the frequency domain approach. The eigenvalues of the correlation matrix are plotted in Fig. 3(a). Clearly, a 12-dof reduced model can reproduce the system dynamics of the full 300-dof system. A few eigenvectors corresponding to the dominant eigenvalues of the correlation matrix are plotted in Fig. 3(b). The fluctuations of the eigenvectors in rod A are not visible in the scale of the graph as their amplitude is much smaller compared to those for rod B which displays, as expected, much more rapid fluctuations.

Figures 4 and 5 show the eigensubspace of the correlation matrices for the fixed-fixed and free-free substructure approaches. From both Figs. 4(a) and (c) and 5(a) and (c),

the reduced models of the rod A and rod B having 100 and 200 dofs can be constructed with only 5 and 10 dofs, respectively, leading to a 15 dofs model of the global reduced model. While a 12 dofs reduced model without substructuring can capture the system behavior, the 15 dofs reduced model obtained the substructure approach is somewhat more efficient from the computational point of view. Of course the significance of this efficiency is expected to increase with the scale of the problem, and in particular as the size of each substructure relative to that of the full system decreases.

Figures 6 and 7 present the result for the case of strong coupling with $K_1 = K_2 = K_3 = 100\,000$ N. In Fig. 6(a), the reference cross-receptance function is plotted. The receptance functions from the reduced models constructed with and without the substructure approach are plotted in Fig. 6(b) for the frequency band 16.5–20 rad/s for the deterministic system. The receptance functions constructed using fixed-fixed substructure approach and that without substructure approach show better agreement with the reference solution than that obtained from the free-free substructure method. In fact, the free-free substructure approach shows unacceptably poor performance. This is not surprising as the subsystems are strongly coupled. Interestingly, the result from the fixed-fixed substructure displays remarkable agreement with the reference solution with the two curves being nearly superimposed in Fig. 6(b).

Similar to Fig. 2(c), the results for the stochastic system are shown in Fig. 6(c). The mean and standard deviations of the receptance functions obtained with fixed-fixed substructure and without the substructure approach display good agreement with the Monte Carlo simulations while the performance of the free-free substructure approach is poor, as expected. In Fig. 7, the eigensubspace of the correlation matrix obtained from the full scale analysis (i.e., without sub-

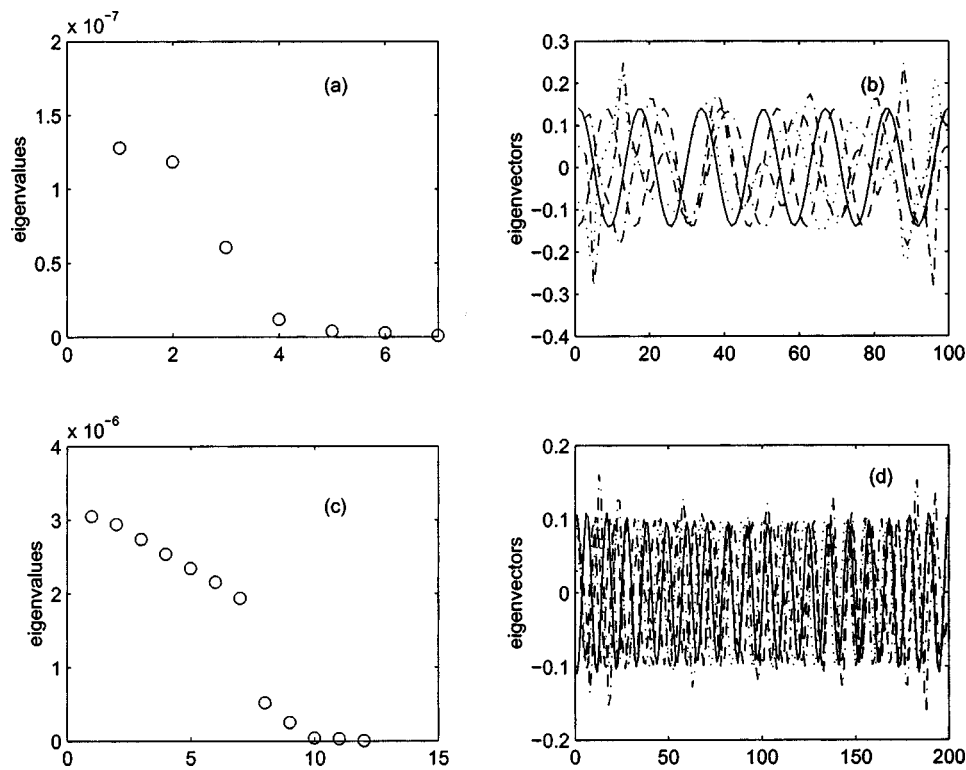


FIG. 5. Eigensubspace of correlation matrix of free-free substructure approach. For rod A: (a) distribution of eigenvalues, (b) first few dominant eigenvectors; for rod B: (c) distribution of eigenvalues, (d) first few dominant eigenvectors.

structure approach) is plotted. It can again be identified from Fig. 7(a) that a 12 dofs reduced model can capture the behavior of the full scale system with 300 dofs. The eigenvectors in Fig. 7(b) display a much different pattern than the previous example due to the strong coupling. It should be noted that the eigensubspace of the subsystems for the substructure analysis remains the same as in Figs. 3 and 4 as they are independent of the coupling parameters. Thus, once

the reduced substructure models are constructed, they can be used repeatedly for global system assembly for different values of coupling parameters. This identifies another important feature of the substructure approach.

Next, the results obtained from the correlation matrix constructed in the time domain approach are presented in Fig. 8 for the case of strong subsystem coupling. Only the case of fixed-fixed substructure approach will be considered

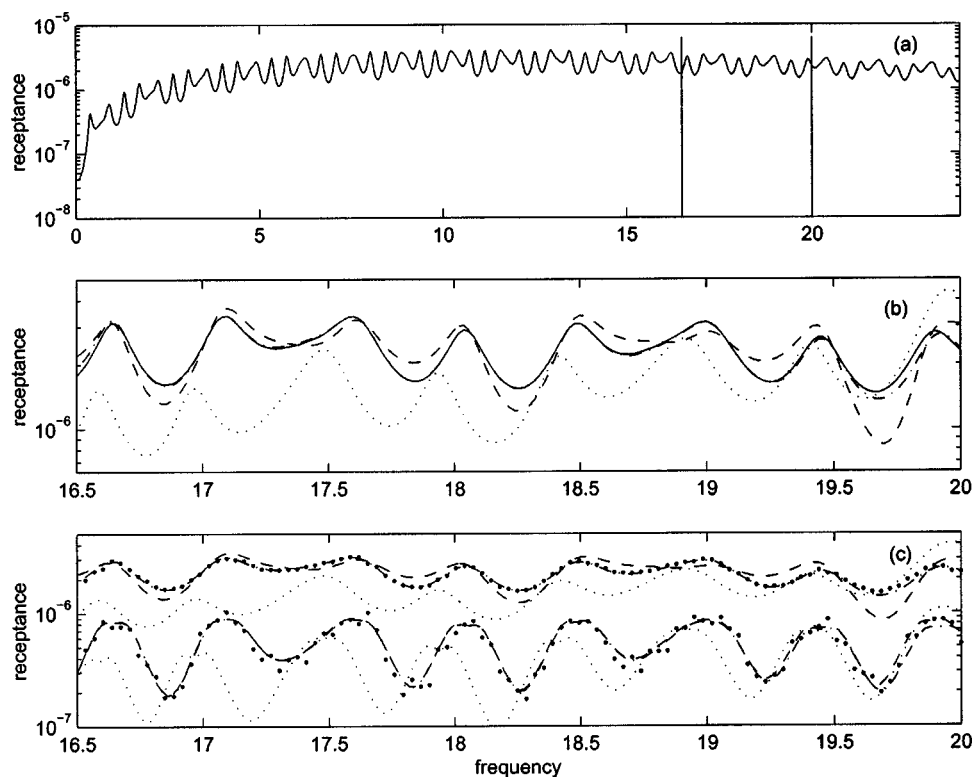


FIG. 6. Receptance function for the case of strong subsystem couplings: (a) reference solution of deterministic system; (b) (—) reference solution, reduced models: (---) without substructure approach, (---) fixed-fixed substructure approach, (---) free-free substructure approach; (c) mean and standard deviation of receptance function of stochastic system: (·) Monte Carlo simulation results on full system; reduced stochastic finite element results: (---) without substructure approach, (---) fixed-fixed substructure approach and (---) free-free substructure approach.

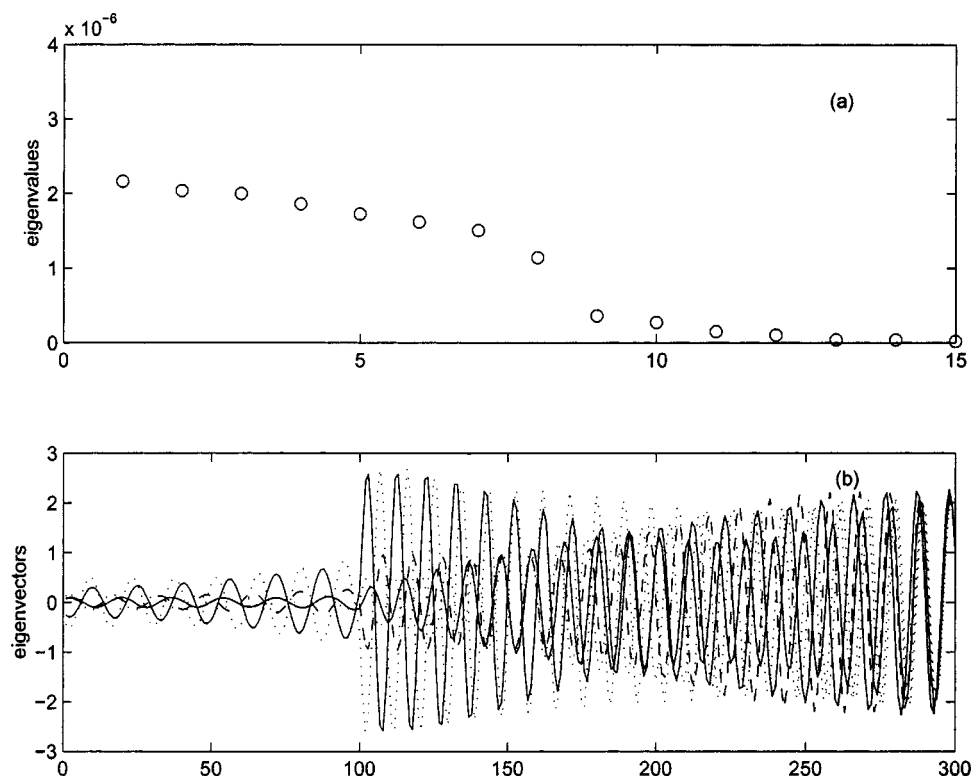


FIG. 7. Eigensubspace of correlation matrix of full system without substructure approach for the case of strong subsystem couplings: (a) distribution of eigenvalues, (b) first few dominant eigenvectors.

as its superiority has already been demonstrated in the case of both weak and strong subsystem coupling. Figure 8(a) presents the receptance function for the deterministic system where the result obtained from reduced models without substructure approach and with the fixed-fixed substructure method are plotted against the reference solution. Again, the results from the fixed-fixed substructure method show excellent agreement with the reference solution whereas the performance of the reduced model obtained without substructure approach is uneven.

The mean and standard deviation of the receptance function are plotted in Fig. 8(b). As expected, the results from the fixed-fixed substructure approach display excellent agreement with Monte Carlo simulations.

In Fig. 9, the dominant eigensubspace of the correlation function obtained by the full-scale system without substructure approach is plotted. In contrast to the case in Fig. 7(a)

obtained from the frequency domain analysis, a 14 dof reduced model is necessary to reproduce the dynamics for the full system with 300 dofs as evident from Fig. 9(a) displaying the eigenvalue distributions of the correlation matrix. A few eigenvectors corresponding to the dominant eigenvalues of the correlation matrix are plotted in Fig. 9(b) which differ from those shown in Fig. 7(b). The dominant eigensubspace of the subsystem correlation matrices are plotted in Fig. 10 which again display differences with the corresponding frequency domain results in Fig. 4. These differences are of course expected because of the approximation involved in the time domain approach due to approximating the ensemble averaging operation as explained before. Notwithstanding the differences in the POMs in the frequency- and time-domain methods, the receptance functions obtained

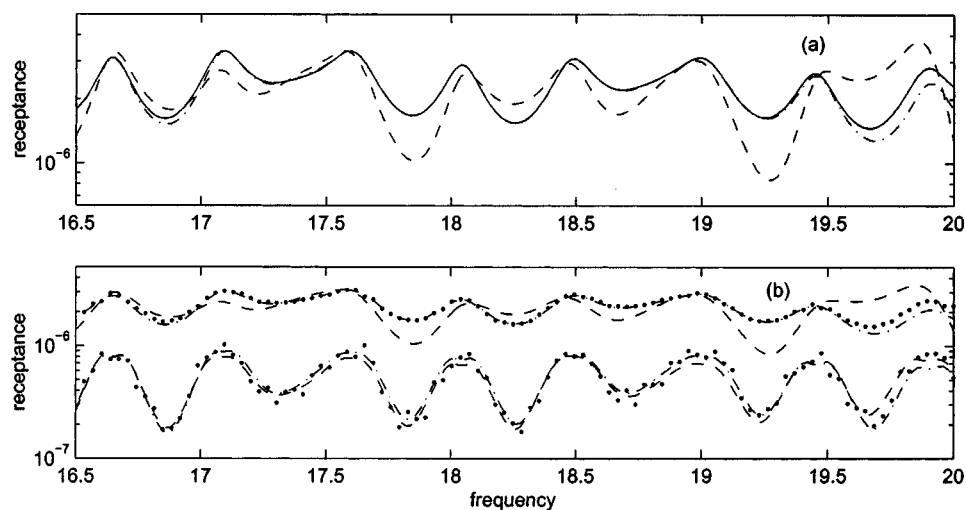


FIG. 8. Substructure approach in *time domain*: receptance function for the case of strong subsystem couplings: receptance function for deterministic systems, (a) (—) reference solution; reduced model: (---) without substructure approach, (---) fixed-fixed substructure approach, (b) mean and standard deviation of receptance function of stochastic system: (···) Monte Carlo simulation results on full system; reduced stochastic finite element results: (---) without substructure approach, and (---) fixed-fixed substructure approach.

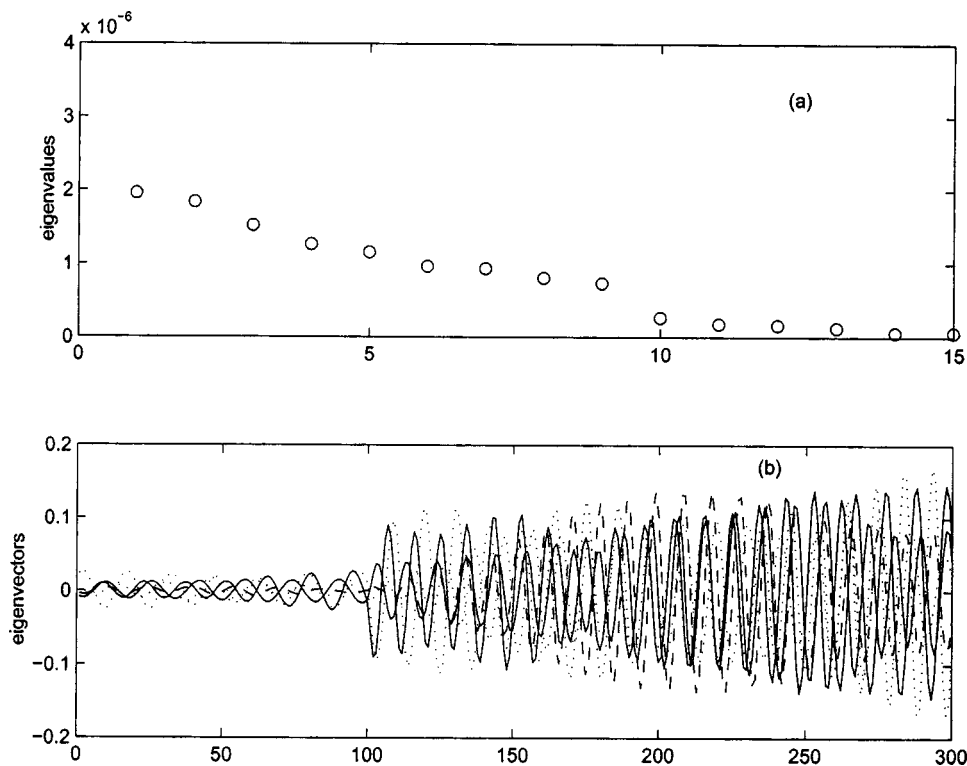


FIG. 9. Substructure approach in *time domain*: eigensubspace of correlation matrix of full system without substructure approach for the case of strong subsystem couplings: (a) distribution of eigenvalues, (b) first few dominant eigenvectors.

from the reduced models using these two approaches show very good agreement with the reference solution in Figs. 6 and 8. This fact can be elegantly exploited for the numerical calculations of the proposed substructure approach in the time domain for the uncertainty analysis of large-scale structures in the midfrequency range.

IX. CONCLUSIONS

A substructure approach is presented for the midfrequency structural dynamics of randomly parametered sys-

tems using the proper orthogonal decomposition. At least for the cases investigated in this paper, the proposed approach based on fixed-fixed proper orthogonal modes, which bears analogy with the popular Craig and Bampton component mode synthesis, exhibits superior performance in comparison to the case when the reduced substructure model is obtained from the free-free substructure proper orthogonal modes. The difference between the present methodology and the traditional component mode synthesis approach for low-

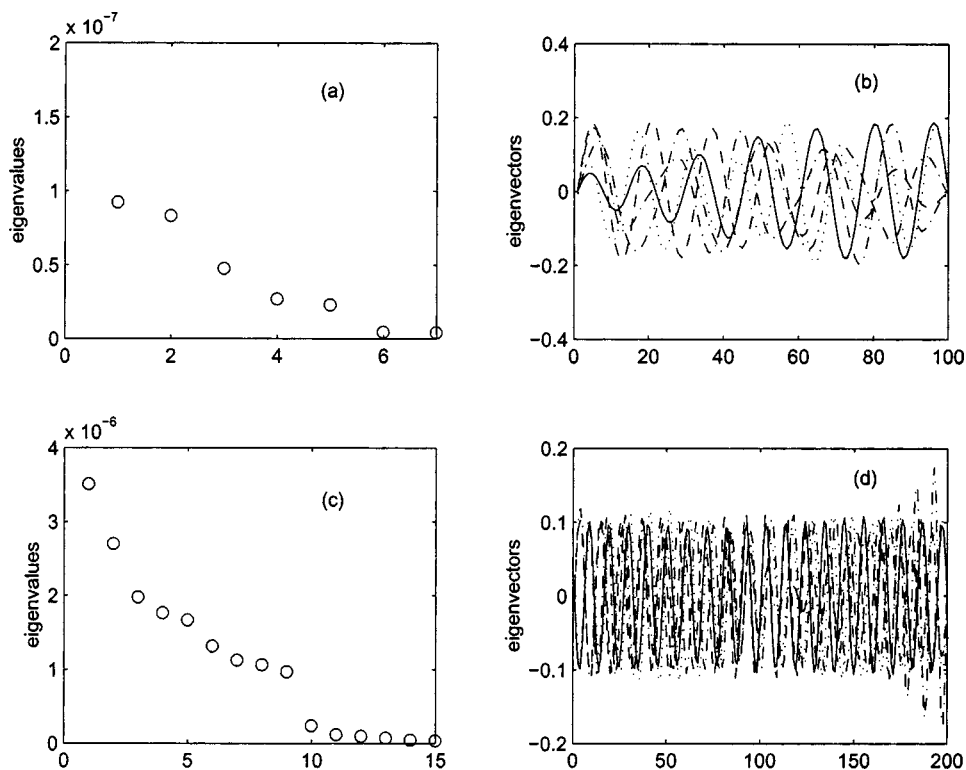


FIG. 10. Substructure approach in *time domain*: eigensubspace of correlation matrix of fixed-fixed substructure approach. For rod A: (a) distribution of eigenvalues, (b) first few dominant eigenvectors; for rod B: (c) distribution of eigenvalues, (d) first few dominant eigenvectors.

frequency range vibration has also been pointed out. It consists mainly in accounting for the inertial effect in performing the substructuring, as well as using the response itself to guide the development of a basis for the coordinate transformation. A further reduction strategy of the eigenvalue problem arising in the proper orthogonal mode decomposition method, suitable for subsystems having a considerably large number of dofs, has also been described and referred to as the *snapshot* approach. The substructure approach can significantly reduce the cost of numerical computations for mid-frequency vibration analysis by avoiding to deal with large scale global system matrices both from the standpoint of deterministic and probabilistic system analysis. Once the reduced model of the substructure is constructed, the probabilistic analysis can also be easily carried out with minimal numerical cost.

ACKNOWLEDGMENT

The authors wish to acknowledge the financial support of the Office of Naval Research under Grant No. N000149910900.

- ¹S. Adhikari and C. S. Manohar, "Dynamic analysis of framed structures with statistical uncertainties," *Int. J. Numer. Methods Eng.* **44**, 1157–1178 (1999).
- ²C. S. Manohar and S. Adhikari, "Dynamic stiffness of randomly parametered beams," *Probab. Eng. Mech.* **13**, 39–49 (1998).
- ³C. Soize, "Reduced models in the medium-frequency range for general external structural-acoustic systems," *J. Acoust. Soc. Am.* **103**, 3393–3406 (1998b).
- ⁴C. Soize, "Reduced models for structures in the medium-frequency range coupled with internal acoustic cavities," *J. Acoust. Soc. Am.* **106**, 3362–3374 (1999).
- ⁵C. Soize, "A nonparametric model of random uncertainties for reduced matrix models in structural dynamics," *Probab. Eng. Mech.* **15**, 277–294 (2000).
- ⁶R. Ohayon and C. Soize, *Structural Acoustics and Vibration* (Academic, San Diego, 1998).
- ⁷R. Ghanem and A. Sarkar, "The medium-frequency range vibration of coupled fluid-structure assembly with system uncertainty," *The 14th ASCE Special. Conf. Eng. Mech.*, The University of Texas, Austin, 22–24 May 2000.
- ⁸R. S. Langley and P. Bremner, "A hybrid method for the vibration analysis of complex structural-acoustic systems," *J. Acoust. Soc. Am.* **105**, 1657–1671 (1998).
- ⁹N. Vlahopoulos and X. Zhao, "Basic development of hybrid finite element method for midfrequency structural vibrations," *AIAA J.* **37**, 1495–1505 (1999).
- ¹⁰M. Geradin and D. Rixen, *Mechanical Vibration* (Wiley, Chichester, 1994).
- ¹¹A. Curnier, "On the three modal synthesis variants," *J. Sound Vib.* **90**, 527–540 (1983).
- ¹²R. R. Craig, Jr. and M. C. C. Bampton, "Coupling of substructures for dynamic analysis," *AIAA J.* **6**, 1313–1319 (1968).
- ¹³C. Soize and S. Mziou, "Dynamic substructuring in the medium-frequency range," *ISMA 25, International Conference on Noise and Vibration Engineering*, Leuven (Belgium), 2000.
- ¹⁴A. Sarkar and R. Ghanem, "Mid-frequency structural dynamics with parameter uncertainty," *42nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Seattle, 2001.
- ¹⁵A. Sarkar and R. Ghanem, "Mid-frequency structural dynamics with parameter uncertainty," *Comput. Methods Appl. Mech. Eng.* **191**, 5499–5513 (2002).
- ¹⁶M. Loeve, *Probability Theory*, 4th ed. (Springer, New York, 1977).
- ¹⁷R. H. Cameron and W. T. Martin, "The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals," *Ann. Math.* **48**, 385–392 (1947).
- ¹⁸N. Wiener, "The homogeneous chaos," *Am. J. Math.* **60**, 897–936 (1938).
- ¹⁹G. Kallianpur, *Stochastic Filtering Theory* (Springer, Berlin, 1980).
- ²⁰J. L. Lumley, "The structure of inhomogeneous turbulent flows," *Atmospheric Turbulence and Radio Wave Propagation* (Nauka, Moscow, 1967), pp. 166–178.
- ²¹X. Ma and A. F. Vakakis, "Nonlinear transient localization and beat phenomena due to backlash in a coupled flexible systems," *J. Vib. Acoust.* **123**, 36–44 (2001).
- ²²L. Sirovich, "Turbulence and dynamics of coherent structures. I. Coherent structures," *Q. Appl. Math.* **45**, 561–571 (1987a).
- ²³L. Sirovich, "Turbulence and dynamics of coherent structures. II. Symmetries and transformations," *Q. Appl. Math.* **45**, 573–582 (1987b).
- ²⁴L. Sirovich, "Turbulence and dynamics of coherent structures. III. Dynamics and scaling," *Q. Appl. Math.* **45**, 583–590 (1987c).
- ²⁵S. B. Breuer and L. Sirovich, "The use of Karhunen–Loeve procedure for the calculation of linear eigenfunctions," *J. Comput. Phys.* **96**, 277–296 (1991).
- ²⁶A. Sarkar and M. P. Paidoussis, "Nonlinear dynamics of cantilever conveying fluid: Coherent modes versus beam modes," *Int. J. Non-linear Mech.* (to be published).
- ²⁷A. Sarkar and M. P. Paidoussis, "A compact limit cycle oscillation (LCO) model for a cantilevered fluid-conveying pipe," *J. Fluids Struct.* (to be published).
- ²⁸K. J. Bathe and E. L. Wilson, *Numerical Methods in Finite Element Analysis* (Prentice–Hall, Englewood Cliffs, NJ, 1976).
- ²⁹R. Ghanem and P. Spanos, *Stochastic Finite Elements: A Spectral Approach* (Springer, Berlin, 1991).
- ³⁰C. Soize, P. M. Hutin, A. Desanti, J. M. David, and F. Chabas, "Linear dynamic analysis of mechanical systems in the medium frequency range," *Comput. Struct.* **23**, 605–637 (1986).
- ³¹T. P. Gialamas, D. T. Tsahalis, D. Otte, H. Van der Auwaraer, and D. A. Manolas, "Substructuring technique: Improvement by means of singular value decomposition (SVD)," *Appl. Acoust.* **62**, 1211–1219 (2001).

Modal characteristics of in-plane vibration of circular plates clamped at the outer edge

N. H. Farag and J. Pan

School of Mechanical Engineering, The University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia

(Received 11 January 2002; revised 13 December 2002; accepted 16 December 2002)

The equations of in-plane vibration in thin flat plates are solved for free vibration in circular plates clamped at the outer edge. The mode shapes are represented by trigonometric functions in the circumferential direction and by series summation of Bessel functions in the radial direction. Accuracy of the predictions of natural frequencies and mode shapes is assessed by comparisons with finite-element predictions and with previously reported results. The present solution gives very accurate predictions. The work also highlights the nature of coupling between the different circumferential and radial modes and the response of different vibrational modes at the center of the plate. It is shown that the center point of the plate vibrates only for modes with unity circumferential wave number (number of nodal diameters). Nondimensional frequency parameters are listed and the radial mode shapes of natural vibration are depicted to illustrate the free-vibration behavior in the frequency range of practical interest. © 2003 Acoustical Society of America.

[DOI: 10.1121/1.1553456]

PACS numbers: 43.40.Dx [ANN]

LIST OF SYMBOLS AND BASIC RELATIONS

E	Young's modulus of elasticity	f	(torsional) directions, respectively frequency (Hz)
h	plate thickness	ω	angular frequency (radians per second)
j	the imaginary number $\sqrt{-1}$	ν	Poisson's ratio
m	circumferential wave number (equals number of nodal diameters)	ρ	mass density
$N_{rr}, N_{\theta\theta}, N_{r\theta}$	in-plane force components (force per unit length) in the directions shown in Fig. 1	$C_L^2 = E/\rho(1 - \nu^2)$	square of the quasilongitudinal wave speed in the plate
\bar{r}	plate radius	$C_T^2 = E/2\rho(1 + \nu)$	square of the in-plane shear wave speed in the plate
u, v	components of in-plane displacement in radial (extensional) and circumferential	$K_L = \omega\bar{r}/C_L$	nondimensional frequency parameter

I. INTRODUCTION

Circular plates exist in many engineering applications. In particular, they form partitions in aircraft fuselages and external fuel tanks and end plates in storage tanks. Some of these applications are subjected to dynamic excitations with large components lying in the middle plane of the plate. The in-plane dynamic loads, being continuous, intermittent, or even impacts, will excite in-plane modes with resonance frequencies in the frequency bands of excitation.

Investigation of the free in-plane vibrational response in circular thin flat plates with clamped edge is the subject of this work. The main objective is to provide the structural dynamics analyst with a quick and easy tool to predict in-plane natural frequencies and depict their mode shapes, and to reveal some important aspects of modal characteristics of in-plane vibration.

While much work has been done on the investigation and documentation of the natural frequencies and modes shapes for flexural vibration (e.g., Refs. 1 and 2), the characteristics of in-plane vibration have not been completely

investigated and are not well documented. However, increasing attention has been given to the in-plane vibration in single and complex plate and plate-like structures in the past few years (e.g., Refs. 3–7). The results of recent investigations emphasize the importance of the in-plane response at high frequencies and in large coupled plate-like structures.

Predictions of the natural frequencies of circular plates are treated in few references. In particular, the problems of in-plane vibration and stability of rotating disks have attracted attention of researchers for decades due to their obvious practical importance in many engineering applications. Also, the in-plane vibrational response was investigated for piezoelectric disks and computer disks. These investigations have been dealing with circular and annular thin or thick plates. Free-boundary conditions at the plate edge(s) were considered in most of the published work.

Holland⁸ investigated the free in-plane vibration in circular plates with free edges using trigonometric and Bessel functions and published frequency parameters for five to ten modes of the first nine circumferential modes (circumferential wave numbers $m = 1$ to 9) for different values of Pois-

son's ratio. Ambati *et al.*⁹ examined the in-plane vibrations in circular and annular plates, and rings with free boundaries. Irie *et al.*¹⁰ used a transfer matrix formulation to solve the equations of free in-plane vibrations for the natural frequencies of annular plates with combinations of free and clamped conditions at the inner and outer edges. In Ref. 10, the circular plate was considered as a limiting case when the inner diameter tends to zero. For this specific case of circular plates, the frequency parameters (nondimensional natural frequencies) are listed for the first two radial modes of five circumferential wave numbers ($m=0$ to 4). The mode shapes were not examined in Ref. 10. The first eight references cited in Ref. 10 represent the bulk of the work done on in-plane free vibration of annular and circular thin plates up to 1984. Chen *et al.*¹¹ presented displacement potentials' formulation to examine the effect of the angular velocity of spinning disks on the in-plane vibrations and natural frequencies.

To conclude, it can be seen that limited work is reported on the in-plane vibration of circular clamped plates. This type of structure was only considered as a limited case of annular plate when the inner radius goes to zero¹⁰ and the radial mode shapes of in-plane vibration have not been reported.

The present work is a comprehensive investigation of the modal characteristics of in-plane vibration in circular plates with clamped edge (i.e., rigidly restrained edge in the plane of the plate). It concentrates on the following new areas: (i) The circular plate is investigated directly not as a special case of annular plate when the inner radius goes to zero. (ii) The work clarifies the modal response at the plate center and the nature of the coupling between different circumferential modes and different radial modes. (iii) The mode shapes of in-plane vibration in the radial direction are examined in detail. (iv) Nondimensional frequency parameters and mode shapes are tabulated. (v) Elements of the characteristic equation are presented, which can be used to compute the natural frequencies and mode shapes for other material properties or higher frequency ranges.

In this paper, the equations for in-plane free vibration are presented in polar coordinates suitable for the type of structure under investigation. Assumed mode shapes are expressed in terms of trigonometric functions in the circumferential direction. The nature of coupling between the circumferential modes is investigated. The physical behavior of the center point of the plate, during free vibration in each of the circumferential modes, is examined. Mode shapes in the radial direction are assumed as series summation of Bessel functions. The mathematical model for free vibration is written in the form of an eigenvalue problem so that natural frequencies and modes shapes can be obtained by solving for eigenvalues and eigenvectors employing any available mathematical software. The frequency parameters obtained by the present method are tabulated and the mode shapes are depicted to illustrate the free-vibration behavior in the frequency range of practical interest. Comparisons with finite-element results and with the previously reported results confirm the accuracy of the predictions of the present

method. The effect of Poisson's ratio on the natural frequencies is also examined.

The mathematical formulations are presented in Sec. II followed by computational examples and discussions in Sec. III. Tables I and II include a tabulation of the frequency parameters and mode shape functions for in-plane vibration in circular clamped plates.

II. MATHEMATICAL SOLUTIONS FOR FREE IN-PLANE VIBRATION IN CIRCULAR CLAMPED PLATES

In this section the equations of free in-plane vibration, in polar coordinates, are presented. The assumed solution represents the circumferential distribution of in-plane vibration by cosine/sine functions and the radial distribution by series summation of Bessel functions. The coupling between the circumferential modes is investigated, as well as their response at the center point of the circular plate. Suitable mathematical forms of the radial mode shapes are used accordingly. The assumed modal response is substituted into the equations of motion. After mathematical manipulation the equations are presented in the form of eigenvalue problems suitable for solution for the nondimensional frequency parameters and mode shapes.

A. Equations of motion

The equations governing free in-plane vibration of thin flat plates, in polar coordinates, can be found in many references (e.g., Refs. 10 and 11). They may be written in the following form:

$$\frac{\partial^2 u}{\partial t^2} - C_L^2 \left[\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} - \frac{u}{r^2} \right] - C_T^2 \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} - C_T^2 \frac{1}{r} \frac{1+v}{1-v} \frac{\partial^2 v}{\partial r \partial \theta} + C_T^2 \frac{1}{r^2} \frac{3-v}{1-v} \frac{\partial v}{\partial \theta} = 0, \quad (1a)$$

$$\frac{\partial^2 v}{\partial t^2} - C_T^2 \left[\frac{\partial^2 v}{\partial r^2} + \frac{1}{r} \frac{\partial v}{\partial r} - \frac{v}{r^2} \right] - C_L^2 \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} - C_T^2 \frac{1}{r} \frac{1+v}{1-v} \frac{\partial^2 u}{\partial r \partial \theta} - C_T^2 \frac{1}{r^2} \frac{3-v}{1-v} \frac{\partial u}{\partial \theta} = 0. \quad (1b)$$

Please see the list of symbols at the beginning of the paper. The above equations are based on the assumptions of thin plate theory¹ and plane stress conditions.⁸ The positive directions of displacement and internal force components are illustrated in Fig. 1.

The in-plane internal force intensities (force per unit length perpendicular to the force direction) are expressed in terms of displacements by the following relations:¹⁰

$$N_r = \frac{-Eh}{1-\nu^2} \left(\frac{\partial u}{\partial r} + \frac{\nu}{r} \frac{\partial v}{\partial \theta} + \frac{\nu}{r} u \right), \quad (2a)$$

$$N_\theta = \frac{-Eh}{1-\nu^2} \left(\frac{1}{r} \frac{\partial v}{\partial \theta} + \frac{u}{r} + \nu \frac{\partial u}{\partial r} \right), \quad (2b)$$

$$N_{r\theta} = \frac{-Eh}{2(1+\nu)} \left(\frac{1}{r} \frac{\partial u}{\partial \theta} + \frac{\partial v}{\partial r} - \frac{v}{r} \right). \quad (2c)$$

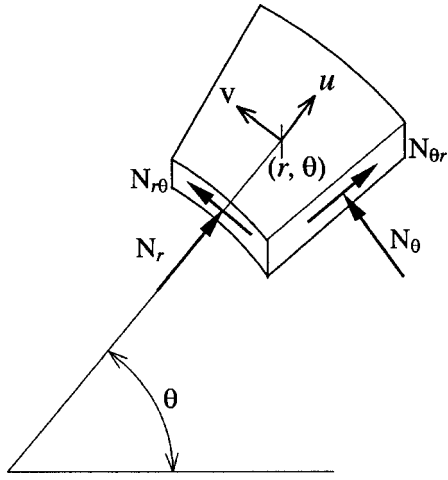


FIG. 1. Positive directions of in-plane force and displacement components on an infinitesimal element of a circular plate.

For harmonic vibration with time dependence $e^{-j\omega t}$, the free in-plane vibrational response is assumed in the form

$$u(r, \theta, t) = \sum_{m_1=0}^{\infty} U_{m_1}(r) \cos(m_1 \theta) e^{-j\omega t}, \quad (3a)$$

$$v(r, \theta, t) = \sum_{m_2=0}^{\infty} V_{m_2}(r) \sin(m_2 \theta + \varphi) e^{-j\omega t}. \quad (3b)$$

The angle φ is introduced in Eq. (3b) to accommodate a nonzero response $v(r, \theta, t)$ for the case of $m_2=0$. Therefore, $\varphi=\pi/2$ for $m_2=0$ and $\varphi=0$ for all other values of m_2 .

The following equations are obtained after substitution of the assumed solution from Eqs. (3) into Eqs. (1), multiplying the resulting two equations by $\cos(m_1 \theta)$ and $\sin(m_2 \theta + \varphi)$, respectively, integrating with respect to θ from $\theta=0$ to $\theta=2\pi$, and employing the orthogonal properties of the trigonometric functions

$$\begin{aligned} C_L^2 \frac{\partial^2 U_{m_1}}{\partial r^2} + C_L^2 \frac{1}{r} \frac{\partial U_{m_1}}{\partial r} + \left(\omega^2 - \frac{C_L^2 + m_1^2 C_T^2}{r^2} \right) U_{m_1} \\ + C_T^2 \frac{1+v}{1-v} \frac{1}{r} \frac{1}{\pi} \sum_{m_2=0}^{\infty} I_{m_2, m_1} \frac{\partial V_{m_2}}{\partial r} \\ - C_T^2 \frac{3-v}{1-v} \frac{1}{r^2} \frac{1}{\pi} \sum_{m_2=0}^{\infty} I_{m_2, m_1} V_{m_2} = 0, \end{aligned} \quad (4a)$$

$$\begin{aligned} C_T^2 \frac{\partial^2 V_{m_2}}{\partial r^2} + C_T^2 \frac{1}{r} \frac{\partial V_{m_2}}{\partial r} + \left(\omega^2 - \frac{C_T^2 + m_2^2 C_L^2}{r^2} \right) V_{m_2} \\ - C_T^2 \frac{1+v}{1-v} \frac{1}{r} \frac{1}{\pi} \sum_{m_1=0}^{\infty} I_{m_1, m_2} \frac{\partial U_{m_1}}{\partial r} \\ - C_T^2 \frac{3-v}{1-v} \frac{1}{r^2} \frac{1}{\pi} \sum_{m_1=0}^{\infty} I_{m_1, m_2} U_{m_1} = 0, \end{aligned} \quad (4b)$$

where

$$\begin{aligned} I_{m_1, m_2} &= m_1 \int_0^{2\pi} \sin(m_1 \theta) \sin(m_2 \theta + \varphi) d\theta \\ &= \begin{cases} 0 & \text{for } m_1 \neq m_2 \\ m_1 \pi \cos \varphi & \text{for } m_1 = m_2 \end{cases}, \end{aligned} \quad (5a)$$

$$\begin{aligned} I_{m_2, m_1} &= m_2 \int_0^{2\pi} \cos(m_2 \theta + \varphi) \cos(m_1 \theta) d\theta \\ &= \begin{cases} 0 & \text{for } m_1 \neq m_2 \\ m_2 \pi \cos \varphi & \text{for } m_1 = m_2 \end{cases}. \end{aligned} \quad (5b)$$

Equations (4) and (5) indicate that $m_1 = m_2 = m$ is the condition for u and v to be coupled. The governing equations of free vibration take the following form for coupled circumferential modes:

$$\begin{aligned} C_L^2 \frac{\partial^2 U_m}{\partial r^2} + C_L^2 \frac{1}{r} \frac{\partial U_m}{\partial r} + \left(\omega^2 - \frac{C_L^2 + m^2 C_T^2}{r^2} \right) U_m \\ + C_T^2 \frac{1+v}{1-v} \frac{m}{r} \frac{\partial V_m}{\partial r} - C_T^2 \frac{3-v}{1-v} \frac{m}{r^2} V_m = 0, \end{aligned} \quad (6a)$$

$$\begin{aligned} C_T^2 \frac{\partial^2 V_m}{\partial r^2} + C_T^2 \frac{1}{r} \frac{\partial V_m}{\partial r} + \left(\omega^2 - \frac{C_T^2 + m^2 C_L^2}{r^2} \right) V_m \\ - C_T^2 \frac{1+v}{1-v} \frac{m}{r} \frac{\partial U_m}{\partial r} - C_T^2 \frac{3-v}{1-v} \frac{m}{r^2} U_m = 0. \end{aligned} \quad (6b)$$

Using the nondimensional parameter $x = r/\bar{r}$, where \bar{r} is the outer radius of the circular plate, and the relation $C_T^2 = [(1 - \nu)/2]C_L^2$, the above equations take the form

$$\begin{aligned} x^2 U_m''(x) + x U_m'(x) + \left(\frac{\omega^2 \bar{r}^2}{C_L^2} x^2 - 1 - m^2 \frac{1-v}{2} \right) U_m(x) \\ + \frac{1+v}{2} m x V_m'(x) - \frac{3-v}{2} m V_m(x) = 0, \end{aligned} \quad (7a)$$

$$\begin{aligned} x^2 V_m''(x) + x V_m'(x) + \left(\frac{\omega^2 \bar{r}^2}{C_T^2} x^2 - 1 - m^2 \frac{2}{1-v} \right) V_m(x) \\ - \frac{1+v}{1-v} m x U_m'(x) - \frac{3-v}{1-v} m U_m(x) = 0. \end{aligned} \quad (7b)$$

Henceforth, the prime is used to denote differentiation with respect to x .

B. Natural frequencies and mode shapes of the axisymmetric modes ($m=0$)

In the above equations the circumferential wave number m represents the number of nodal diameters in the free vibrational response. The case $m=0$ represents the vibrational modes where u and v are uncoupled. Hence, the uncoupled modes are axisymmetric because there are no nodal diameters. Equations (7) take the following form for the axisymmetric modes ($m=0$):

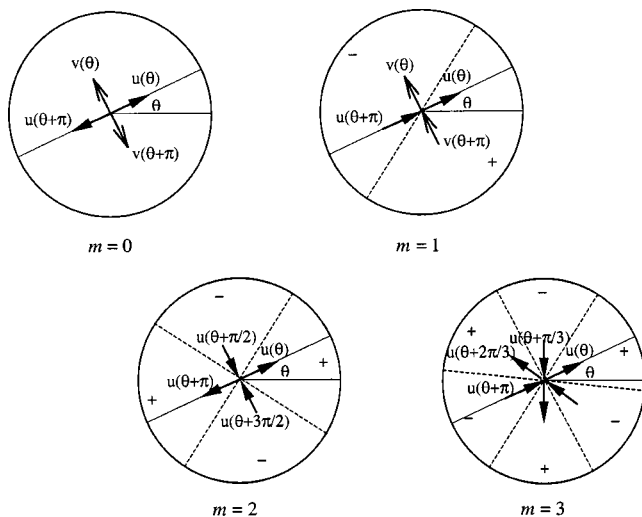


FIG. 2. Free vibration response at the center point of a circular clamped plate for different circumferential wave numbers m (note that for $m=2$ and $m=3$, v is not plotted; it follows the same behavior as u).

$$x^2 U_0''(x) + x U_0'(x) + (\lambda_p^2 x^2 - 1) U_0(x) = 0, \quad (8a)$$

$$x^2 V_0''(x) + x V_0'(x) + (\lambda_q^2 x^2 - 1) V_0(x) = 0, \quad (8b)$$

where $\lambda_p^2 = \omega^2 \bar{r}^2 / C_L^2$ and $\lambda_q^2 = \omega^2 \bar{r}^2 / C_T^2$.

Each of Eqs. (8) is a Bessel equation of order unity.¹² Solution of these equations is: $U_{0,p}(x) = J_1(\lambda_p x)$ and $V_{0,q}(x) = J_1(\lambda_q x)$, where $J_1(z)$ is Bessel function of the first kind of order unity. The displacement components must assume zero values at the clamped edge of the plate where $x = 1$. Hence, λ_p and λ_q ($= \lambda_p$ in this case) are the roots of the equation: $J_1(z) = 0$, excluding the root at the origin. The subscripts p and q stand for the sequential number of the root. Natural frequencies of the radial (extensional) and tangential (torsional) axisymmetric modes are given, respectively, by

$$\omega_{L,p} = \frac{C_L \lambda_p}{\bar{r}} \quad \text{and} \quad \omega_{T,q} = \frac{C_T \lambda_q}{\bar{r}}. \quad (9)$$

Radial distributions of the free vibrational modes take the form of Bessel functions of the first kind of order unity. It has to be stated here that Bessel function of the second kind is also a solution that satisfies Eqs. (8) but it is discarded here because it has a singularity at the origin.¹² In the following analysis, Bessel function of the first kind will be referred to as the Bessel function for brevity.

C. Natural frequencies and mode shapes of the coupled modes ($m > 0$)

Although Bessel function of the first order satisfies Eqs. (8) for axisymmetric modes, it does not satisfy Eqs. (7) due to the coupling between the radial, $U(x)$, and tangential, $V(x)$, components of in-plane vibration. However, the radial mode shapes of vibration can be expressed as a series of Bessel functions of any order. The physical in-plane response of the plate at the center $r = 0$ has to be examined first to help choosing a function that best represents the behavior at plate center point.

The in-plane response at $r = 0$ is examined and presented in Fig. 2, which also depicts the nodal diameters for the first four circumferential wave numbers ($m = 0:3$). For $m = 0$, if $U(0)$ and $V(0)$ components exist along and normal to any radial direction θ , it will be counteracted by similar response along and normal to the radial direction $\theta + \pi$ due to the axisymmetry. It follows that the axisymmetric mode, $m = 0$, must have zero response at the plate center point. This finding agrees with the value of Bessel function of order unity at $x = 0$, which is zero.

For $m = 1$, it is shown in Fig. 2 that u and v components of the response at $\theta + \pi$ are added to u and v components at θ . This is due to the change of sign of the sine and cosine functions [see Eqs. (3)] when the angle is increased by π . This indicates that the modes $m = 1$ have nonzero response at the plate center point.

Figure 2 also presents analysis of the response of the wave numbers $m = 2$ and $m = 3$ at the center of the plate. Following the same reasoning as before, it is shown that the plate center must have zero response for these two modes. It can also be shown that this always is the case for all the higher order modes ($m \geq 4$).

In summary, the modes with circumferential wave number $m = 1$ are the only ones that have nonzero response at the center point of the plate.

Radial mode shapes of the modes with circumferential wave numbers $m \geq 2$ are assumed as series summations of Bessel functions of integral order $n > 0$, which satisfies the condition of zero response at the plate center point

$$U_m(x) = \sum_{p=1}^{\infty} \bar{U}_{m,p} J_n(\lambda_p x), \quad (10a)$$

$$V_m(x) = \sum_{p=1}^{\infty} \bar{V}_{m,p} J_n(\lambda_p x), \quad (10b)$$

where λ_p are nonzero roots of the equation $J_n(z) = 0$.

Radial mode shapes of the modes with circumferential wave number $m = 1$ are assumed as a series summation of Bessel functions of the order 0.5 divided by \sqrt{x} as follows:

$$U_m(x) = \sum_{p=1}^{\infty} \bar{U}_{m,p} \frac{J_{0.5}(\lambda_p x)}{\sqrt{x}}, \quad (11a)$$

$$V_m(x) = \sum_{p=1}^{\infty} \bar{V}_{m,p} \frac{J_{0.5}(\lambda_p x)}{\sqrt{x}}, \quad (11b)$$

where λ_p are roots of the equation $J_{0.5}(z) = 0$.

It can be proved,¹² using expansion of $J_{0.5}(\lambda_p x)$ in terms of x , that $J_{0.5}(\lambda_p x) / \sqrt{x}$ assumes finite value at $x = 0$.

Henceforth, $J_{0.5}(\lambda_p x) / \sqrt{x}$ and $J_n(\lambda_p x)$ will be called “elemental mode shapes” for $m = 1$ and $m \geq 2$, respectively.

1. Natural frequencies and mode shapes of the modes $m \geq 2$

Three properties of Bessel functions will be considered in the following analysis:¹²

- (a) $U_{m,p}(x) = J_n(\lambda_p x)$ is a solution to the equation
- $$x^2 U_m''(x) + x U_m'(x) + (\lambda_p^2 x^2 - n^2) U_m(x) = 0. \quad (12)$$
- (b) Differentiation of Eq. (10a) is
- $$\frac{dU_m}{dx} = \sum_{p=1}^{\infty} \bar{U}_{m,p} \left[\frac{n}{x} J_n(\lambda_p x) - \lambda_p J_{n+1}(\lambda_p x) \right]. \quad (13)$$
- (c) The orthogonality of Bessel functions is expressed by the relations

$$\int_{x=0}^1 x J_n(\lambda_p x) J_n(\lambda_{p'} x) dx = \begin{cases} 0 & p \neq p' \\ \frac{1}{2} [J_{n+1}(\lambda_p)]^2 & p = p' \end{cases} \quad (14)$$

For brevity and clarity, the subscript m will be removed in the following analysis.

The following equations are obtained after substitution of (10) into (7) and using p' to denote a general elemental mode shape:

$$\begin{aligned} \sum_{p'=1}^{\infty} a_1 J_n(\lambda_{p'} x) \bar{U}_{p'} + \sum_{p'=1}^{\infty} (K_L^2 - \lambda_{p'}^2) x^2 J_n(\lambda_{p'} x) \bar{U}_{p'} \\ + \sum_{p'=1}^{\infty} a_2 J_n(\lambda_{p'} x) \bar{V}_{p'} \\ + \sum_{p'=1}^{\infty} a_3 x \lambda_{p'} J_{n+1}(\lambda_{p'} x) \bar{V}_{p'} = 0, \end{aligned} \quad (15a)$$

$$\begin{aligned} \sum_{p'=1}^{\infty} b_1 J_n(\lambda_{p'} x) \bar{V}_{p'} + \sum_{p'=1}^{\infty} (K_T^2 - \lambda_{p'}^2) x^2 J_n(\lambda_{p'} x) \bar{V}_{p'} \\ + \sum_{p'=1}^{\infty} b_2 J_n(\lambda_{p'} x) \bar{U}_{p'} \\ + \sum_{p'=1}^{\infty} b_3 x \lambda_{p'} J_{n+1}(\lambda_{p'} x) \bar{U}_{p'} = 0, \end{aligned} \quad (15b)$$

where

$$\begin{aligned} a_1 &= n^2 - 1 - m^2 \frac{1-\nu}{2}, & b_1 &= n^2 - 1 - m^2 \frac{2}{1-\nu}, \\ a_2 &= \frac{1+\nu}{2} mn - \frac{3-\nu}{2} m, & b_2 &= -\frac{1+\nu}{1-\nu} mn - \frac{3-\nu}{1-\nu} m, \\ a_3 &= -\frac{1+\nu}{2} m, & b_3 &= \frac{1+\nu}{1-\nu} m, \\ K_L^2 &= \frac{\omega^2 r^2}{C_L^2}, & \text{and } K_T^2 &= \frac{\omega^2 r^2}{C_T^2}. \end{aligned}$$

Multiplying Eqs. (15) by $[J_n(\lambda_p x)]/x$, integrating with respect to x from $x=0$ to $x=1$, and employing the orthogonal property give rise to the following equations:

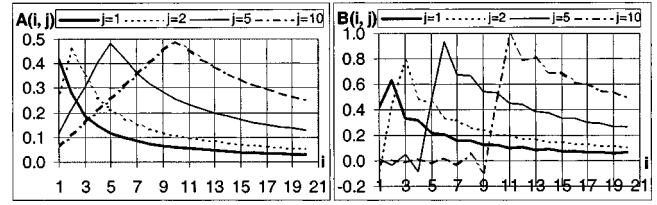


FIG. 3. Modal coupling factors $A_{p',p}$ and $B_{p',p}$ according to Eqs. (17) for $n=1$.

$$\begin{aligned} a_1 \sum_{p'=1}^{\infty} A_{p',p} \bar{U}_{p'} + (K_L^2 - \lambda_p^2) C_p \bar{U}_p + a_2 \sum_{p'=1}^{\infty} A_{p',p} \bar{V}_{p'} \\ + a_3 \sum_{p'=1}^{\infty} B_{p',p} \bar{V}_{p'} = 0, \end{aligned} \quad (16a)$$

$$\begin{aligned} b_1 \sum_{p'=1}^{\infty} A_{p',p} \bar{V}_{p'} + (K_T^2 - \lambda_p^2) C_p \bar{V}_p + b_2 \sum_{p'=1}^{\infty} A_{p',p} \bar{U}_{p'} \\ + b_3 \sum_{p'=1}^{\infty} B_{p',p} \bar{U}_{p'} = 0, \end{aligned} \quad (16b)$$

where $C_p = \frac{1}{2} [J_{n+1}(\lambda_p)]^2$,

$$\begin{aligned} A_{p',p} &= \int_{x=0}^1 \frac{1}{x} J_n(\lambda_{p'} x) J_n(\lambda_p x) dx, \\ B_{p',p} &= \int_{x=0}^1 \lambda_{p'} J_{n+1}(\lambda_{p'} x) J_n(\lambda_p x) dx. \end{aligned} \quad (17)$$

Equations (16) can be arranged in the form of an eigenvalue problem as follows:

$$\begin{aligned} K_L^2 \bar{U}_p &= \left(\lambda_p^2 - \frac{a_1}{C_p} A_{p,p} \right) \bar{U}_p - \sum_{\substack{p'=1 \\ (p' \neq p)}}^{\infty} \frac{a_1}{C_p} A_{p',p} \bar{U}_{p'} \\ &\quad - \sum_{p'=1}^{\infty} \frac{a_2 A_{p',p} + a_3 B_{p',p}}{C_p} \bar{V}_{p'}, \end{aligned} \quad (18a)$$

$$\begin{aligned} K_T^2 \bar{V}_p &= \left(\lambda_p^2 - \frac{b_1 A_{p,p}}{C_p} \right) \frac{1-\nu}{2} \bar{V}_p \\ &\quad - \sum_{\substack{p'=1 \\ (p' \neq p)}}^{\infty} \frac{b_1}{C_p} A_{p',p} \frac{1-\nu}{2} \bar{V}_{p'} \\ &\quad - \sum_{p'=1}^{\infty} \frac{b_2 A_{p',p} + b_3 B_{p',p}}{C_p} \frac{1-\nu}{2} \bar{U}_{p'}. \end{aligned} \quad (18b)$$

$A_{p',p}$ and $B_{p',p}$ represent the coupling strength between the elemental modes in Eqs. (10). They are plotted in Fig. 3 for different values of p' and p to give a qualitative measure of the expected convergence of the series summation in Eqs. (10). It can be proved that the integrand of $A_{p',p}$ assumes a zero value at $x=0$ for $n>0.5$.

For each circumferential wave number m , Eqs. (18) have to be written for a number, say k , of $U_{m,p}$ and $V_{m,p}$ coupled to $U_{m,p'}$ and $V_{m,p'}$ leading to $2 \times k$ equations. The resulting eigenvalue problem can be solved for the eigenvalues (non-dimensional frequency parameters K_L) and the eigenvectors

(corresponding mode shapes). The natural frequencies are obtained from the relation $f = (K_L C_L / 2\pi\bar{r})$ Hz.

It has to be mentioned here that Chen and Liu,¹³ in their solutions for in-plane vibration in plates with free edges, used series summation of Bessel functions of the first kind of order m (m is the circumferential wave number). The mathematical formulation presented in this section uses Bessel functions of integral order $n > 0$ (any arbitrary integer). This implies that m may be used as the order of Bessel functions in the series summation in Eqs. (10). However, Bessel functions of order unity were used in the computational examples. It was also confirmed by computational examples (not presented in this paper) that Bessel functions of order m give the same accuracy at the same computational effort [same number of terms in (10) to obtain certain accuracy].

2. Natural frequencies and mode shapes of the modes $m=1$

The same procedure will be followed in the analysis of this case as for $m \geq 2$.

The following properties of Bessel functions are used in the mathematical analysis:¹²

- (a) $U_{1,p}(x) = [J_{0.5}(\lambda_p x) / \sqrt{x}]$ is a solution of the equation
- $$x^2 U_{1,p}''(x) + 2x U_{1,p}'(x) + \lambda_p^2 x^2 U_{1,p}(x) = 0. \quad (19)$$
- (b) Differentiation of the elemental mode shape in Eqs. (11) is

$$\frac{d}{dx} \left[\frac{J_{0.5}(\lambda_p x)}{x^{0.5}} \right] = -\lambda_p \frac{J_{1.5}(\lambda_p x)}{x^{0.5}}. \quad (20)$$

TABLE I. Frequency parameters $K_L = \omega\bar{r}/C_L$ (see the list of symbols) for the first eight natural frequencies of the first ten circumferential wave numbers of in-plane vibration in circular clamped plates for two values of poisson's ratio $\nu=0.28$ and $\nu=0.33$.

Wave number	Poisson's ratio	1st mode	2nd mode	3rd mode	4th mode	5th mode	6th mode	7th mode	8th mode
$m=0$, radial	any	3.8317	7.0156	10.1735	13.3237	16.4706	19.6159	22.7601	25.9037
$m=0$, tangential	$\nu=0.28$	2.2990	4.2094	6.1041	7.9942	9.8824	11.7700	13.6560	15.5420
	$\nu=0.33$	2.2178	4.0606	5.8883	7.7116	9.5331	11.3530	13.1730	14.9930
$m=1$	$\nu=0.28$	1.9655	3.2210	5.0696	5.3802	7.0227	8.5254	8.9305	10.8060
	$\nu=0.33$	1.9441	3.1126	4.9104	5.3570	6.7763	8.4938	8.6458	10.4250
$m=2$	$\nu=0.28$	3.0658	4.1344	5.9357	6.7304	7.9147	9.7408	10.0340	11.7070
	$\nu=0.33$	3.0185	4.0127	5.7398	6.7079	7.6442	9.4356	9.9894	11.2970
$m=3$	$\nu=0.28$	3.9956	5.0741	6.7755	7.9853	8.8061	10.6370	11.3680	12.5930
	$\nu=0.33$	3.9116	4.9489	6.5537	7.9342	8.5336	10.2790	11.3380	12.1620
$m=4$	$\nu=0.28$	4.8244	6.0289	7.6072	9.1255	9.7416	11.4966	12.6630	13.4786
	$\nu=0.33$	4.7021	5.8985	7.3648	8.9816	9.5296	11.1087	12.5940	13.0582
$m=5$	$\nu=0.28$	5.5944	6.9753	8.4407	10.1174	10.7708	12.3452	13.8738	14.4020
	$\nu=0.33$	5.4370	6.8306	8.1834	9.8642	10.6324	11.9353	13.6350	14.1196
$m=6$	$\nu=0.28$	6.3301	7.8961	9.2829	11.0012	11.8604	13.1924	14.9187	15.4500
	$\nu=0.33$	6.1410	7.7265	9.0167	10.6837	11.7423	12.7712	14.5015	15.3093
$m=7$	$\nu=0.28$	7.0444	8.7821	10.1376	11.8398	12.9484	14.0485	15.8195	16.6034
	$\nu=0.33$	6.8259	8.5787	9.8666	11.4838	12.8054	13.6373	15.3240	16.4911
$m=8$	$\nu=0.28$	7.7444	9.6309	11.0042	12.6615	13.9993	14.9280	16.6701	17.7658
	$\nu=0.33$	7.4982	9.3887	10.7281	12.2792	13.7879	14.5598	16.1348	17.6027
$m=9$	$\nu=0.28$	8.4340	10.4450	11.8780	13.4780	14.9910	15.8480	17.5050	18.8890
	$\nu=0.33$	8.1612	10.1630	11.5920	13.0770	14.6840	15.5420	16.9450	18.5840
$m=10$	$\nu=0.28$	9.1159	11.2310	12.7500	14.2960	15.9150	16.8120	18.3350	19.9300
	$\nu=0.33$	8.8172	10.9100	12.4470	13.8830	15.5210	16.5510	17.7630	19.4540

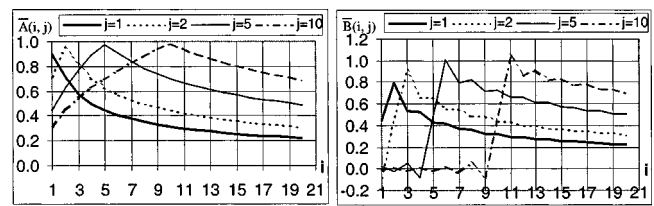


FIG. 4. Modal coupling factors $\bar{A}_{p',p}$ and $\bar{B}_{p',p}$ according to Eqs. (17) for $n=0.5$.

- (c) The orthogonality of Bessel functions according to Eqs. (14) is valid for $n=0.5$.

Substitution of the assumed solution from Eqs. (11) into Eqs. (7) gives rise to:

$$\sum_{p'=1}^{\infty} \bar{B}_{p',p} \bar{U}_{p'} + \bar{a}_1 \sum_{p'=1}^{\infty} \bar{A}_{p',p} \bar{U}_{p'} + (K_L^2 - \lambda_p^2) \bar{C}_p \bar{U}_p + \bar{a}_2 \sum_{p'=1}^{\infty} \bar{B}_{p',p} \bar{V}_{p'} + \bar{a}_3 \sum_{p'=1}^{\infty} \bar{A}_{p',p} \bar{V}_{p'} = 0, \quad (21a)$$

$$\sum_{p'=1}^{\infty} \bar{B}_{p',p} \bar{V}_{p'} + \bar{b}_1 \sum_{p'=1}^{\infty} \bar{A}_{p',p} \bar{V}_{p'} + (K_T^2 - \lambda_p^2) \bar{C}_p \bar{V}_p + \bar{b}_2 \sum_{p'=1}^{\infty} \bar{B}_{p',p} \bar{U}_{p'} + \bar{b}_3 \sum_{p'=1}^{\infty} \bar{A}_{p',p} \bar{U}_{p'} = 0, \quad (21b)$$

where

$$\bar{a}_1 = \bar{a}_3 = -\frac{3-\nu}{2},$$

$$\bar{b}_1 = \bar{b}_3 = -\frac{3-\nu}{1-\nu},$$

$$\bar{a}_2 = -\frac{1+\nu}{2}, \quad \bar{b}_2 = \frac{1+\nu}{1-\nu}.$$

\bar{C}_p , $\bar{A}_{p',p}$ and $\bar{B}_{p',p}$ are obtained from relations similar to those of Eqs. (17) with n replaced by 0.5. It can be proved that the integrand of $\bar{A}_{p',p}$ assumes finite value at $x=0$. This is another reason for the choice of the elemental mode shapes in Eqs. (11). Multiplying Eqs. (21) by $J_{0.5}(\lambda_p x)/\sqrt{x}$, integrating with respect to x from $x=0$ to $x=1$, and using the properties of Bessel functions, the following equations are obtained:

TABLE II. Coefficients ($\bar{U}_{m,p}$ and $\bar{V}_{m,p}$) of the elemental modes in $U_{1,p}$ (or $V_{1,p}$) = $\sum_{p=1}^{\infty} \bar{U}_{1,p}$ (or $\bar{V}_{1,p}$) $[J_{0.5}(\lambda_p x)]/\sqrt{x}$ [Eqs. (11)] and in $U_{m,p}$ (or $V_{m,p}$) = $\sum_{p=1}^{\infty} \bar{U}_{m,p}$ (or $\bar{V}_{m,p}$) $J_1(\lambda_p x)$ [Eqs. (10), $m \geq 2$] for in-plane vibration of circular clamped plate ($\nu=0.3$); the bold numbers highlight the elemental modes with the highest contribution to the radial mode shape.

Circum. wave #	Mode	P									
		1	2	3	4	5	6	7	8	9	10
$m=1$	#1, \bar{U}	0.7650	-0.0515	0.0170	-0.0080	0.0046	-0.0028	0.0020	-0.0013	0.0011	-0.0008
	, \bar{V}	-0.6376	-0.0686	0.0200	-0.0091	0.0051	-0.0031	0.0022	-0.0015	0.0012	-0.0009
	#2, \bar{U}	-0.2197	-0.1631	0.0257	-0.0107	0.0054	-0.0035	0.0022	-0.0017	0.0011	-0.0010
	, \bar{V}	-0.5907	0.7568	-0.0473	0.0159	-0.0074	0.0045	-0.0028	0.0020	-0.0014	0.0011
	#3, \bar{U}	-0.1133	0.4293	0.0940	-0.0153	0.0058	-0.0034	0.0019	-0.0015	0.0009	-0.0009
	, \bar{V}	-0.0072	0.4375	-0.7732	0.0614	-0.0212	0.0112	-0.0065	0.0045	-0.0031	0.0024
	#4, \bar{U}	-0.4573	0.6097	-0.1163	0.0243	-0.0113	0.0064	-0.0042	0.0028	-0.0022	0.0015
	, \bar{V}	-0.2038	-0.4850	0.3566	0.0261	-0.0119	0.0066	-0.0043	0.0028	-0.0022	0.0016
	#5, \bar{U}	0.0051	-0.0279	-0.1971	-0.0924	0.0165	-0.0077	0.0041	-0.0029	0.0018	-0.0015
	, \bar{V}	-0.0252	-0.0643	-0.5233	0.8186	-0.0487	0.0170	-0.0081	0.0051	-0.0032	0.0024
$m=2$	#1, \bar{U}	-0.8160	-0.0176	0.0021	-0.0006	0.0002	-0.0002	0.0000	-0.0001	0.0000	-0.0002
	, \bar{V}	0.5329	0.2164	-0.0489	0.0212	-0.0117	0.0072	-0.0049	0.0034	-0.0027	0.0019
	#2, \bar{U}	0.2670	0.3034	-0.0360	0.0146	-0.0078	0.0047	-0.0032	0.0022	-0.0017	0.0012
	, \bar{V}	0.5739	-0.7109	-0.0180	0.0084	-0.0047	0.0029	-0.0020	0.0014	-0.0012	0.0008
	#3, \bar{U}	0.0444	0.5463	0.2075	-0.0298	0.0126	-0.0069	0.0044	-0.0030	0.0022	-0.0016
	, \bar{V}	-0.0078	0.2578	-0.7672	0.0182	-0.0057	0.0027	-0.0016	0.0010	-0.0007	0.0005
	#4, \bar{U}	0.3393	-0.5820	0.1632	0.0020	0.0007	-0.0004	0.0005	-0.0001	0.0003	0.0001
	, \bar{V}	0.3813	0.5334	-0.2575	-0.1447	0.0409	-0.0198	0.0119	-0.0077	0.0057	-0.0039
	#5, \bar{U}	-0.0033	-0.1766	-0.3899	-0.1847	0.0261	-0.0113	0.0064	-0.0040	0.0029	-0.0020
	, \bar{V}	-0.0117	-0.0728	-0.3018	0.8278	-0.0004	-0.0010	0.0009	-0.0006	0.0006	-0.0004
$m=3$	#1, \bar{U}	-0.8417	0.2073	0.0753	0.0121	0.0134	0.0032	0.0050	0.0011	0.0022	0.0000
	, \bar{V}	0.4224	0.1579	-0.1930	0.0228	-0.0337	0.0083	-0.0132	0.0037	-0.0073	0.0013
	#2, \bar{U}	-0.2071	-0.2869	0.1804	-0.0007	0.0241	-0.0017	0.0088	-0.0011	0.0043	-0.0008
	, \bar{V}	-0.5237	0.7361	-0.1430	-0.0679	0.0020	-0.0165	0.0022	-0.0073	0.0012	-0.0045
	#3, \bar{U}	-0.1359	-0.5560	-0.0106	0.1815	0.0096	0.0308	0.0021	0.0118	0.0002	0.0052
	, \bar{V}	0.0418	-0.1464	0.7284	-0.2858	-0.0419	-0.0255	-0.0107	-0.0092	-0.0058	-0.0059
	#4, \bar{U}	-0.2312	0.4622	-0.4221	0.0570	0.0384	0.0027	0.0069	0.0003	0.0024	-0.0008
	, \bar{V}	-0.4231	-0.3412	0.4088	0.2348	-0.1743	0.0265	-0.0303	0.0102	-0.0131	0.0042
	#5, \bar{U}	0.0107	0.3265	0.3505	0.0053	-0.1776	-0.0175	-0.0323	-0.0051	-0.0126	-0.0010
	, \bar{V}	-0.0453	0.0002	0.1537	-0.7878	0.2909	0.0730	0.0229	0.0206	0.0093	0.0123
$m=4$	#1, \bar{U}	0.8096	-0.3702	-0.0668	0.0204	-0.0102	0.0058	-0.0040	0.0028	-0.0020	0.0020
	, \bar{V}	-0.3271	-0.1188	0.2703	-0.0775	0.0374	-0.0218	0.0144	-0.0096	0.0077	-0.0046
	#2, \bar{U}	-0.1635	-0.2873	0.2925	-0.0380	0.0172	-0.0099	0.0064	-0.0042	0.0034	-0.0017
	, \bar{V}	-0.4647	0.7186	-0.2491	-0.0823	0.0370	-0.0210	0.0138	-0.0091	0.0075	-0.0046
	#3, \bar{U}	0.1950	0.4962	-0.1631	-0.2314	0.0473	-0.0220	0.0131	-0.0083	0.0063	-0.0041
	, \bar{V}	-0.0760	0.1053	-0.6327	0.4639	-0.0140	0.0025	-0.0005	0.0000	0.0004	-0.0005
	#4, \bar{U}	0.1672	-0.1975	0.5279	-0.2436	-0.1045	0.0256	-0.0127	0.0074	-0.0051	0.0042
	, \bar{V}	0.3667	0.1351	-0.4147	-0.2921	0.4077	-0.0721	0.0331	-0.0186	0.0133	-0.0072
	#5, \bar{U}	0.0066	0.4648	0.1429	-0.1121	-0.2257	0.0359	-0.0169	0.0096	-0.0069	0.0036
	, \bar{V}	-0.1635	-0.0490	0.1785	-0.6596	0.4463	0.0461	-0.0232	0.0135	-0.0101	0.0059

TABLE II. (Continued.)

Circum. wave #	Mode	p									
		1	2	3	4	5	6	7	8	9	10
$m=5$	#1, \bar{U}	0.7585	-0.4823	-0.0128	0.0519	-0.0156	0.0116	-0.0068	0.0050	-0.0037	0.0034
	, \bar{V}	-0.2621	-0.0821	0.2984	-0.1416	0.0506	-0.0318	0.0192	-0.0133	0.0102	-0.0058
	#2, \bar{U}	0.1510	0.2845	-0.3874	0.1012	-0.0113	0.0123	-0.0054	0.0046	-0.0031	0.0012
	, \bar{V}	0.3988	-0.6788	0.3167	0.0790	-0.0832	0.0352	-0.0249	0.0151	-0.0129	0.0070
	#3, \bar{U}	0.2313	0.4073	-0.2783	-0.2132	0.1238	-0.0261	0.0221	-0.0105	0.0096	-0.0050
	, \bar{V}	-0.1060	0.1073	-0.5438	0.5570	-0.1100	-0.0137	-0.0006	-0.0042	0.0018	-0.0026
	#4, \bar{U}	0.1386	0.0266	0.4357	-0.4154	-0.0895	0.0950	-0.0193	0.0177	-0.0088	0.0078
	, \bar{V}	0.2699	-0.0078	-0.3332	-0.2477	0.5650	-0.1845	0.0250	-0.0230	0.0104	-0.0073
	#5, \bar{U}	0.0075	0.5104	-0.1077	-0.0839	-0.2188	0.1007	-0.0059	0.0132	-0.0044	0.0028
	, \bar{V}	-0.2789	-0.0155	0.2504	-0.5582	0.4490	0.0176	-0.0832	0.0280	-0.0261	0.0105
	#1, \bar{U}	-0.7048	0.5568	-0.0609	-0.0657	0.0291	-0.0162	0.0110	-0.0071	0.0059	-0.0048
	, \bar{V}	0.2181	0.0484	-0.2954	0.1954	-0.0754	0.0408	-0.0255	0.0168	-0.0131	0.0067
	#2, \bar{U}	0.1594	0.2677	-0.4645	0.1798	-0.0148	0.0076	-0.0042	0.0030	-0.0021	-0.0002
	, \bar{V}	0.3328	-0.6238	0.3563	0.0675	-0.1307	0.0615	-0.0365	0.0235	-0.0185	0.0099
$m=6$	#3, \bar{U}	-0.2507	-0.3128	0.3341	0.1679	-0.1900	0.0511	-0.0264	0.0158	-0.0116	0.0064
	, \bar{V}	0.1296	-0.1311	0.4805	-0.5940	0.2041	0.0111	-0.0109	0.0078	-0.0069	0.0043
	#4, \bar{U}	-0.1318	-0.1279	-0.2973	0.5067	-0.0027	-0.1493	0.0453	-0.0228	0.0157	-0.0097
	, \bar{V}	-0.2125	0.0858	0.2568	0.1485	-0.5953	0.3138	-0.0457	0.0191	-0.0109	0.0042
	#5, \bar{U}	-0.0257	-0.4727	0.2438	0.0082	0.2333	-0.1591	0.0061	-0.0041	0.0021	0.0016
	, \bar{V}	0.3222	-0.0647	-0.2411	0.5080	-0.4291	-0.0188	0.1481	-0.0635	0.0401	-0.0190
	#1, \bar{U}	-0.6541	0.6044	-0.1386	-0.0599	0.0441	-0.0224	0.0154	-0.0097	0.0084	-0.0064
	, \bar{V}	0.1869	0.0197	-0.2754	0.2329	-0.1063	0.0521	-0.0322	0.0210	-0.0160	0.0077
	#2, \bar{U}	0.1771	0.2356	-0.5183	0.2645	-0.0318	-0.0013	-0.0006	0.0004	0.0002	-0.0022
	, \bar{V}	0.2714	-0.5598	0.3748	0.0500	-0.1721	0.0973	-0.0510	0.0328	-0.0246	0.0124
	#3, \bar{U}	-0.2580	-0.2258	0.3449	0.1231	-0.2397	0.0907	-0.0325	0.0204	-0.0138	0.0070
	, \bar{V}	0.1447	-0.1599	0.4410	-0.6008	0.2787	-0.0043	-0.0280	0.0160	-0.0140	0.0076
	#4, \bar{U}	-0.1312	-0.1681	-0.1716	0.5349	-0.1177	-0.1665	0.0862	-0.0307	0.0225	-0.0121
	, \bar{V}	-0.1822	0.1348	0.1869	0.0627	-0.5584	0.4200	-0.0976	0.0164	-0.0101	0.0019
$m=7$	#5, \bar{U}	0.0526	0.4144	-0.2830	0.0369	-0.2747	0.2310	-0.0138	-0.0152	0.0058	-0.0079
	, \bar{V}	-0.3147	0.1366	0.1657	-0.4616	0.4284	0.0347	-0.2310	0.1203	-0.0571	0.0259
	#1, \bar{U}	-0.6078	0.6327	-0.2122	-0.0384	0.0554	-0.0305	0.0201	-0.0128	0.0112	-0.0081
	, \bar{V}	0.1636	-0.0035	-0.2474	0.2548	-0.1376	0.0668	-0.0394	0.0258	-0.0190	0.0086
	#2, \bar{U}	-0.1959	-0.1919	0.5457	-0.3464	0.0635	0.0093	-0.0063	0.0032	-0.0038	0.0046
	, \bar{V}	-0.2182	0.4939	-0.3789	-0.0264	0.2012	-0.1371	0.0696	-0.0425	0.0308	-0.0145
	#3, \bar{U}	0.2572	0.1536	-0.3297	-0.0904	0.2773	-0.1377	0.0434	-0.0235	0.0155	-0.0067
	, \bar{V}	-0.1501	0.1818	-0.4155	0.5933	-0.3319	0.0240	0.0471	-0.0305	0.0235	-0.0121
	#4, \bar{U}	0.1310	0.1812	0.0662	-0.5191	0.2240	0.1494	-0.1295	0.0472	-0.0291	0.0152
	, \bar{V}	0.1646	-0.1692	-0.1192	-0.0076	0.4945	-0.4909	0.1660	-0.0225	0.0061	-0.0001
	#5, \bar{U}	0.0819	0.3527	-0.2641	0.0293	-0.3111	0.3205	-0.0340	-0.0453	0.0242	-0.0159
	, \bar{V}	-0.2731	0.1775	0.0588	-0.3823	0.4260	0.0490	-0.3280	0.2013	-0.0792	0.0288
	#1, \bar{U}	-0.6078	0.6327	-0.2122	-0.0384	0.0554	-0.0305	0.0201	-0.0128	0.0112	-0.0081
	, \bar{V}	0.1636	-0.0035	-0.2474	0.2548	-0.1376	0.0668	-0.0394	0.0258	-0.0190	0.0086
$m=8$	#2, \bar{U}	-0.1959	-0.1919	0.5457	-0.3464	0.0635	0.0093	-0.0063	0.0032	-0.0038	0.0046
	, \bar{V}	-0.2182	0.4939	-0.3789	-0.0264	0.2012	-0.1371	0.0696	-0.0425	0.0308	-0.0145
	#3, \bar{U}	0.2572	0.1536	-0.3297	-0.0904	0.2773	-0.1377	0.0434	-0.0235	0.0155	-0.0067
	, \bar{V}	-0.1501	0.1818	-0.4155	0.5933	-0.3319	0.0240	0.0471	-0.0305	0.0235	-0.0121
	#4, \bar{U}	0.1310	0.1812	0.0662	-0.5191	0.2240	0.1494	-0.1295	0.0472	-0.0291	0.0152
	, \bar{V}	0.1646	-0.1692	-0.1192	-0.0076	0.4945	-0.4909	0.1660	-0.0225	0.0061	-0.0001
	#5, \bar{U}	0.0819	0.3527	-0.2641	0.0293	-0.3111	0.3205	-0.0340	-0.0453	0.0242	-0.0159
	, \bar{V}	-0.2731	0.1775	0.0588	-0.3823	0.4260	0.0490	-0.3280	0.2013	-0.0792	0.0288

$$\begin{aligned}
K_L^2 \bar{U}_p = & \left(\lambda_p^2 - \frac{\bar{B}_{p,p} + \bar{a}_1 \bar{A}_{p,p}}{\bar{C}_p} \right) \bar{U}_p \\
& - \sum_{\substack{p'=1 \\ (p' \neq p)}}^{\infty} \frac{\bar{B}_{p',p} + \bar{a}_1 \bar{A}_{p',p}}{\bar{C}_p} \bar{U}_{p'} \\
& - \sum_{p'=1}^{\infty} \frac{\bar{a}_2 \bar{B}_{p',p} + \bar{a}_3 \bar{A}_{p',p}}{\bar{C}_p} \bar{V}_{p'}, \quad (22a)
\end{aligned}$$

$$\begin{aligned}
K_L^2 \bar{V}_p = & \left(\frac{\lambda_p^2}{2} - \frac{B_{p,p} + \bar{b}_1 \bar{A}_{p,p}}{\bar{C}_p \frac{2}{1-\nu}} \right) \bar{V}_p \\
& - \sum_{\substack{p'=1 \\ (p' \neq p)}}^{\infty} \frac{\bar{B}_{p',p} + \bar{b}_1 \bar{A}_{p',p}}{\bar{C}_p \frac{2}{1-\nu}} \bar{V}_{p'} \\
& - \sum_{p'=1}^{\infty} \frac{\bar{b}_2 \bar{B}_{p',p} + \bar{b}_3 \bar{A}_{p',p}}{\bar{C}_p \frac{2}{1-\nu}} \bar{U}_{p'}. \quad (22b)
\end{aligned}$$

TABLE III. Comparisons of natural frequency predictions by the present method and by Finite Element Analysis for the plate described in Sec. III (“ p ” is number of terms in the series summation; “ e ” is total number of finite elements). (a) Axisymmetric modes, circumferential wave number $m=0$. (b) Circumferential wave number, $m=1$. (c) Circumferential wave numbers, $m=2$ and $m=3$. (d) Circumferential wave numbers, $m=7$ and $m=8$.

$m=0$, tangential modes					$m=0$, radial modes				
(a) Mode number	Eqs. (8)		Finite element		Mode number	Eqs. (8)		Finite element	
	Hz	p	Hz	e		Hz	p	Hz	e
1	3 860	1	3 868	1 200	1	6 434	1	6 439	1 200
			3 863	3 000				6 437	3 000
2	7 068	1	7 101	1 200	2	11 780	1	11 822	1 200
			7 084	3 000				11 807	3 000
3	10 249	1	10 347	1 200	3	17 082	1	17 229	1 200
			10 304	3 000				17 175	3 000
4	13 423	1	13 644	1 200					
			13 553	3 000					
5	16 593	1	17 014	1 200					
			16 849	3 000					
(b) Mode number	Eqs. (23)		Finite element		Mode number	Eqs. (22)		Finite element	
	Hz	p	Hz	e		Hz	p	Hz	e
1	3 301	5	3 303	1 200	6	14 321	5	14 398	1 200
	3 300	10	3 301	3 000		14 314	10	14 365	3 000
	3 300	20				14 315	20		
2	5 420	5	5 425	1 200	7	14 908	5	15 292	1 200
	5 406	10	5 416	3 000		15 001	10	15 176	3 000
	5 408	20				14 995	20		
3	8 484	5	8 558	1 200	8		5	18 691	1 200
	8 517	10	8 538	3 000		18 137	10	18 485	3 000
	8 512	20				18 144	20		
4	9 036	5	9 063	1 200	9	19 653	5		
	9 034	10	9 047	3 000		19 658	10	19 793	3 000
	9 034	20				19 657	20		
5	11 831	5	11 939	1 200					
	11 787	10	11 879	3 000					
	11 792	20							
$m=2$					$m=3$				
(c) Mode number	Eqs. (18)		Finite element		Eqs. (18)		Finite element		
	Hz	p	Hz	e	Hz	p	Hz	e	
1	5 149	5	5 155	1 200	6 710	10	6 729	1 200	
	5 148	10	5 151	3 000	6 709	20	6 716	3 000	
	5 148	20			6 709	30			
2	6 943	5	6 976	1 200	8 520	10	8 583	1 200	
	6 942	10	6 956	3 000	8 520	20	8 542	3 000	
	6 942	20			8 520	30			
3	9 969	5	10 048	1 200	11 378	10	11 489	1 200	
	9 967	10	10 013	3 000	11 377	20	11 435	3 000	
	9 967	20			11 376	30			
4	11 304	5	11 351	1 200	13 409	10	13 504	1 200	
	11 301	10	11 323	3 000	13 408	20	13 450	3 000	
	11 301	20			13 408	30			
5	13 295	5	13 486	1 200	14 790	10	15 015	1 200	
	13 290	10	13 410	3 000	14 786	20	14 922	3 000	
	13 289	20			14 786	30			
6	16 368	5	16 648	1 200	17 868	10	18 306	1 200	
	16 356	10	16 544	3 000	17 861	20	18 139	3 000	
	16 356	20			17 861	30			
7	16 862	5	17 060	1 200					
	16 849	10	16 970	3 000					
	16 848	20							

$\bar{A}_{p',p}$ and $\bar{B}_{p',p}$ represent the coupling strength between the elemental modes in Eqs. (11). They are plotted in Fig. 4 for different values of p' and p to give a qualitative measure of the expected convergence of the series summation in Eqs.

(11). \bar{C}_p is the same as in Eq. (17) with $n=0.5$.

Solution of the eigenvalue problem of Eqs. (22) will produce the frequency parameters K_L and the mode shapes.

Table I is a tabulation of the frequency parameters for

TABLE III. (Continued.)

(d) Mode number	$m=7$				$m=8$			
	Eqs. (18)		Finite element		Eqs. (18)		Finite element	
	Hz	p	Hz	e	Hz	p	Hz	e
1	11 832	10	12 090	1 200	13 009	10	13 367	1 200
	11 828	20	11 891	3 000	13 003	20	13 094	3 000
2	14 751	10	15 152	1 200	16 179	10	16 718	1 200
	14 746	20	14 857	3 000	16 171	20	16 323	3 000
3	17 028	10	17 484	1 200	18 486	10	19 077	1 200
	17 022	20	17 172	3 000	18 477	20	18 668	3 000
4	19 889	10						
	19 880	20	20 158	3 000				

the first eight modes of circumferential wave numbers $m=0$ to $m=10$. The frequency parameters are calculated for two values of Poisson's ratio $\nu=0.28$ and $\nu=0.33$ representing typical values for steel and aluminum, respectively. Coefficients of the elemental modes in Eqs. (11) and (10), are listed in Table II (calculated for $\nu=0.3$). Mode shapes are depicted in Table V as part of the computational example.

III. COMPUTATIONAL RESULTS AND COMPARISONS

In this section: (i) The natural frequencies and mode shapes are computed using the procedures of Sec. II. (ii) The results are compared with finite-element analysis predictions and with previously published data to assess the accuracy of the new method. (iii) Effect of the number of terms in the series summation of Eqs. (10) and (11) on the accuracy of the results is investigated. (iv) The effect of Poisson's ratio on the natural frequencies is examined computationally. (v) The effect of plate thickness is discussed. The procedures presented in Sec. II are used to compute the natural frequencies and mode shapes of for a 1-m-diameter circular thin plate clamped at the outer edge. MATLAB programming and eigenvalue solutions were used. The plate is made of steel of Young's modulus $E=200 \times 10^9$ N/m², density $\rho=7800$ kg/m³, and Poisson's ratio $\nu=0.28$. Finite-element analysis (FEA) was also used to compute the natural frequencies and mode shapes for comparison and assessment of the accuracy of the new method. Membrane shell elements were used in the finite-element analysis and two axisymmetric mesh models were built: (1) 1200-element mesh, 20 elements along the radius and 60 elements along the circumference; (2) 3000-element model, 25 elements along the radius and 120 element along the circumference.

Tables III(a)–(d) summarize the results of the numerical computations of the in-plane natural frequencies of the plate in the frequency range up to 20 kHz.

It is expected that the accuracy of the predictions of the present method will depend upon the number of terms used in the series summation of Eqs. (10) and (11). The predictions listed in Tables III(a)–(d) indicate that ten terms in the series summation are sufficient for accuracy to three significant figures.

The accuracy of FE predictions decreases as the frequency increases (for the same number of elements). In par-

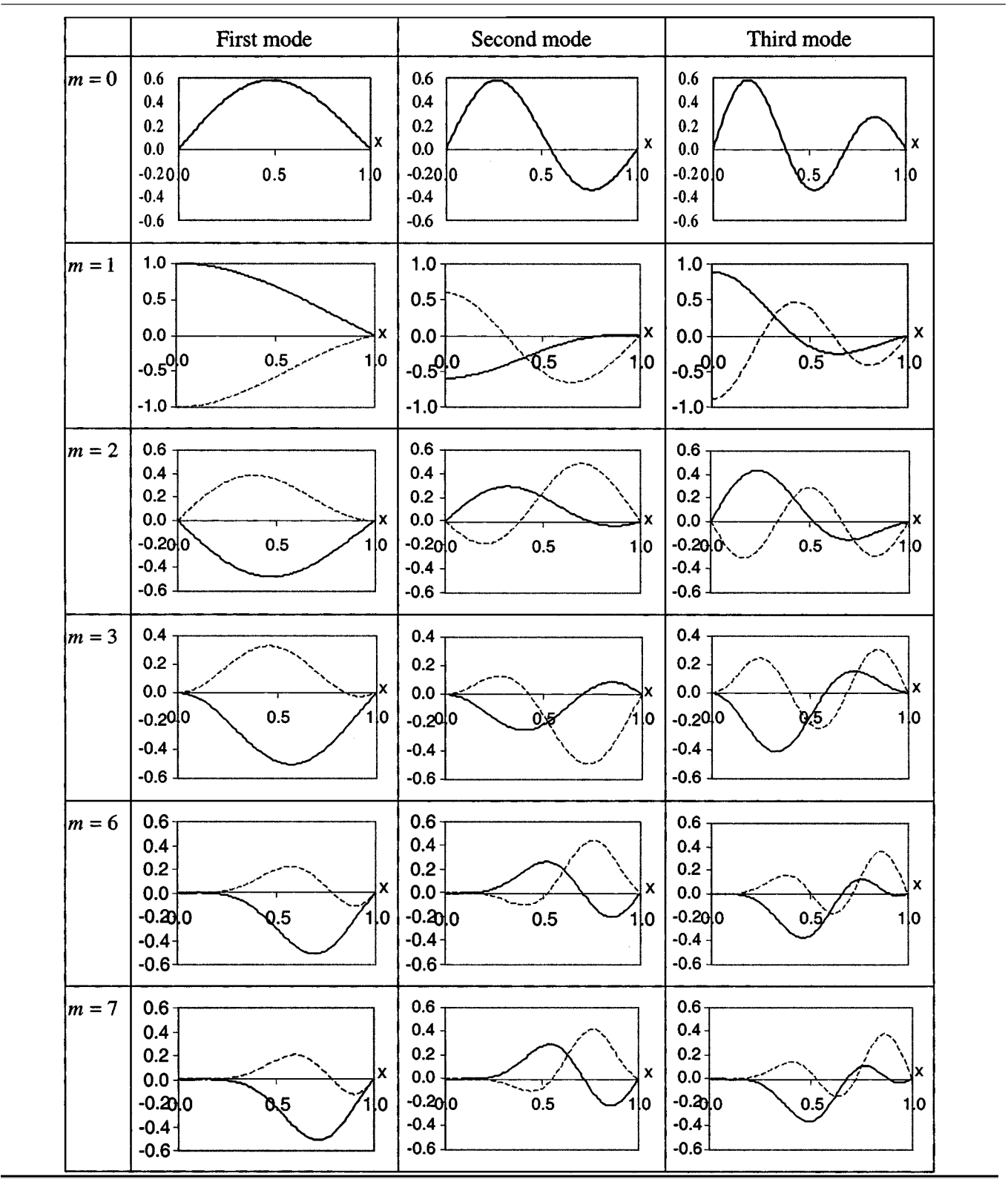
ticular, FE predictions are less accurate for the modes with large number of nodal diameters (i.e., large circumferential wave number m). It is clear from Tables III(a)–(d) that FE accuracy increased when the number of elements along the circumference was increased from 60 in the first mesh model to 120 in the second mesh model. It is also evident that FE predictions are always higher than the predictions of the present method. At low frequencies, i.e., lower circumferential wave numbers and lower-order modes, predictions of the present method and of FE are nearly identical. Consequently, the deviations at higher frequencies and for higher-order modes are judged to be due to insufficient number of elements in the FE mesh model.

It can be noted from Table I that the frequency parameters for the radial axisymmetric modes ($m=0$) are independent of Poisson's ratio because the radial and tangential components of in-plane response are uncoupled as explained in Sec. II B. On the contrary, the natural frequency of any asymmetric mode ($m>0$) decreases as Poisson's ratio increases. Similar behavior is reported in Ref. 8 for free disks. Calculations of frequency parameters in Table I are based on 20 terms in the series summation of Eqs. (10) and (11), which ensures accuracy to four significant figures. Comparisons with the results of Ref. 10 are presented in Table IV. It can be seen that the agreement is at least to three significant figures. The difference in the second line, between 7.013 and 7.0156, is worth comment. The present solution, for this case, is exact as indicated in Sec. II B and the frequency

TABLE IV. Comparisons of frequency parameters with the results of Ref. 10 ($\nu=0.3$ was used to calculate frequency parameters for $m>0$ (radial)).

Circumferential wave number	Irie <i>et al.</i> , Ref. 10	Present work
$m=0$	3.831	3.8317
(radial)	7.013	7.0156
$m=0$	2.267	2.2669
(tangential)	4.151	4.1505
$m=1$	1.958	1.9571
	3.180	3.1781
$m=2$	3.046	3.0474
	4.085	4.0859
$m=3$	3.964	3.9631
	5.024	5.0239
$m=4$	4.777	4.7768
	5.977	5.9767

TABLE V. Tabulation of the mode shapes, solid lines $U_{m,p}$, dashed lines $V_{m,p}$.



parameter computed by the present method is the second nonzero root of the equation $J_1(z)=0$.

The thickness of the plate is not appearing in the frequency equations [see Eqs. (8), (18) and (22)], implying that it has no effect on the free in-plane vibrational response of the plate. This is only true in the range of validity of Eqs. (1). This issue has been investigated and discussed in detail in

Ref. 14. It is shown, in Ref. 14, that the results of the two-dimensional plane stress equations [Eqs. (1)] are accurate in the frequency range where the frequency of the first thickness mode is much higher than the highest in-plane natural frequency in the range. The first thickness mode occurs when the plate thickness equals a half wavelength [see Eq. (29) in Ref. 14]. The coupling between the thickness modes and

in-plane modes increases as the two approach each other. Numerical and experimental results reported in Ref. 14 show that the effect of coupling between the fundamental thickness mode and the in-plane modes starts to diminish rapidly as the thickness to diameter ratio decreases to less than 0.2. In the transportation structures, of interest in this study, this ratio is usually far less than 0.2, indicating that the results of the present analysis are expected to be accurate through the frequency range of interest which is probably up to 20 kHz.

Radial distribution of the mode shapes are depicted in Table V for the first three modes of circumferential wave numbers $m=0, 1, 2, 3, 6$, and 7. It can be seen that the number of nodal circles is not always the same for the radial and tangential components of in-plane vibration of a mode. Also, a node of one component at a certain point is not necessarily associated with an antinode for the orthogonal component. In general, lower-order modes have more movement near the plate center, while higher-order modes have more movement near the clamped edge.

IV. CONCLUSIONS

The modal characteristics of in-plane vibration in circular thin flat plates with clamped edge are investigated in this work. The equations of in-plane vibration are solved for the natural frequencies and mode shapes. Assumed mode shapes are expressed in terms of trigonometric functions in the circumferential direction. It is proved that the circumferential modes are completely uncoupled. This means that the displacement components in the tangential and in the radial directions always have the same mode shape in the circumferential direction, though shifted by $\pi/2m$ where m is the circumferential wave number representing the number of nodal diameters [see Eqs. (3)].

It is shown that the modes with circumferential wave number equal to unity ($m=1$) are the only modes with displacement at the plate center point. All other modes have zero motion at the plate center point. As a consequence, only a combination of modes with $m=1$ is expected to constitute the response of the plate to in-plane excitation at the center point of the circular clamped plate.

It is proved that the mode shapes, in the radial direction, for the axisymmetric modes ($m=0$), are Bessel functions of the first type of order unity. Mode shapes in the radial direction for $m=1$ are assumed as series summation of elemental mode shapes of the form $J_{0.5}(\lambda_p x)/\sqrt{x}$ to suit the finite response at the plate center point. Mode shapes in the radial

direction for $m \geq 2$ are assumed as series summation of elemental mode shapes of the form $J_n(\lambda_p x)$ that represent the physical behavior of zero response at the plate center point.

The mathematical solution for free vibration is written in the form of an eigenvalue problem so that natural frequencies and modes shapes can be obtained by solving for the eigenvalues and the eigenvectors employing any suitable mathematical software. The frequency parameters (nondimensional natural frequencies) obtained by the present method are tabulated and the mode shapes are depicted to illustrate the free-vibration behavior in the frequency range of practical interest. Accuracy of the predictions of natural frequencies and mode shapes is assessed by comparisons with finite-element predictions and with previously reported results. The present method gives very accurate predictions. Effect of Poisson's ratio on the natural frequencies has also been examined. Natural frequency of the radial axisymmetric modes ($m=0$) is independent of Poisson's ratio, while the natural frequency of any other mode decreases as Poisson's ratio increases. Extension of the present solution to the case of elastically restrained edges, which practically is more realistic, is a possibility for future work.

¹A. Leissa, *Vibration of Plates* (Acoustical Society of America, Woodbury, NY, 1993).

²R. D. Blevins, *Formulas for Natural Frequency and Mode Shape* (Krieger, Malabar, FL, 1995).

³N. H. Farag and J. Pan, "Modal characteristics of in-plane vibration of rectangular plates," *J. Acoust. Soc. Am.* **105**, 3295–3310 (1999).

⁴N. H. Farag and J. Pan, "Free and forced in-plane vibration of rectangular plates," *J. Acoust. Soc. Am.* **103**, 408–413 (1998).

⁵J. So and A. W. Leissa, "Three-dimensional vibrations of thick circular and annular plates," *J. Sound Vib.* **209**(1), 15–41 (1998).

⁶R. M. Grice and R. J. Pinnington, "Vibration analysis of a thin-plate box using a finite element model which accommodates only in-plane motion," *J. Sound Vib.* **232**(2), 449–471 (2000).

⁷C. F. Liu and Y. T. Lee, "Finite element analysis of three-dimensional vibrations of thick circular and annular plates," *J. Sound Vib.* **233**(1), 63–80 (2000).

⁸R. Holland, "Numerical studies of elastic-disk contour modes lacking axial symmetry," *J. Acoust. Soc. Am.* **40**, 1051–1057 (1966).

⁹G. Ambati, J. F. W. Bell, and J. C. K. Sharp, "In-plane vibrations of annular rings," *J. Sound Vib.* **47**(3), 415–432 (1976).

¹⁰T. G. Irie, Yamada, and Y. Muramoto, "Natural frequencies of in-plane vibration of annular plates," *J. Sound Vib.* **97**(1), 171–175 (1984).

¹¹J. S. Chen and J. L. Jhu, "On the in-plane vibration and stability of a spinning annular disk," *J. Sound Vib.* **195**(4), 585–593 (1996).

¹²O. J. Farrell and B. Ross, *Solved Problems: Gamma and Beta Functions, Legendre Polynomials, Bessel Functions* (Macmillan, New York, 1963).

¹³S. S. H. Chen and T. M. Liu, "Extensional vibration of thin plates of various shapes," *J. Acoust. Soc. Am.* **58**, 828–831 (1975).

¹⁴T. R. Kane and R. D. Mindlin, "High-frequency extensional vibrations of plates," *J. Appl. Mech.* **23**, 277–283 (1956).

The wave transmission coefficients and coupling loss factors of point connected structures

R. S. Langley^{a)} and P. J. Shorter

Vibro-Acoustic Sciences, Inc., 12555 High Bluff Drive, Suite 310, San Diego, California 92130

(Received 18 December 2001; revised 3 August 2002; accepted 29 August 2002)

This analysis is concerned with the calculation of the elastic wave transmission coefficients and coupling loss factors between an arbitrary number of structural components that are coupled at a point. A general approach to the problem is presented and it is demonstrated that the resulting coupling loss factors satisfy reciprocity. A key aspect of the method is the consideration of cylindrical waves in two-dimensional components, and this builds upon recent results regarding the energetics of diffuse wavefields when expressed in cylindrical coordinates. Specific details of the method are given for beam and thin plate components, and a number of examples are presented.

© 2002 Acoustical Society of America. [DOI: 10.1121/1.1515791]

PACS numbers: 43.40.Qi, 43.20.Ks, 43.40.Dx [JHG]

I. INTRODUCTION

The calculation of the elastic wave transmission coefficients between coupled structural components is of interest in many areas of structural acoustics. For example, the wave transmission coefficients are often used to determine the coupling loss factors between the subsystems in a statistical energy analysis model. Given the importance of the coupling loss factors to the prediction of the vibrational energy distribution in complex structures it is perhaps not surprising that they have been the subject of a number of previous studies.

A general derivation of the wave transmission coefficients between an arbitrary number of semi-infinite plate substructures coupled along an infinite beam has been presented by Langley and Heron (1990). The analysis was based on a Fourier series decomposition of the plate response into various incident and reflected trace wave components. The analysis was extended to encompass curved shells by Langley (1994), and semi-infinite “strip” plates contained within the junction by Heron (1998). In many instances, the connection between various substructures is not continuous but instead takes the form of a number of discrete point connections. At low frequencies, when the wavelength is long in comparison to the spacing between the connection points, the junction behaves like a continuous line junction. However, when the wavelength becomes comparable to the connection spacing, this approach is no longer valid. The analysis of equally spaced point connections (or “combs”) has been discussed by Heron (1998) and also by Bosmans and Vermeir (1997) and Bosmans and Nightingale (2001). The dynamic stiffness of the connection points can be phrased in terms of a Fourier series, and the response of the junction to various incident trace wavenumbers in the connected substructures calculated.

At high frequencies (and for randomly spaced point connections), the motions of the individual connection points become incoherent and the point connections can be analyzed individually. In principle, the analysis presented by

Heron (1998) and Bosmans and Nightingale (2001) could be used to determine the response of an individual point connection. However, such an approach requires a large number of trace wavenumbers to be considered to achieve an accurate description of the scattering that occurs at the connection point. Given the computational expense associated with such a calculation, it is natural to question whether an alternative, more computationally efficient, approach exists.

Situations also arise in which a number of beams and plates are coupled together at a single point connection. The analysis of the transmission of elastic waves between a number of point connected beams has been studied extensively in the past (see, for example, Cremer and Heckl, 1988; Sablik, 1982; Moore, 1990; Horner and White, 1991; Guo, 1995; Tso and Norwood, 1995). However, the general analysis of wave transmission between a number of point connected beam and plate substructures, and the calculation of the associated coupling loss factors, appears to have received relatively little attention. One possible explanation is that the choice of wavetypes used to describe the response of a plate is not immediately obvious. Clearly, excitation of the junction generates a number of outgoing cylindrical waves in the plate and a natural wave description of the plate is therefore in terms of incoming and outgoing cylindrical waves. However, when deriving the SEA coupling loss factors it is typically assumed that a diffuse field exists within the excited subsystem. Since the concept of a diffuse field is generally defined in terms of propagating plane waves, the nature of the incident wavefield is not immediately obvious.

The traditional definition of a diffuse field in terms of plane waves may perhaps be attributed to the widespread use of a Cartesian coordinate system to describe the response of rectangular rooms in classical room acoustics. A diffuse field can, however, be defined in terms of the natural wavetypes associated with any coordinate system in which the wave equation is separable. For example, Langley and Shorter (2001) have recently provided expressions for the diffuse bending, extensional, and shear wavefields of a thin plate, expressed in cylindrical coordinates. These expressions are employed in the current paper to provide a general derivation of the wave transmission coefficients and coupling loss fac-

^{a)}Visiting Research Scientist. Permanent address: Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK. Electronic mail: rsl21@eng.cam.ac.uk

tors between an arbitrary number of semi-infinite beams and infinite plates coupled at a point.

It is worth mentioning that the coupling loss factors for a point connection that has a single degree of freedom can be written directly in terms of the point impedances (or mobilities) of the various substructures—see, for example, Eq. (20) of Manning (1994) or Eq. (10.1.19) of Lyon and DeJong (1995). In fact, this approach was employed by Lyon and Eichler (1964) for a beam-plate structure in one of the first predictions of a SEA coupling loss factor. There are a number of references in the literature that discuss the calculation of the point impedance of various substructures (see, for example, Cremer *et al.*, 1988; Dyer, 1960; Ljunggren, 1984; Leung and Pinnington, 1986; Petersson and Heckl, 1996; Sarradj, 2000). However, the previously mentioned coupling loss factor formulas (Manning, 1994; Lyon and DeJong, 1995) do not apply when the junction has more than one degree of freedom, in which case a wave approach is normally adopted.

This paper presents a general derivation of the elastic wave transmission coefficients and coupling loss factors between an arbitrary number of semi-infinite beams and infinite plates coupled at a point connection. The beams and plates may be oriented at any given angle and the internal dynamics and finite dimension of the junction can be fully accounted for through the use of a dynamic stiffness formulation. The analysis assumes that the substructures are connected to the junction via a number of massless, rigid discs. It is shown that SEA reciprocity holds for such junctions (for an appropriate choice of SEA subsystems) and a number of numerical examples are presented to validate and demonstrate the approach.

II. COUPLING LOSS FACTORS FOR POINT CONNECTIONS

A. General methodology

This work is concerned with the wave transmission properties of a structural junction at which a number N_C of structural components (for example, beams and/or plates) are coupled at a single point, as shown in Fig. 1. Each component j is taken to have an embedded rigid massless disc which is connected to the coupling point and which has six degrees of freedom \mathbf{w}_j in a local Cartesian reference frame: (w_{j1}, w_{j2}, w_{j3}) representing displacements and (w_{j4}, w_{j5}, w_{j6}) representing rotations. The present concern is with the steady-state situation where an elastic wave of frequency ω propagates through one of the components towards the junction, so that all time dependencies have the form $\exp(i\omega t)$. In general this wave will be partially reflected and partially transmitted, leading to outgoing waves in each component. The aim of the following analysis is to compute the amplitudes of these waves and thereby find the wave transmission coefficients and SEA coupling loss factors of the junction. Initially compatibility and equilibrium conditions for each component must be established, and these are then assembled into equations that govern the dynamics of the complete junction. The solution of these equations then yields the required wave amplitudes.

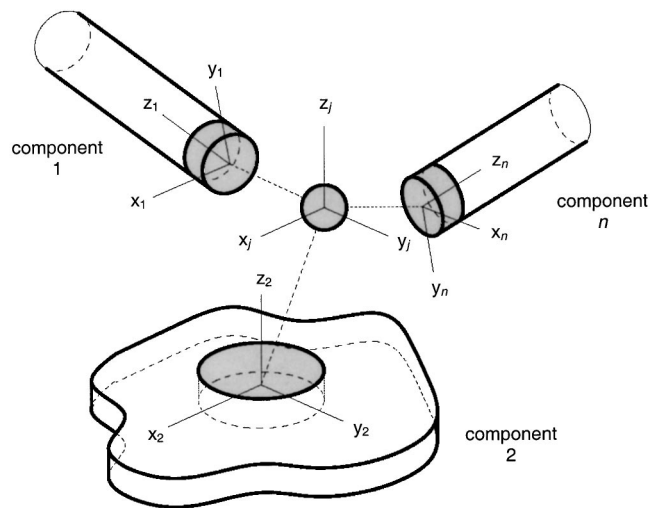


FIG. 1. A general point connection between an arbitrary number of structural components.

B. Compatibility conditions for the j th component

It will be considered initially that the j th component does not carry an incident wave towards the junction. Furthermore, it will be assumed that any other component boundaries (apart from the point connection) are absorbing, in the sense that a wave which is incident on one of these boundaries is not reflected back. This is equivalent to taking the component to be of infinite extent, so that the transmission properties of the junction can be isolated from the complex dynamics of the actual system. It then follows that the motion of the component can be expressed in terms of waves that either propagate away from the junction, or (in the case of evanescent waves) decay in amplitude away from the junction. The number and nature of these waves will depend upon the structure under consideration: in the case of a thin beam (see Sec. III) there will be six waves, corresponding to four propagating waves in torsion, extension, and bending (in each of two planes), and two evanescent bending waves. For a thin plate (see Sec. IV) there will be three propagating bending waves, three evanescent bending waves, two propagating longitudinal waves, and three propagating shear waves, giving a total of 11 waves. To formulate the equations of motion of the junction, the amplitudes of the waves must be related to the displacements \mathbf{w}_j of the component connection disc. Each wave will produce motion of this disc, and the various contributions will add linearly, so that it is possible to write

$$\mathbf{w}_j = \sum_k \mathbf{q}_{1jk} a_{jk} = \mathbf{Q}_{1j} \mathbf{a}_j. \quad (1)$$

Here a_{jk} is the amplitude of the k th wave propagating or decaying away from the junction in the j th component, and \mathbf{q}_{1jk} is a vector that contains the contribution to \mathbf{w}_j produced by a wave of unit amplitude—this vector can be found routinely from the detailed mechanics of the wave motion. The matrix \mathbf{Q}_{1j} has columns \mathbf{q}_{1jk} , and it is important to note that this matrix may not be square: for the case of a beam the matrix is square (6×6) but for the case of a plate the matrix is rectangular (6×11). If \mathbf{Q}_{1j} is rectangular, then additional

conditions must be imposed on \mathbf{a}_j if the wave amplitudes are to be determined uniquely from the motion of the junction. Such additional conditions can be written in the form

$$\mathbf{Q}_{2j}\mathbf{a}_j = \mathbf{0}, \quad (2)$$

where \mathbf{Q}_{2j} is a constraint matrix of appropriate dimensions. In the case of a plate for example, the additional conditions might state that the plate is clamped to the connection disc, so that there is no relative rotation about the circumferential boundary; in this case \mathbf{Q}_{2j} can be shown to be of dimension 5×11 (see Sec. IV). Equations (1) and (2) can be combined in the form

$$\begin{pmatrix} \mathbf{Q}_{1j} \\ \mathbf{Q}_{2j} \end{pmatrix} \mathbf{a}_j = \mathbf{Q}_j \mathbf{a}_j = \begin{pmatrix} \mathbf{w}_j \\ \mathbf{0} \end{pmatrix}, \quad (3)$$

The square matrix \mathbf{Q}_j can be inverted to yield

$$\mathbf{a}_j = \mathbf{Q}_j^{-1} \begin{pmatrix} \mathbf{w}_j \\ \mathbf{0} \end{pmatrix} = (\mathbf{P}_j \quad \mathbf{R}_j) \begin{pmatrix} \mathbf{w}_j \\ \mathbf{0} \end{pmatrix}, \quad (4)$$

$$\mathbf{a}_j = \mathbf{P}_j \mathbf{w}_j, \quad (5)$$

where \mathbf{P}_j and \mathbf{R}_j are (possibly nonsquare) partitions of \mathbf{Q}_j^{-1} . Equation (5), representing compatibility between the junction motion and the wave amplitudes, is the main result of this section.

C. Elastic forces in the j th component

The elastic forces imparted to the j th component by the connection disc can be written in the form

$$\mathbf{F}_j = \mathbf{S}_j \mathbf{a}_j, \quad (6)$$

where the matrix \mathbf{S}_j can be determined from the mechanics of the wave motion. In general this matrix will be rectangular, being of dimension 6×11 for a plate, although for the case of a beam the matrix is square and 6×6 .

D. The dynamic stiffness matrix of the j th component

Equations (5) and (6) can be combined to yield a relationship between the elastic forces imparted by component j and the motion of the coupling disk in the form

$$\mathbf{F}_j = \mathbf{S}_j \mathbf{P}_j \mathbf{w}_j = \mathbf{D}_j \mathbf{w}_j, \quad (7)$$

where \mathbf{D}_j is the (6×6) dynamic stiffness matrix of the component. Often in vibro-acoustics the behavior of a system is phrased in terms of the impedance matrix (which relates the system velocity to the applied force) rather than the dynamic stiffness matrix. Although the impedance matrix is not employed here, for completeness it can be noted that in the present notation this matrix has the form $\mathbf{Z}_j = \mathbf{D}_j / i\omega$ —in fact this is the *point* impedance matrix of the component, and the inverse would be the point mobility matrix.

E. The effect of an incoming wave in the j th component

Equations (1)–(7) are based on the premise that the j th component carries only outgoing waves. If the component carries an incoming wave, then the equations must be modified accordingly. In particular, Eq. (3) becomes

$$\mathbf{Q}_j \mathbf{a}_j + \begin{pmatrix} \mathbf{w}_{in} \\ \mathbf{q}_{in} \end{pmatrix} a_{in} = \begin{pmatrix} \mathbf{w}_j \\ \mathbf{0} \end{pmatrix}, \quad (8)$$

where a_{in} is the amplitude of the incident wave, \mathbf{w}_{in} is the motion of the connection disk produced by an incident wave of unit amplitude, and \mathbf{q}_{in} arises from the appropriate addition to the constraint Eq. (2). Similarly Eq. (6) is modified to become

$$\mathbf{F}_j = \mathbf{S}_j \mathbf{a}_j + \mathbf{F}_{in} a_{in} \quad (9)$$

where \mathbf{F}_{in} is the force produced at the connection disc by an incoming wave of unit amplitude. By combining Eqs. (8) and (9), and noting Eq. (4), the revised version of Eq. (7) is found to be

$$\mathbf{F}_j = \mathbf{D}_j \mathbf{w}_j + [\mathbf{F}_{in} - \mathbf{S}_j(\mathbf{P}_j \mathbf{w}_{in} + \mathbf{R}_j \mathbf{q}_{in})] a_{in}. \quad (10)$$

In this equation all of the terms on the right-hand side, other than \mathbf{w}_j (and hence \mathbf{a}_j), are known, in the sense that they can be found by considering in detail the mechanics of wave motion in the component, as explained in Secs. III and IV, respectively, for a beam and a plate. To find the wave amplitudes \mathbf{a}_j it is necessary to formulate the equations of the complete system, as explained in what follows.

F. Assembly and solution of the junction equations

Given that there is no external force applied at the junction, equilibrium requires that the forces \mathbf{F}_j in the various components must sum to zero. It then follows from Eq. (10) that the equations of motion that govern the complete junction have the form

$$\mathbf{w}_j = \mathbf{T}_j \mathbf{w}, \quad (11)$$

$$\left(\sum_{j=1}^{N_C} \mathbf{T}_j^T \mathbf{D}_j \mathbf{T}_j \right) \mathbf{w} = -\mathbf{T}_r^T [\mathbf{F}_{in} - \mathbf{S}_r(\mathbf{P}_r \mathbf{w}_{in} + \mathbf{R}_r \mathbf{q}_{in})] a_{in}, \quad (12)$$

where the incident wave has been taken to occur in component r —by problem definition there is only one incident wave, since the aim is to compute the wave amplitudes generated by the reflection and transmission of a single wave. The matrix \mathbf{T}_j that appears in Eq. (11) is a rotation matrix between the local component degrees of freedom \mathbf{w}_j and a global set \mathbf{w} common to the whole junction. Equation (12) can be solved to yield the junction motion \mathbf{w} , and the wave amplitudes can then be recovered from Eq. (5) for any component $j \neq r$, and Eq. (8) for component r .

It has been assumed in the forgoing analysis that the components are rigidly attached at the junction. If this is not the case, but rather (for example) the components are connected via an arbitrary linear dynamic system, then Eq. (12) can readily be modified; in brief, the component equations are added in to the equations of motion of the dynamic coupling.

G. Wave transmission coefficients

Having found the amplitudes of the generated waves, the wave (power) transmission coefficients can be computed. The propagating wave components transmit power away from the junction, and the mean rate of transmission is pro-

portional to the product of the amplitude squared and the group velocity c_g . For a wave of amplitude A this can be written as

$$P = \alpha |A|^2 c_g. \quad (13)$$

For a one-dimensional component $\alpha |A|^2$ is the energy density, so that $\alpha = m\omega^2/2$ where m is the mass or moment of inertia per unit length as appropriate. The energy density is equal to the sum of the kinetic and potential energy densities, and for wave motion in the far field this is equal to twice the kinetic energy density; the factor of $\frac{1}{2}$ that appears in α accounts for the fact that harmonic motion is considered, so that the mean value of the energy is equal to half the peak value. For a two-dimensional component the waves generated by a point coupling are cylindrical waves, and the energetics of such waves have been considered by Langley and Shorter (2001). In this case $\alpha = m\omega c$ unless the wave is radially symmetric, when $\alpha = 2m\omega c$; here m is the mass per unit area and c is the phase speed.

The wave transmission coefficient τ_{jk} between any two wave components A_j and A_k is defined as the ratio of the power transmitted away from the junction by component A_j to the power carried towards the junction by an incident wave A_k . It follows from Eq. (13) that this can be written as

$$\tau_{jk} = \frac{\alpha_j c_{gj} |A_j|^2}{\alpha_k c_{gk} |A_k|^2}. \quad (14)$$

To relate Eq. (14) to the analysis of the previous sections, A_j can be identified with any wave amplitude a_{jk} (the k th wave in the j th component) and A_k can be identified with the incident wave amplitude a_{in} . By repeating the analysis for a range of incident waves, the full set of junction transmission coefficients can be found. It can be noted that the transmission coefficient is symmetric with $\tau_{jk} = \tau_{kj}$ (Cremer *et al.*, 1988; Mace, 1992).

H. Coupling loss factors and reciprocity

1. General definition

In statistical energy analysis (SEA) the net energy flow between two coupled “subsystems” j and k (from subsystem j to subsystem k) is written in the form

$$P_{jk} = \omega \eta_{jk} n_j (E_j/n_j - E_k/n_k), \quad (15)$$

where E_r is the vibrational energy (twice the time averaged kinetic energy) of subsystem r ($r = j, k$), n_r is the modal density of the subsystem (the average number of resonant modes in a unit frequency band), and η_{jk} is the coupling loss factor (CLF) between the two subsystems j and k . In this context a “subsystem” does not generally refer to a complete structural component, but rather a particular wave type in a particular component—for example, axial waves in a beam or bending waves in a plate. The modal density is well established for most forms of subsystem, and the standard results for one- and two-dimensional wavefields are (Lyon and De-Jong, 1995)

$$n_{1-D} = L/(\pi c_g), \quad (16)$$

$$n_{2-D} = \omega A/(2\pi c_g c). \quad (17)$$

Here L is the length of the one-dimensional component, A is the area of the two-dimensional component, and c and c_g are, respectively, the appropriate group and phase velocities. Apart from the subsystem energies, the only other potentially unknown quantity on the right-hand side of Eq. (15) is the coupling loss factor η_{jk} , and the aim of the following subsections is to demonstrate that this factor can be expressed in terms of the appropriate transmission coefficient for a point connection. In line with standard SEA assumptions, each subsystem is taken to consist of a diffuse wavefield, which means that energy propagates with equal intensity in each direction, and the constituent waves are statistically independent. The form of the CLF η_{jk} will depend on the dimensions of the subsystems j and k , and various cases are considered in what follows.

2. 2-D subsystem $j \rightarrow$ 1-D subsystem k

In this case subsystem j might represent, for example, bending waves in a plate, and subsystem k might represent axial waves in a beam. For a point connection the 2-D subsystem is conveniently expressed in terms of cylindrical waves, and Langley and Shorter (2001) have derived various results for the energetics of a diffuse wavefield that is described in this way. The field consists of an infinite number of pairs of cylindrical waves (one in-coming and one out-going), and the power carried by *each* wave can be written in the form

$$P = E_j c_j c_{gj}/(\omega A), \quad (18)$$

where E_j is the *total* energy of the diffuse wavefield, A is the area of the component, c_{gj} is the group velocity, and c_j is the phase velocity. The energy transmitted to subsystem k can therefore be written as

$$P_{jk} = \left(\sum_{r_j} \tau_{kr_j} \right) E_j c_j c_{gj}/(\omega A), \quad (19)$$

where the sum over r_j represents a sum over all incoming cylindrical waves that are part of subsystem j . Equation (19) constitutes the first term that appears on the right-hand side of Eq. (15), and by noting Eq. (17) it then follows that

$$\eta_{jk} n_j = \frac{1}{2\pi\omega} \left(\sum_{r_j} \tau_{kr_j} \right), \quad (20)$$

$$\eta_{jk} = \frac{c_{gj} c_j}{\omega^2 A} \left(\sum_{r_j} \tau_{kr_j} \right). \quad (21)$$

3. 1-D subsystem $k \rightarrow$ 2-D subsystem j

In this case the energy propagated towards the junction in the 1-D subsystem is partially transmitted into cylindrical waves in the 2-D subsystem. The energy density in the 1-D subsystem can be written as E_k/L , and if the waves are diffuse then half of this energy is associated with waves that are propagating towards the junction. The power incident on the junction is thus

$$P = c_{gk} E_k/(2L), \quad (22)$$

and the power transmitted is

$$P_{kj} = \left(\sum_{r_j} \tau_{r_j k} \right) E_k c_{gk} / (2L). \quad (23)$$

It then follows from Eq. (15) (with j and k interchanged) and Eq. (16) that

$$\eta_{kj} n_k = \frac{1}{2\pi\omega} \left(\sum_{r_j} \tau_{r_j k} \right), \quad (24)$$

$$\eta_{kj} = \frac{c_{gk}}{2\omega L} \left(\sum_{r_j} \tau_{r_j k} \right). \quad (25)$$

Given that the wave transmission coefficient is symmetric, it follows from Eqs. (20) and (24) that the coupling loss factors satisfy *reciprocity* in that $\eta_{jk} n_j = \eta_{kj} n_k$.

4. Other cases

Following similar arguments to those employed in the previous subsections, it can be shown that the coupling loss factor between two 2-D subsystems is given by

$$\eta_{jk} n_j = \frac{1}{2\pi\omega} \left(\sum_{r_j} \sum_{s_k} \tau_{s_k r_j} \right), \quad (26)$$

where the sum over s_k covers all the relevant cylindrical waves in subsystem k . It is clear from Eq. (26) that the coupling loss factors satisfy reciprocity. In the case of two 1-D subsystems, the coupling loss factor is again given by Eq. (26), although in this case no summation is required, and s_k and r_j are replaced by k and j , respectively; this latter result coincides with the established literature for 1-D systems (Lyon and DeJong, 1995). The foregoing analysis does not assume that there are only two subsystems at the junction, but rather demonstrates that reciprocity holds between *any* two subsystems that are coupled at the junction.

III. WAVE PROPERTIES FOR A BEAM

In this section the various matrices that appear in Eq. (10) are derived for a component j that consists of a Euler-Bernoulli beam. The local coordinate system for the beam is shown in Fig. 2. The beam supports four propagating waves associated with bending in two orthogonal planes, torsion and extension, along with two evanescent waves associated with bending. The following analysis assumes that the shear center and neutral axis of the beam are coincident; however, the extension of this analysis to encompass an offset between the shear center and neutral axis is straightforward.

The displacement of the beam at the interface with the disc ($z=0$) is related to the amplitudes of the various out-

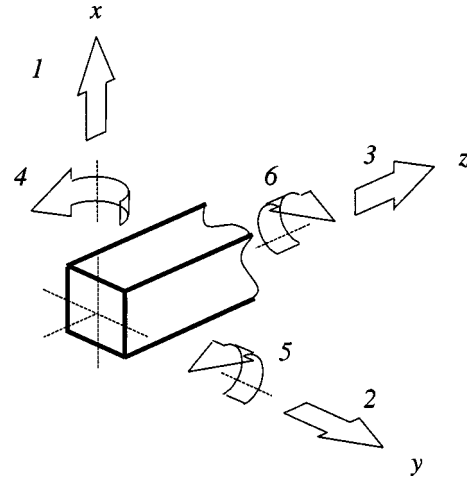


FIG. 2. Coordinate system and degrees of freedom numbering at the end of a beam component. A particular number, k say, shows the direction of the k th component of the displacement vector \mathbf{w}_j .

going watypes by Eq. (1). In this case the matrix \mathbf{Q}_{1j} that appears in this equation is 6×6 and has the entries (Cremer *et al.*, 1988)

$$\mathbf{Q}_{1j} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & ik_{by} & 0 & 0 & 0 & k_{bx} \\ -ik_{bx} & 0 & 0 & 0 & -k_{bx} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{bx} \\ a_{by} \\ a_e \\ a_t \\ a_{bxn} \\ a_{byn} \end{bmatrix}, \quad (27)$$

where the various wavenumbers are given by

$$k_t^2 = \omega^2 \frac{\rho J}{GQ}, \quad k_e^2 = \omega^2 \frac{\rho}{E}, \quad k_{bx}^4 = \omega^2 \frac{m}{B_x}, \quad k_{by}^4 = \omega^2 \frac{m}{B_y}. \quad (28)$$

Here G is the shear modulus, E is the Young's modulus, ρ is the mass per unit volume, m is the mass per unit length, J is the polar moment of area, GQ is the torsional rigidity, and B_y and B_x are the bending rigidities about the x and y axes, respectively. In the notation adopted here the subscript bx or by denotes a propagating bending wave which causes motion in the x or y directions, respectively, the subscript bxn or byn denotes an evanescent bending wave, and the subscripts e and t denote extensional (axial) and torsional waves. The matrix \mathbf{Q}_{2j} is not required for a beam, so that Eq. (27) gives the complete matrix \mathbf{Q}_1 that appears in Eq. (3); also in this case $\mathbf{R}_j = \mathbf{0}$ in Eq. (4). Taking the inverse of Eq. (27) gives the following expression for the matrix \mathbf{P}_j :

$$\mathbf{P}_j = \begin{bmatrix} (1+i)/2 & 0 & 0 & 0 & (1+i)/2k_{bx} & 0 \\ 0 & (1+i)/2 & 0 & -(1+i)/2k_{by} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ (1-i)/2 & 0 & 0 & 0 & -(1+i)/2k_{bx} & 0 \\ 0 & (1-i)/2 & 0 & (1+i)/2k_{by} & 0 & 0 \end{bmatrix}. \quad (29)$$

From simple beam theory, the bending moment and shear force associated with a bending wave propagating in the positive z direction with displacement u are given by

$$M = -B \frac{\partial^2 u}{\partial z^2}, \quad (30)$$

$$F = B \frac{\partial^3 u}{\partial z^3}, \quad (31)$$

where the positive directions of the previous quantities are illustrated in Fig. 3 for the x - z plane. The matrix \mathbf{S}_j in Eq. (6) is then given by:

$$\begin{bmatrix} f_u \\ f_v \\ f_w \\ M_{\theta_x} \\ M_{\theta_y} \\ M_{\theta_z} \end{bmatrix} = \mathbf{S}_j \mathbf{a}_j = \begin{bmatrix} iB_x k_{bx}^3 & 0 & 0 & 0 & -B_x k_{bx}^3 & 0 \\ 0 & iB_y k_{by}^3 & 0 & 0 & 0 & -B_y k_{by}^3 \\ 0 & 0 & ik_e EA & 0 & 0 & 0 \\ 0 & -B_y k_{by}^2 & 0 & 0 & 0 & B_y k_{by}^2 \\ B_x k_{bx}^2 & 0 & 0 & 0 & -B_x k_{bx}^2 & 0 \\ 0 & 0 & 0 & ik_t GQ & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{bx} \\ a_{by} \\ a_e \\ a_t \\ a_{bxn} \\ a_{byn} \end{bmatrix}. \quad (32)$$

It then follows from Eq. (7) that the dynamic stiffness matrix is given by

$$\mathbf{D}_j = \begin{bmatrix} (i-1)B_x k_{bx}^3 & 0 & 0 & 0 & iB_x k_{bx}^2 & 0 \\ 0 & (i-1)B_y k_{by}^3 & 0 & -iB_y k_{by}^2 & 0 & 0 \\ 0 & 0 & iE A k_e & 0 & 0 & 0 \\ 0 & -iB_y k_{by}^2 & 0 & (1+i)B_y k_{by} & 0 & 0 \\ iB_x k_{bx}^2 & 0 & 0 & 0 & (1+i)B_x k_{bx} & 0 \\ 0 & 0 & 0 & 0 & 0 & iGQ k_t \end{bmatrix}, \quad (33)$$

which is consistent with the expressions for the beam impedance given by Cremer *et al.* (1988). The vector \mathbf{w}_{in} that appears in Eq. (10) can be found for any particular incoming wave by selecting the relevant column of the matrix \mathbf{Q}_{1j} and replacing k with $-k$ for each wavenumber. Similarly, the vector \mathbf{F}_{in} can be found by selecting the relevant column of \mathbf{S}_j and again replacing k with $-k$.

IV. WAVE PROPERTIES FOR A THIN PLATE

A. Wave motion in the plate

This section derives the various detailed terms that are required to include a thin plate as one of the structural components in a point connected junction. The point coupling is taken to be attached to a rigid disc that is embedded in the plate, as shown in Fig. 4. The waves that will be generated by motion of the disk are all cylindrical waves: bending waves for the case of the out-of-plane motion w and the rotations θ_x and θ_y , shear and longitudinal waves for the case of the in-plane motions u and v , and shear waves for the

rotation θ_z . The total out-of-plane motion can be written as the sum of six waves as follows [see, for example, Leissa, 1993, Eq. (1.18) with the Bessel functions and modified Bessel functions of the first and second kinds written in terms of Hankel functions]:

$$\begin{aligned} u_z(r, \theta) = & a_0^{(b)} H_0^{(2)}(k_b r) + a_{1c}^{(b)} H_1^{(2)}(k_b r) \cos \theta \\ & + a_{1s}^{(b)} H_1^{(2)}(k_b r) \sin \theta + a_0^{(bn)} H_0^{(2)}(-ik_b r) \\ & + a_{1c}^{(bn)} H_1^{(2)}(-ik_b r) \cos \theta \\ & + a_{1s}^{(bn)} H_1^{(2)}(-ik_b r) \sin \theta. \end{aligned} \quad (34)$$

Here $H_n^{(2)}$ is the Hankel function of the second kind of order n , r and θ are polar coordinates centered on the attachment

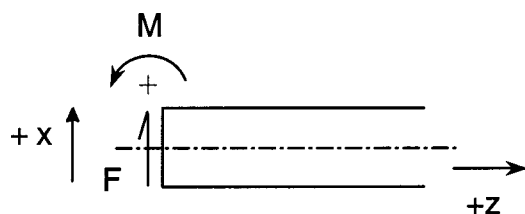


FIG. 3. Positive sign convention for shear force and bending moment in a beam.

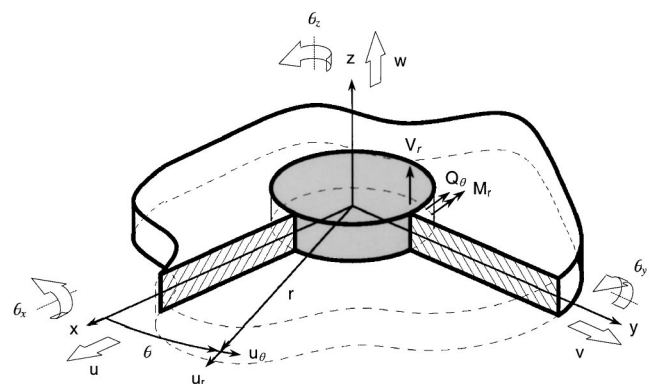


FIG. 4. Coordinate systems used to describe displacement of a disc embedded in a thin plate and the associated forces and moments.


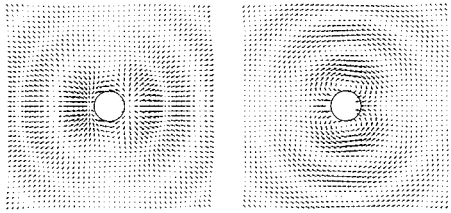

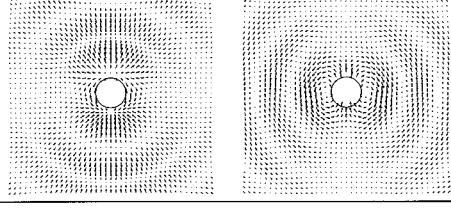

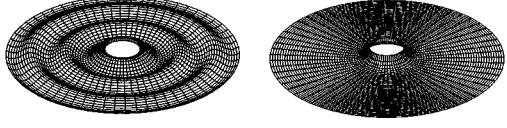
Disc DOF	Displacement field associated with outgoing wave	Wavetype
1 		$a_{1c}^{(e)}, a_{1s}^{(e)}$
2 		$a_{1s}^{(e)}, a_{1c}^{(e)}$
3 		$a_0^{(b)}, a_0^{(bn)}$

FIG. 5. Outgoing waves associated with translation of the disk.

disc, and k_b is the wavenumber for bending waves in the plate at the prescribed frequency ω , so that

$$k_b = (m/D)^{1/4} \omega^{1/2}, \quad (35)$$

where m is the mass per unit area and D is the flexural rigidity of the plate. The wave amplitudes $a_0^{(b)}$, $a_{1c}^{(b)}$, and $a_{1s}^{(b)}$ are associated with propagating waves that carry energy away from the junction, while the amplitudes $a_0^{(bn)}$, $a_{1c}^{(bn)}$, and $a_{1s}^{(bn)}$ are associated with evanescent waves that transmit no energy and decay away from the junction.

The in-plane waves that are generated by the motion of the disc can be expressed in terms a longitudinal wave potential ψ (sometimes referred to as the extensional wave potential) and a shear wave potential ϕ so that

$$u_r(r, \theta) = \frac{\partial \psi}{\partial r} + \frac{1}{r} \frac{\partial \phi}{\partial \theta}, \quad (36)$$

$$u_\theta(r, \theta) = \frac{1}{r} \frac{\partial \psi}{\partial \theta} - \frac{\partial \phi}{\partial r}, \quad (37)$$

where u_r and u_θ are, respectively, the radial and transverse components of the in-plane displacement. The wave potentials can be written as (see, for example, Morfey, 1994)

$$k_e \psi(r, \theta) = a_{1c}^{(e)} H_1^{(2)}(k_e r) \cos \theta + a_{1s}^{(e)} H_1^{(2)}(k_e r) \sin \theta, \quad (38)$$

$$k_s \phi(r, \theta) = a_0^{(s)} H_0^{(2)}(k_s r) + a_{1c}^{(s)} H_1^{(2)}(k_s r) \cos \theta + a_{1s}^{(s)} H_1^{(2)}(k_s r) \sin \theta. \quad (39)$$

Here $a_{1c}^{(e)}$ and $a_{1s}^{(e)}$ are longitudinal wave amplitudes, and $a_0^{(s)}$, $a_{1c}^{(s)}$ and $a_{1s}^{(s)}$ are shear wave amplitudes; all the in-plane waves are propagating, so that there are no evanescent components. The wavenumbers that appear in Eqs. (38) and (39) are defined as

$$k_e = \omega \sqrt{\rho(1 - \nu^2)/E}, \quad (40)$$

$$k_s = \omega \sqrt{\rho/G}, \quad (41)$$

where ρ is the density of the plate material, ν is the Poisson ratio, E is the Young's modulus, and G is the shear modulus.

It follows from Eqs. (34)–(41) that 11 wave amplitudes a_j are produced by motion of the connection point for a thin plate component j ; the displacement field associated with each of these components is shown in Figs. 5 and 6. Compatibility and equilibrium conditions must now be considered in order to derive the component matrices that appear in Eq. (12), and these conditions are detailed in what follows.

B. Compatibility conditions

1. Out-of-plane motion

Compatibility between the out-of-plane displacement of the edge of the disk and the motion of the plate u_z at $r=a$ leads to the condition

$$u_z(a, \theta) = w + a \theta_x \sin \theta - a \theta_y \cos \theta. \quad (42)$$

Applying this condition to Eq. (34) leads to three separate equations, since the coefficients of $\sin \theta$, $\cos \theta$, and the θ -independent terms can be equated independently. These conditions are to equivalent Eq. (1) for the out-of-plane motion of the plate, and three additional conditions in the form of Eq. (2) are required to fully determine the six bending wave amplitudes. These conditions can be derived by noting that the plate is rigidly clamped to the perimeter of the disc, so that the slope of the plate in the radial direction must match the corresponding slope of the disk. This yields

$$\frac{\partial u_z(a, \theta)}{\partial r} - \theta_x \sin \theta + \theta_y \cos \theta = 0, \quad (43)$$

and again this equation leads to three separate conditions. Equation (42) contributes to the terms in the matrix \mathbf{Q}_{1j} in Eq. (3), while Eq. (43) contributes to the terms in the matrix \mathbf{Q}_{2j} . Full details of these matrices for the whole plate (out-of-plane and in-plane motion) are given in the Appendix.

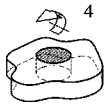
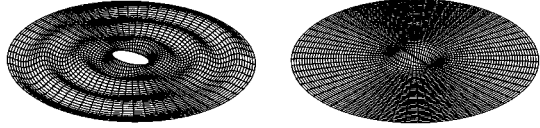
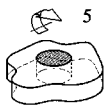
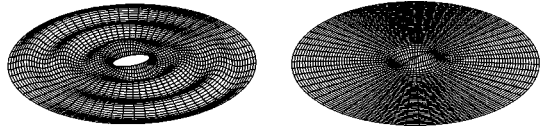
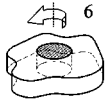
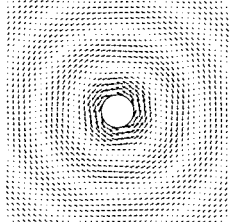
Disc DOF	Displacement field associated with outgoing wave	Wavetype
		$a_{1s}^{(b)}, a_{1s}^{(bn)}$
		$a_{1c}^{(b)}, a_{1c}^{(bn)}$
		$a_0^{(s)}$

FIG. 6. Outgoing waves associated with rotation of the disk.

2. In-plane motion

The compatibility conditions for in-plane motion have the form

$$u_r(a, \theta) = u \cos \theta + v \sin \theta, \quad (44)$$

$$u_\theta(a, \theta) = a \theta_z - u \sin \theta + v \cos \theta. \quad (45)$$

These equations lead to five separate conditions ($\sin \theta$, $\cos \theta$, and the θ -independent terms) on the five in-plane wave amplitudes. The first three of these conditions provide the relevant entries of the matrix \mathbf{Q}_{1j} , while the last two contribute to \mathbf{Q}_{2j} . Full details of the matrix \mathbf{Q}_j are given in the Appendix, together with the partition \mathbf{P}_j of the inverse matrix \mathbf{Q}_j^{-1} .

C. Elastic forces

1. Out-of-plane motion

The elastic forces provided by the disc due to out-of-plane motion consist of shear forces and bending moments acting on the disc perimeter. The Kirchhoff shear force per unit circumferential length can be written in the form (Leissa, 1993)

$$V_r(\theta) = D \left\{ \frac{\partial^3 u_z}{\partial r^3} + \frac{1}{a} \frac{\partial^2 u_z}{\partial r^2} - \frac{1}{a^2} \frac{\partial u_z}{\partial r} - \frac{(3-\nu)}{a^3} \frac{\partial^3 u_z}{\partial \theta^3} + \frac{(2-\nu)}{a^2} \frac{\partial^3 u_z}{\partial r \partial \theta^2} \right\}. \quad (46)$$

Similarly, the bending moment per unit circumferential length has the form (Leissa, 1993)

$$M_r(\theta) = D \left\{ \frac{\partial^2 u_z}{\partial r^2} + \frac{\nu}{a^2} \frac{\partial^2 u_z}{\partial \theta^2} + \frac{\nu}{a} \frac{\partial u_z}{\partial r} \right\}. \quad (47)$$

The net out-of-plane force and the net out-of-plane moments provided by the disk can thus be written as

$$M_x = a \int_0^{2\pi} (a V_r - M_r) \sin \theta \, d\theta, \quad (48)$$

$$M_y = -a \int_0^{2\pi} (a V_r - M_r) \cos \theta \, d\theta, \quad (49)$$

$$F_z = a \int_0^{2\pi} V_r \, d\theta. \quad (50)$$

Equation (34) can be inserted into Eqs. (46)–(50) to yield a relationship between the out-of-plane forces and the wave amplitudes in the form of Eq. (6). The resulting entries in the matrix \mathbf{S}_j for the complete plate are detailed in the Appendix.

2. In-plane motion

Assuming plane stress conditions, the in-plane stresses in the plate can be related to the in-plane displacements as follows (Graff, 1975):

$$\sigma_{rr} = \frac{E}{1-\nu^2} \left(\frac{\partial u_r}{\partial r} + \frac{\nu}{r} u_r + \frac{\nu}{r} \frac{\partial u_\theta}{\partial \theta} \right), \quad (51)$$

$$\sigma_{r\theta} = G \left(\frac{1}{r} \frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r} + \frac{\partial u_\theta}{\partial r} \right). \quad (52)$$

The resulting net forces provided by the disk are thus

$$F_x = -ha \int_0^{2\pi} (\sigma_{rr} \cos \theta - \sigma_{r\theta} \sin \theta) \, d\theta, \quad (53)$$

$$F_y = -ha \int_0^{2\pi} (\sigma_{rr} \sin \theta + \sigma_{r\theta} \cos \theta) \, d\theta, \quad (54)$$

$$M_z = -ha^2 \int_0^{2\pi} \sigma_{r\theta} \, d\theta, \quad (55)$$

where h is the plate thickness. Equations (36)–(39) can be substituted into Eqs. (51)–(55) to yield the in-plane forces in terms of the in-plane wave amplitudes. This leads to the

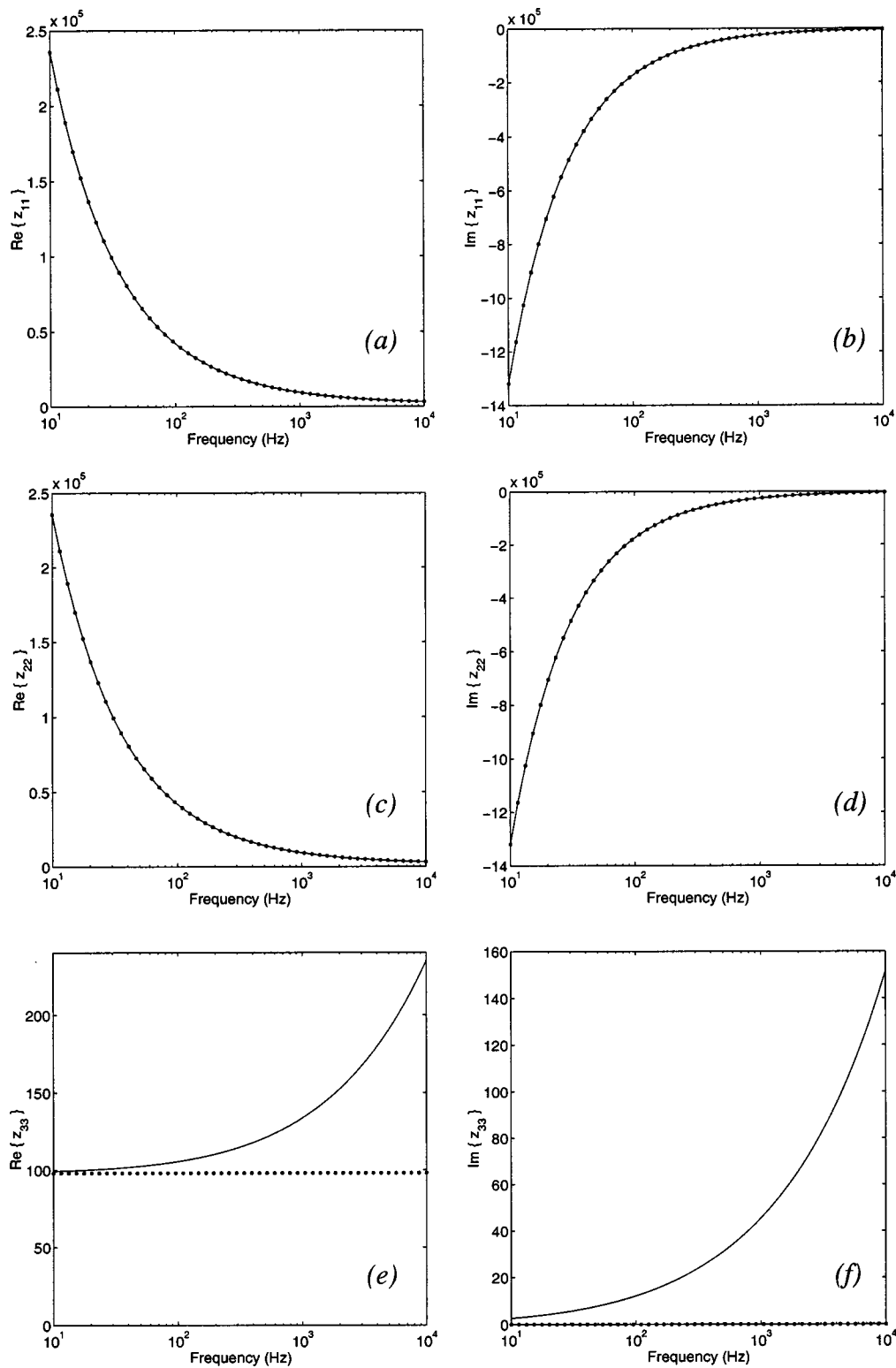


FIG. 7. Translational impedances for a disc with a 10-mm radius in a 1-mm-thick steel plate: (a) $\text{Re}\{Z_{11}\}$, (b) $\text{Im}\{Z_{11}\}$, (c) $\text{Re}\{Z_{22}\}$, (d) $\text{Im}\{Z_{22}\}$, (e) $\text{Re}\{Z_{33}\}$, and (f) $\text{Im}\{Z_{33}\}$. Solid line, predicted impedance; dotted line, impedance from the literature (see text for details).

in-plane components of the matrix \mathbf{S}_j that are detailed in the Appendix.

D. Dynamic stiffness matrix

The matrices \mathbf{S}_j and \mathbf{P}_j arising from the previous analysis can be used to find the dynamic stiffness matrix \mathbf{D}_j of the

plate via Eq. (7); the final form of this matrix is detailed in the Appendix. The impedance matrix $\mathbf{Z}_{\text{plate}}$ is closely related to the dynamic stiffness matrix, and can be written in the form

$$\mathbf{Z}_{\text{plate}} = \frac{1}{i\omega} \mathbf{D}_j. \quad (56)$$

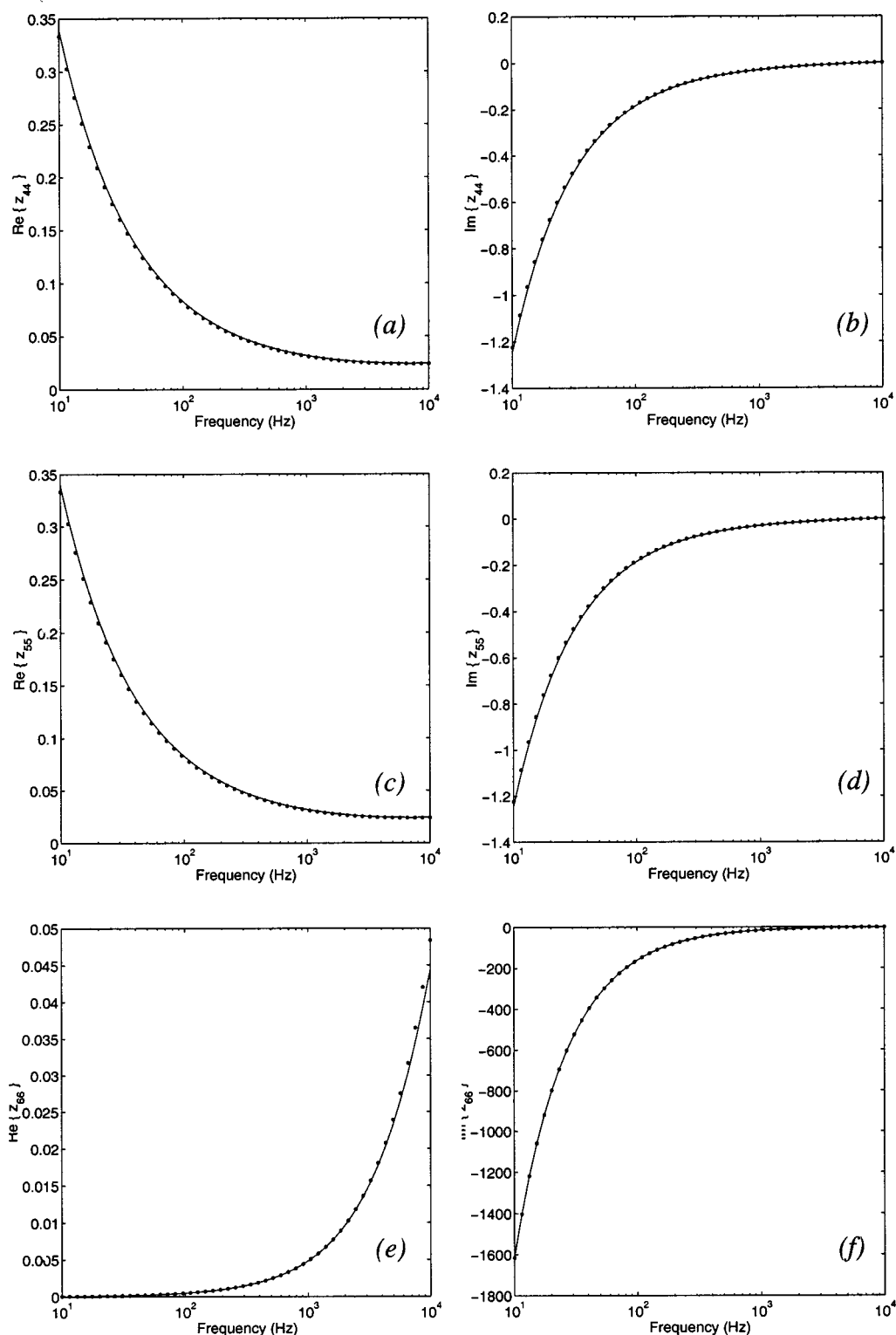


FIG. 8. Rotational impedances for a disc with a 10-mm radius in a 1-mm-thick steel plate: (a) $\text{Re}\{Z_{44}\}$, (b) $\text{Im}\{Z_{44}\}$, (c) $\text{Re}\{Z_{55}\}$, (d) $\text{Im}\{Z_{55}\}$, (e) $\text{Re}\{Z_{66}\}$, and (f) $\text{Im}\{Z_{66}\}$. Solid line, predicted impedance; dotted line, impedance from the literature (see text for details).

Various results for the point impedance of a plate are available in the literature, and these can be used to validate the present formulation, as described in what follows.

The real and imaginary parts of the diagonal entries of the impedance matrix of a disc with radius 10 mm contained within a 1-mm-thick steel plate ($E = 2.1 \times 10^{11} \text{ N/m}^2$, $\nu = 0.3$, $\rho = 7800 \text{ kg/m}^3$) are shown in Figs. 7 and 8. It can be

shown analytically that the present expressions for the in-plane translational impedances are in full agreement with Eq. (7) of Leung and Pinnington (1986); numerical confirmation of this agreement for the present example is shown in Figs. 7(a)–(d). The classical expression for the (transverse) point impedance of a thin plate is given by Eq. (4.63) of Cremer *et al.* (1988), and states that the impedance is real and fre-

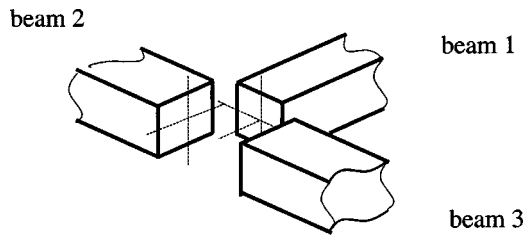


FIG. 9. T-junction between three beam subsystems.

quency independent. This expression is only valid when the size of the disk is small compared with the bending wavelength of the plate; it can be shown analytically that the present result reduces to the standard expression as the size of the disk tends to zero. For the current numerical example $k_b a = 1$ at 2.5 kHz and $k_b a = 2$ at 10 kHz. Figures 7(e) and (f) compare the present result for the transverse translational component of the impedance with that given by Cremer *et al.* (1988). The increased impedance due to the finite size

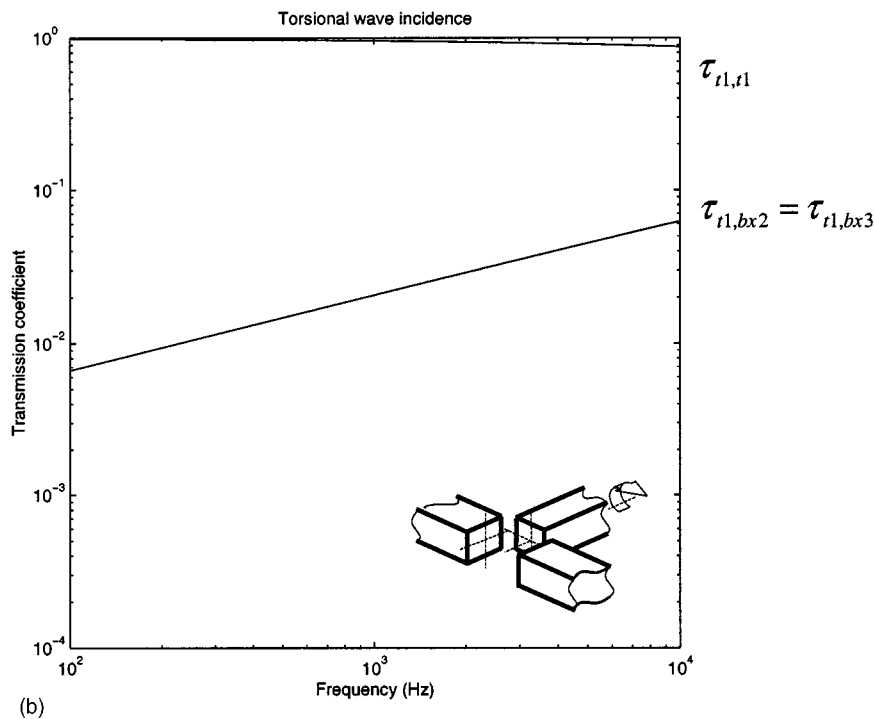
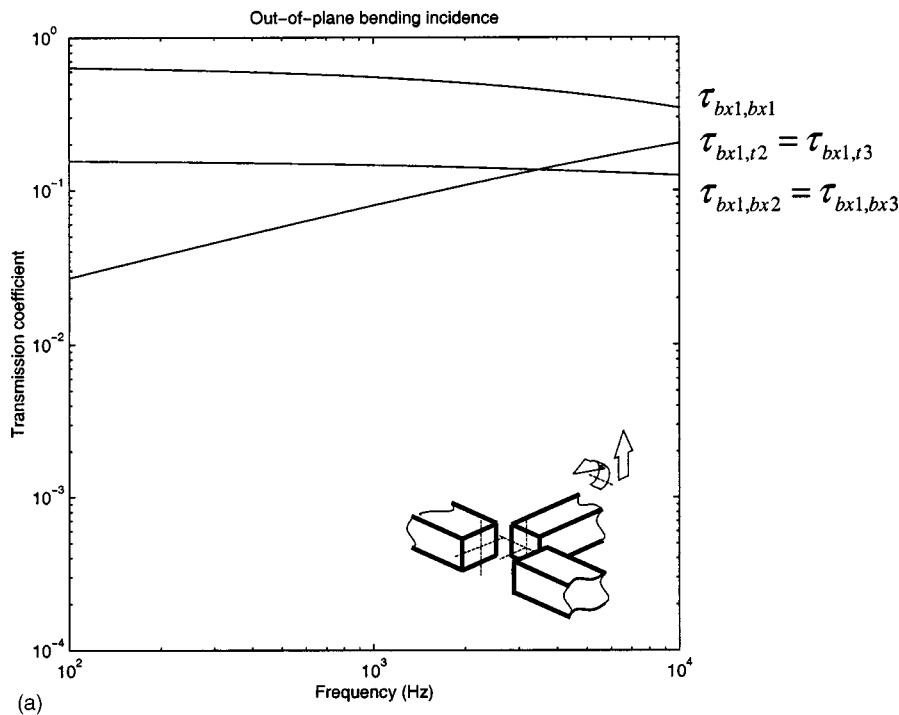


FIG. 10. Transmission coefficients for (a) incident out of plane bending wave and (b) incident torsional wave in beam 1. These figures match Fig. 3 of Moore (1990).

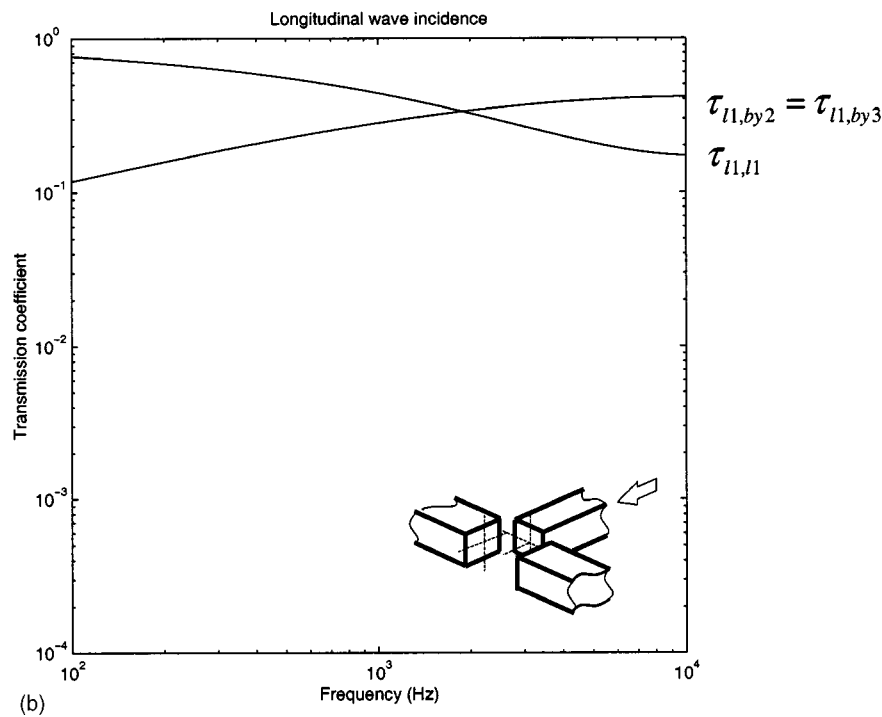
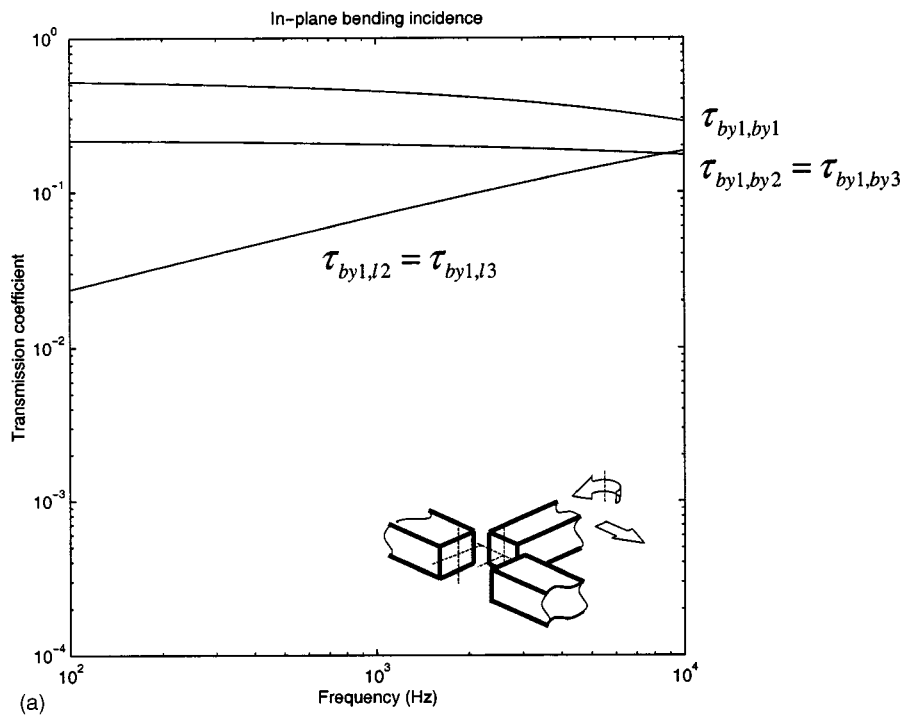


FIG. 11. Transmission coefficients for (a) incident in-plane bending wave and (b) incident longitudinal wave in subsystem 1. These figures match Fig. 5 of Moore (1990).

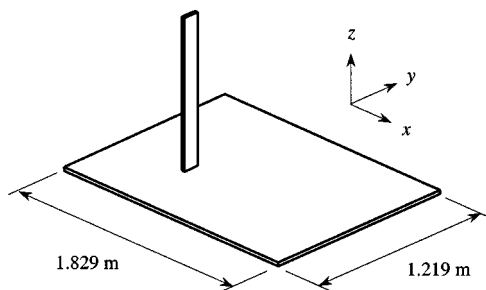


FIG. 12. Beam-plate structure used by Lyon and Eichler (1964).

of the disk is clearly visible with increasing frequency.

Dyer (1960) has provided a number of expressions for the moment impedance of a thick plate. Figures 8(a)–(d) compare the present prediction of the rotational components of the impedance associated with bending to those given by Eq. (18) of Dyer (1960)—due account has been taken of the fact that a time variation of $\exp(-i\omega t)$ has been used by Dyer rather than the variation $\exp(i\omega t)$ adopted in the current analysis. Clearly there is very good agreement between the two sets of results.

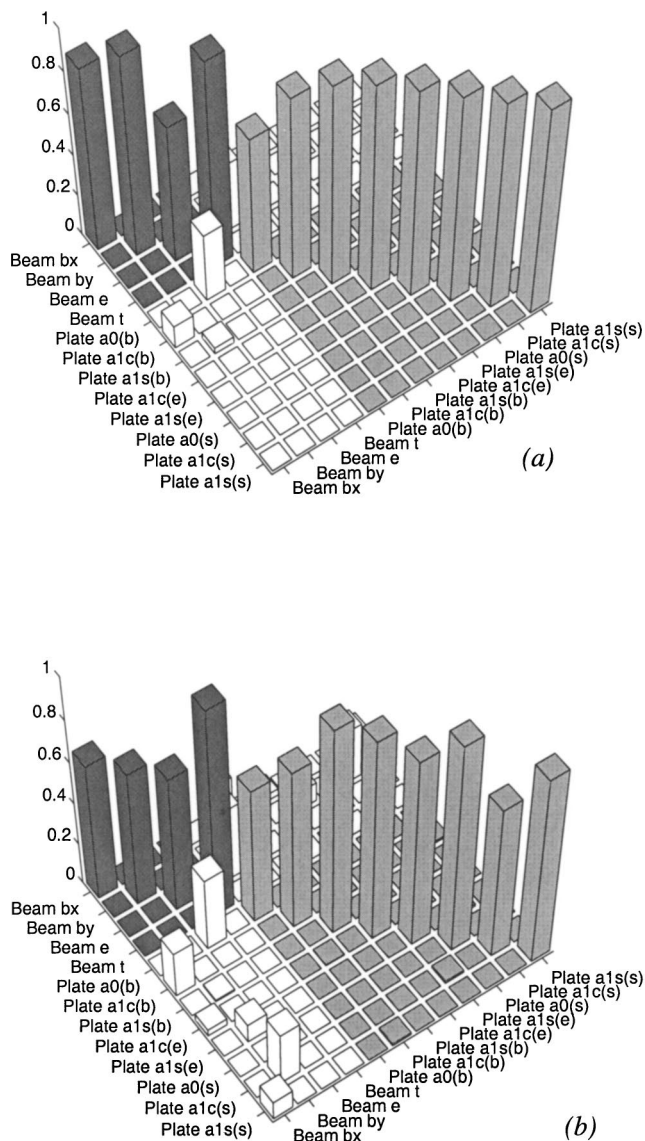


FIG. 13. Transmission matrix for beam-plate junction at (a) 100 Hz and (b) 10 kHz.

Ljunggren (1984) has provided various expressions for the torsional impedance of a plate derived by considering the equations of motion of a three-dimensional continuum. Figures 8(e) and (f) compare the present prediction with that given by Eq. (40) of Ljunggren (1984), which is an approximate expression that is valid for small values of $k_s a$. There is very close agreement between the two sets of results.

In a practical problem, the disk size a should be selected to match the physical dimensions of the appropriate substructures, such as the cross section of an adjoining beam. The results shown in Figs. 7 and 8 demonstrate that for Z_{33} the impedance is insensitive to the disk size providing the disk is small relative to the plate wavelength, while for the other components the real part of the impedance becomes singular as the disk size tends to zero, in agreement with the established literature.

E. Incident waves

The terms required to describe the effects of an incoming wave in Eq. (10) can readily be deduced by noting that

these waves have the same form as those included in Eqs. (34) and (38)–(39), apart from the fact that the Hankel functions of the second kind $H_n^{(2)}$ are replaced with those of the first kind $H_n^{(1)}$. The resulting terms \mathbf{w}_{in} and $\mathbf{R}_j \mathbf{q}_{in}$ are detailed in the Appendix. This completes the analysis required to include a thin plate as a structural component at a point connected junction.

V. NUMERICAL EXAMPLES

The following section provides a number of numerical examples which serve to validate the current derivation. Initially, the proposed methodology is validated by computing the transmission coefficients between a number of coupled beams and comparing the results with previous studies in the literature. Second, the transmission coefficients and coupling loss factors of a simple beam-plate structure are calculated and discussed.

A. Beam validation examples

There are numerous studies in the literature related to the calculation of the reflection and transmission coefficients of coupled beams (for example, Cremer and Heckl, 1998; Sablik, 1982; Moore, 1990; Horner and White, 1991; Guo, 1995; Tso and Norwood, 1995). The results of the previous sections are compared against an example initially presented by Moore (1990) and also validated by Tso and Norwood (1995). Three (thin) beams whose neutral axes lie in the same plane are coupled at a T-junction as illustrated in Fig. 9. The properties of the beams are $E = 7.2 \times 10^{10} \text{ N/m}^2$, $\rho = 2710 \text{ kg/m}^3$, $G = 2.69 \times 10^{10} \text{ N/m}^2$, $GQ = 3.5 \times 10^3 \text{ Nm}^2$, $A = 4.0 \times 10^{-4} \text{ m}^2$, $J = 2.83 \times 10^{-6} \text{ m}^4$, $I_{xx} = 3.48 \times 10^{-8} \text{ m}^4$, and $I_{yy} = 2.8 \times 10^{-6} \text{ m}^4$, with the y axis lying in the plane of the junction. Beam 1 is excited by a given incident wave and the transmission coefficients are plotted as a function of frequency. The computed transmission coefficients are illustrated in Figs. 10 and 11 and are in agreement with Figs. 3 and 5 of Moore (1990).

B. Plate-beam junction

Consider the beam-plate structure illustrated in Fig. 12. The physical and material properties of the beam and plate are plate thickness 3.175 mm, beam width in x direction 3.175 mm, beam width in y direction 38.1 mm, beam length 1.3208 m, $E = 7.21 \times 10^{10} \text{ N/m}^2$, $\rho = 2710 \text{ kg/m}^3$, and $\nu = 0.3$. This example was used by Lyon and Eichler (1964) in one of the first predictions (and measurements) of a SEA coupling loss factor between two coupled structures. When predicting the coupling loss factor between the beam and plate Lyon and Eichler (1964) assumed that the dominant motion of each component is associated with a single bending wavefield and that the junction supports only rotational motion in one degree of freedom. For the special case in which the phase speed, density, and bending stiffness associated with bending waves in the beam and plate are the same, the predicted coupling loss factor from the beam to the plate is given by

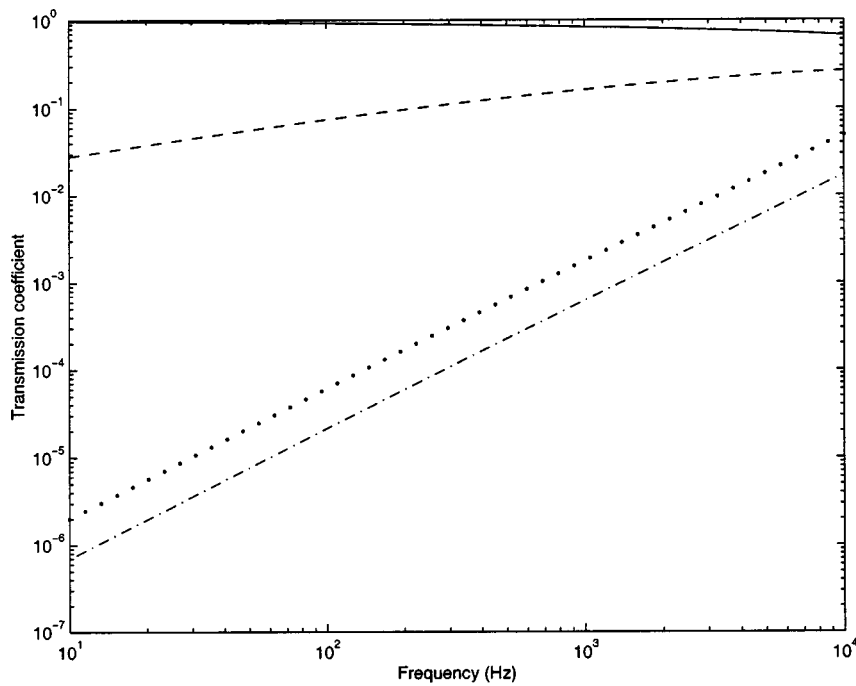


FIG. 14. Transmission coefficients associated with incident bending wave in beam: — $\tau_{bx,bx}$; - - - $\tau_{bx,a_{1c}^{(b)}}$; ··· $\tau_{bx,a_{1s}^{(s)}}$; - · - $\tau_{bx,a_{1c}^{(e)}}$.

$$\eta_{bp}^{(\text{Lyon})} = \frac{W}{4L}, \quad (57)$$

where W and L are the width and length of the beam, respectively.

The transmission coefficients associated with the beam-plate junction can be computed by using the present analysis method. In the numerical example presented here the radius of the rigid disc that connects the beam and plate is assumed to equal one-half of beam cross section in the x direction. While the choice of the disc size is perhaps somewhat arbitrary in this example due to the differing cross-sectional dimensions of the beam, the resulting transmission coefficients are fairly insensitive to perturbations in the size of the disc.

Figure 13 shows the entries of the transmission matrix associated with the propagating wavetypes in the beam and plate at 500 Hz and 1 kHz. The resulting transmission matrix is symmetric and the rows and columns sum to unity as expected. It can be seen that the strongest coupling occurs between the extensional wave in the beam and the $a_0^{(b)}$ bending wave in the plate. There is also some coupling between the bending wavefields of the beam and the $a_{1c}^{(b)}$ and $a_{1s}^{(b)}$ bending wavefields in the plate. Figure 14 shows the (non-zero) entries of the transmission matrix associated with an incident bending wave in the beam (with motion in the x direction), as a function of frequency. The incident bending wave in the beam is scattered into the plate waves associated

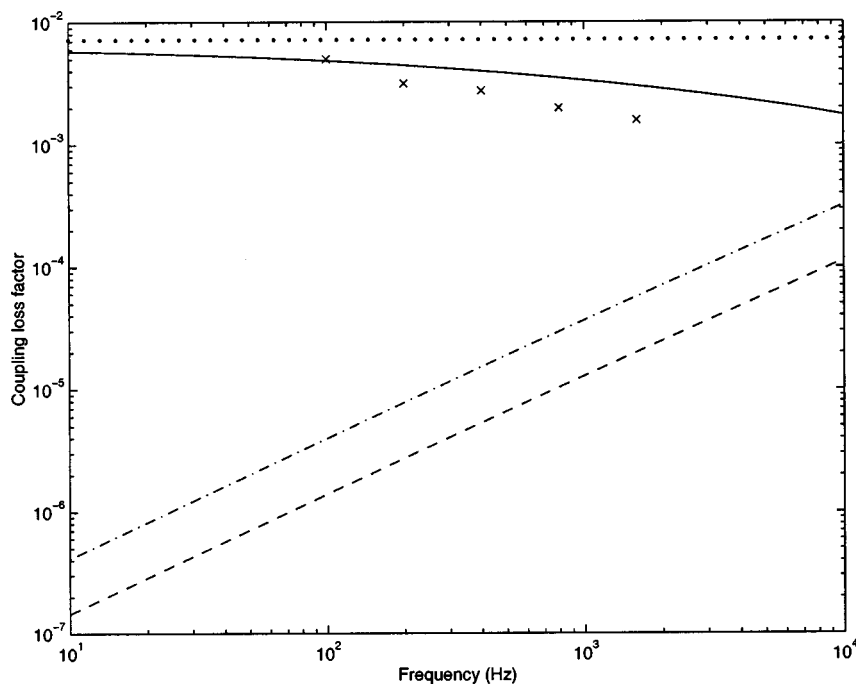


FIG. 15. Coupling loss factors between the bending- x subsystem in the beam and various subsystems of the plate: — plate bending subsystem; - - - plate shear subsystem; - · - plate extensional subsystem; ··· predicted CLF from Lyon and Eichler (1964); × experimentally measured CLF from Lyon and Eichler (1964).

with in-plane displacement in the x direction and bending rotation of the disc about the y axis as expected.

A SEA model of the system can then be created in which the beam substructure is described by four SEA subsystems associated with bending waves in the x and y directions, extensional waves, and torsional waves. The plate substructure can be described by three SEA subsystems associated with bending, extension, and shear. The coupling loss factors between the various SEA subsystems can then be computed using Eqs. (21), (25), and (26). Figure 15 shows the coupling loss factors between the beam subsystem associated with bending motion in the x direction and the various plate subsystems. The CLF predicted by Lyon and Eichler is also plotted along with an experimentally measured CLF (obtained from Fig. 6 of Lyon and Eichler, 1964). It is interesting to note that the finite impedance of the in-plane degrees of freedom of the plate, included in the current analysis, provides some explanation for the discrepancies between the predicted and measured CLF's in Lyon and Eichler's original analysis.

VI. CONCLUDING REMARKS

A general methodology has been presented for computing the wave transmission coefficients and coupling loss fac-

tors between components that are coupled at a point. The work builds upon an earlier study of cylindrical wave motion (Langley and Shorter, 2001), and it has been shown that in all cases the coupling loss factors yielded by the present approach satisfy reciprocity. Full details of the matrices required to include thin plate and beam components have been given and a number of examples have been presented. It can be noted that the general methodology is not limited to structures of this type, and the inclusion of other types of component can be accomplished by computing (via a study of wave motion in the component) the relevant matrices and vectors for use in Eq. (11).

ACKNOWLEDGMENTS

This research has been funded by Vibro-Acoustic Sciences, Inc; and helpful discussions with Tom Burton, Paul Bremner, and Bryce Gardner are gratefully acknowledged.

APPENDIX: DETAILS OF PLATE MATRICES

1. Displacement related matrices

The matrix \mathbf{Q}_{1j} is defined by Eq. (1). For a thin plate this equation takes the form

$$\begin{bmatrix} u \\ v \\ w \\ \theta_x \\ \theta_y \\ \theta_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & B_3 & 0 & 0 & 0 & B_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & B_3 & 0 & -B_4 & 0 & 0 & 0 & 0 \\ B_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_{1n} & 0 & 0 \\ 0 & 0 & B_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_{2n} \\ 0 & -B_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -B_{2n} & 0 \\ 0 & 0 & 0 & 0 & 0 & B_5 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0^{(b)} \\ a_{1c}^{(b)} \\ a_{1s}^{(b)} \\ a_{1c}^{(e)} \\ a_{1s}^{(e)} \\ a_0^{(s)} \\ a_{1c}^{(s)} \\ a_{1s}^{(s)} \\ a_0^{(bn)} \\ a_{1c}^{(bn)} \\ a_{1s}^{(bn)} \end{bmatrix}, \quad (\text{A1})$$

where

$$B_1 = H_0^{(2)}(k_b a), \quad B_{1n} = H_0^{(2)}(-ik_b a), \quad B_2 = \frac{1}{a} H_1^{(2)}(k_b a), \quad B_{2n} = \frac{1}{a} H_1^{(2)}(-ik_b a), \quad (\text{A2})$$

$$B_3 = H_1^{(2)'}(k_e a), \quad B_4 = \frac{1}{k_s a} H_1^{(2)}(k_s a), \quad B_5 = -\frac{1}{a} H_0^{(2)'}(k_s a).$$

The matrix \mathbf{Q}_{2j} is defined by Eq. (2). It follows from Eqs. (43) and (45) that for a thin plate the matrix has the form

$$\mathbf{Q}_{2j} = \begin{bmatrix} B_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_{6n} & 0 & 0 \\ 0 & 0 & B_7 - B_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_{7n} - B_{2n} \\ 0 & B_7 - B_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_{7n} - B_{2n} & 0 \\ 0 & 0 & 0 & B_8 + B_3 & 0 & 0 & 0 & B_9 + B_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -B_8 - B_3 & 0 & B_9 + B_4 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (\text{A3})$$

where

$$\begin{aligned} B_6 &= k_b H_0^{(2)'}(k_b a), \quad B_{6n} = -ik_b H_0^{(2)'}(-ik_b a), \\ B_7 &= k_b H_1^{(2)'}(k_b a), \quad B_{7n} = -ik_b H_1^{(2)'}(-ik_b a), \quad (A4) \\ B_8 &= -\frac{1}{k_e a} H_1^{(2)}(k_e a), \quad B_9 = -H_1^{(2)'}(k_s a). \end{aligned}$$

By inverting the total matrix \mathbf{Q}_j , the matrix \mathbf{P}_j that appears in Eq. (5) can be shown to be

$$\mathbf{P}_j = \begin{bmatrix} 0 & 0 & A_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_2 & 0 \\ 0 & 0 & 0 & -A_2 & 0 & 0 \\ A_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & A_5 \\ 0 & A_4 & 0 & 0 & 0 & 0 \\ -A_4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A_{1n} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_{2n} & 0 \\ 0 & 0 & 0 & -A_{2n} & 0 & 0 \end{bmatrix}, \quad (A5)$$

where

$$\begin{aligned} A_1 &= \frac{H_0^{(2)'}(-ik_b a)}{H_0^{(2)}(k_b a)H_0^{(2)'}(-ik_b a) - iH_0^{(2)}(-ik_b a)H_0^{(2)'}(k_b a)}, \\ A_2 &= \frac{-H_1^{(2)}(-ik_b a) - ik_b a H_1^{(2)'}(-ik_b a)}{ik_b H_1^{(2)}(k_b a)H_1^{(2)'}(-ik_b a) + k_b H_1^{(2)}(-ik_b a)H_1^{(2)'}(k_b a)}, \\ A_3 &= \frac{k_s k_e a^2 H_1^{(2)'}(k_s a) - k_e a H_1^{(2)}(k_s a)}{k_s k_e a^2 H_1^{(2)'}(k_s a)H_1^{(2)'}(k_e a) - H_1^{(2)}(k_s a)H_1^{(2)}(k_e a)}, \\ A_4 &= \frac{-k_s k_e a^2 H_1^{(2)'}(k_e a) + k_s a H_1^{(2)}(k_e a)}{k_s k_e a^2 H_1^{(2)'}(k_s a)H_1^{(2)'}(k_e a) - H_1^{(2)}(k_s a)H_1^{(2)}(k_e a)}, \quad (A6) \\ A_5 &= -\frac{a}{H_0^{(2)'}(k_s a)}, \\ A_{1n} &= \frac{-iH_0^{(2)'}(k_b a)}{H_0^{(2)}(k_b a)H_0^{(2)'}(-ik_b a) - iH_0^{(2)}(-ik_b a)H_0^{(2)'}(k_b a)}, \\ A_{2n} &= \frac{iH_1^{(2)}(k_b a) - ik_b a H_1^{(2)'}(k_b a)}{ik_b H_1^{(2)}(-ik_b a)H_1^{(2)'}(k_b a) - k_b H_1^{(2)}(k_b a)H_1^{(2)'}(-ik_b a)}. \end{aligned}$$

Now Eq. (10) contains a term on the right-hand side that arises from the presence of an incident wave and reads $\mathbf{P}_j \mathbf{w}_{in} + \mathbf{R}_j \mathbf{q}_{in}$. This vector can be derived for each particular incident wavetype, and it can be shown that the results correspond to the columns in the following matrix:

$$\begin{bmatrix} a_0^{(b)} & a_{1c}^{(b)} & a_{1s}^{(b)} & a_{1c}^{(e)} & a_{1s}^{(e)} & a_0^{(s)} & a_{1c}^{(s)} & a_{1s}^{(s)} \\ C_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & C_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & C_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & C_3 & 0 & 0 & 0 & -C_6 \\ 0 & 0 & 0 & 0 & C_3 & 0 & C_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & C_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & -C_4 & 0 & C_7 & 0 \\ 0 & 0 & 0 & C_4 & 0 & 0 & 0 & C_7 \\ C_{1n} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & C_{2n} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & C_{2n} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (A7)$$

where

$$\begin{aligned} C_1 &= \frac{H_0^{(2)'}(-ik_b a)H_0^{(1)}(k_b a) - iH_0^{(1)'}(k_b a)H_0^{(2)}(-ik_b a)}{H_0^{(2)'}(-ik_b a)H_0^{(2)}(k_b a) - iH_0^{(2)'}(k_b a)H_0^{(2)}(-ik_b a)}, \\ C_{1n} &= \frac{H_0^{(2)'}(k_b a)H_0^{(1)}(k_b a) - H_0^{(1)'}(k_b a)H_0^{(2)}(k_b a)}{H_0^{(2)'}(k_b a)H_0^{(2)}(-ik_b a) + iH_0^{(2)'}(-ik_b a)H_0^{(2)}(k_b a)}, \\ C_2 &= \frac{iH_1^{(2)'}(-ik_b a)H_1^{(1)}(k_b a) + H_1^{(1)'}(k_b a)H_1^{(2)}(-ik_b a)}{H_1^{(2)'}(k_b a)H_1^{(2)}(-ik_b a) + iH_1^{(2)'}(-ik_b a)H_1^{(2)}(k_b a)}, \\ C_{2n} &= \frac{H_1^{(1)}(k_b a)H_1^{(2)'}(k_b a) - H_1^{(2)}(k_b a)H_1^{(1)'}(k_b a)}{iH_1^{(2)}(k_b a)H_1^{(2)'}(-ik_b a) + H_1^{(2)}(-ik_b a)H_1^{(2)'}(k_b a)}, \\ C_3 &= \frac{H_1^{(1)}(k_e a)H_1^{(2)}(k_s a) - a^2 k_e k_s H_1^{(1)'}(k_e a)H_1^{(2)'}(k_s a)}{H_1^{(2)}(k_e a)H_1^{(2)}(k_s a) - a^2 k_e k_s H_1^{(2)'}(k_e a)H_1^{(2)'}(k_s a)}, \quad (A8) \\ C_4 &= \frac{a k_e H_1^{(2)'}(k_e a)H_1^{(1)}(k_e a) - a k_e H_1^{(1)'}(k_e a)H_1^{(2)}(k_e a)}{-H_1^{(2)}(k_e a)H_1^{(2)}(k_s a) + a^2 k_e k_s H_1^{(2)'}(k_e a)H_1^{(2)'}(k_s a)}, \\ C_5 &= \frac{H_0^{(1)'}(k_s a)}{H_0^{(2)'}(k_s a)}, \\ C_6 &= \frac{k_s a H_1^{(1)'}(k_s a)H_1^{(2)}(k_s a) - k_s a H_1^{(2)'}(k_s a)H_1^{(1)}(k_s a)}{a^2 k_e k_s H_1^{(2)'}(k_e a)H_1^{(2)'}(k_s a) - H_1^{(2)}(k_e a)H_1^{(2)}(k_s a)}, \\ C_7 &= \frac{H_1^{(1)}(k_s a)H_1^{(2)}(k_e a) - k_e k_s a^2 H_1^{(1)'}(k_s a)H_1^{(2)'}(k_e a)}{H_1^{(2)}(k_e a)H_1^{(2)}(k_s a) - k_e k_s a^2 H_1^{(2)'}(k_e a)H_1^{(2)'}(k_s a)}. \end{aligned}$$

The top row in Eq. (A7) shows the type of incident wave, while the column below shows the appropriate vector $\mathbf{P}_j \mathbf{w}_{in} + \mathbf{R}_j \mathbf{q}_{in}$.

2. Force related matrices

It can be shown from Eqs. (34), (36)–(39), and (46)–(55) that the relation between the wave amplitudes and the elastic forces supplied by the disk is as follows:

$$\begin{bmatrix} f_x \\ f_y \\ f_z \\ M_x \\ M_y \\ M_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & X_3 & 0 & 0 & 0 & X_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & X_3 & 0 & -X_4 & 0 & 0 & 0 & 0 \\ X_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X_{1n} & 0 & 0 \\ 0 & 0 & X_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X_{2n} \\ 0 & -X_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -X_{2n} & 0 \\ 0 & 0 & 0 & 0 & 0 & X_5 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0^{(b)} \\ a_{1c}^{(b)} \\ a_{1s}^{(b)} \\ a_{1c}^{(e)} \\ a_{1s}^{(e)} \\ a_0^{(s)} \\ a_{1c}^{(s)} \\ a_{1s}^{(s)} \\ a_0^{(bn)} \\ a_{1c}^{(bn)} \\ a_{1s}^{(bn)} \end{bmatrix}. \quad (A9)$$

Here

$$\begin{aligned} X_1 &= 2\pi D k_b^3 a \left[-\frac{1}{k_b^2 a^2} H_0^{(2)'}(k_b a) + \frac{1}{k_b a} H_0^{(2)''}(k_b a) \right. \\ &\quad \left. + H_0^{(2)'''}(k_b a) \right], \\ X_{1n} &= 2\pi D k_b^3 a \left[\frac{i}{k_b^2 a^2} H_0^{(2)'}(-ik_b a) - \frac{1}{k_b a} H_0^{(2)''}(-ik_b a) \right. \\ &\quad \left. + i H_0^{(2)'''}(-ik_b a) \right], \\ X_2 &= \pi B k_b^3 a^2 \left[\frac{3}{k_b^3 a^3} H_1^{(2)}(k_b a) - \frac{3}{k_b^2 a^2} H_1^{(2)'}(k_b a) \right. \\ &\quad \left. + H_1^{(2)'''}(k_b a) \right], \\ X_{2n} &= \pi B k_b^3 a^2 \left[\frac{3}{k_b^3 a^3} H_1^{(2)}(-ik_b a) + \frac{3i}{k_b^2 a^2} H_1^{(2)'}(-ik_b a) \right. \\ &\quad \left. + i H_1^{(2)'''}(-ik_b a) \right], \\ X_3 &= \frac{E h \pi}{(1-\nu^2) k_e a} [H_1^{(2)}(k_e a) - k_e a H_1^{(2)'}(k_e a) \\ &\quad - k_e^2 a^2 H_1^{(2)''}(k_e a)], \\ X_4 &= \frac{G h \pi}{k_s a} [H_1^{(2)}(k_s a) - k_s a H_1^{(2)'}(k_s a) \\ &\quad - k_s^2 a^2 H_1^{(2)''}(k_s a)], \\ X_5 &= 2G h \pi a [-H_0^{(2)'}(k_s a) + k_s a H_0^{(2)''}(k_s a)]. \end{aligned} \quad (A10)$$

The matrix that appears in Eq. (A9) is the matrix \mathbf{S}_j of Eq. (6). The dynamic stiffness matrix of the plate is given by $\mathbf{D}_j = \mathbf{S}_j \mathbf{P}_j$, and it follows from Eqs. (A5) and (A9) that the matrix is diagonal with the following entries:

$$\begin{aligned} D_{j11} &= D_{j22} = X_3 A_3 - X_4 A_4, \\ D_{j33} &= X_1 A_1 + X_{1n} A_{1n}, \end{aligned} \quad (A11)$$

$$D_{j44} = D_{j55} = -X_2 A_2 - X_{2n} A_{2n},$$

$$D_{j66} = X_5 A_5.$$

The final item required to complete the description of the plate component is the force vector \mathbf{F}_{in} that appears in Eq. (9). This vector can readily be found by identifying the relevant column of the matrix in Eq. (A9) and replacing all the Hankel functions of the second kind $H_n^{(2)}$ with those of the first kind $H_n^{(1)}$.

- Bosmans, I., and Nightingale, T. (2001). "Modeling vibrational energy transmission at bolted junctions between a plate and a stiffening rib," *J. Acoust. Soc. Am.* **109**, 999–1010.
- Bosmans, I., and Vermeir, G. (1997). "Diffuse transmission of structure-borne sound at periodic junctions of semi-infinite plates," *J. Acoust. Soc. Am.* **101**, 3443–3456.
- Cremer, L., Heckl, M., and Ungar, E. E. (1988). *Structure-borne Sound*, Second Edition (Springer-Verlag, Berlin).
- Dyer, I. (1960). "Moment impedances of plates," *J. Acoust. Soc. Am.* **32**, 1290–1297.
- Graff, K. (1975). *Wave Motion in Elastic Solids* (Dover, New York).
- Guo, Y. (1995). "Flexural wave transmission through angled structural joints," *J. Acoust. Soc. Am.* **97**, 289–297.
- Heron, K. H. (1999). "Predictive SEA using line wave impedances," in *TUTAM Symposium on Statistical Energy Analysis*, edited by F. J. Fahy and W. G. Price, Kluwer Academic Publishers.
- Heron, K. H. (1998). "The wave approach to predictive statistical energy analysis and equally spaced point connections with isolators," *Noise-Con 98*, Ypsilanti, Michigan.
- Horner, J. L., and White, R. G. (1991). "Prediction of vibrational power transmission through bends and joints in beam-like structures," *J. Sound Vib.* **147**(1), 87–103.
- Langley, R. S. (1994). "Elastic wave transmission coefficients and coupling loss factors for structural junctions between curved panels," *J. Sound Vib.* **169**(3), 297–317.
- Langley, R. S., and Heron, K. H. (1990). "Elastic wave transmission through plate/beam junctions," *J. Sound Vib.* **143**(2), 241–253.
- Langley, R. S., and Shorter, P. J. (2002). "Diffuse wavefields in cylindrical coordinates," *J. Acoust. Soc. Am.* **112**, 1465–1470.
- Leissa, A. (1993). *Vibration of Plates* (Acoustical Society of America, Woodbury, NY).
- Leung, R., and Pinnington, R. J. (1986). "Point inertance of an infinite plate with respect to a force acting in its plane," *J. Sound Vib.* **111**(1), 125–129.
- Ljunggren, S. (1984). "Generation of waves in an elastic plate by a torsional moment and a horizontal force," *J. Sound Vib.* **93**(2), 161–187.
- Lyons, R. H., and DeJong, R. G. (1995). *Theory and Application of Statistical Energy Analysis* (Butterworth-Heinemann, Boston).
- Lyons, R. H., and Eichler, E. (1964). "Random vibration of connected structures," *J. Acoust. Soc. Am.* **36**, 1344–1354.

- Mace, B. R. (1992). "Reciprocity, conservation of energy and some properties of reflection and transmission coefficients," *J. Sound Vib.* **155**, 375–381.
- Manning, J. E. (1994). "Formulation of SEA parameters using mobility functions," *Philos. Trans. R. Soc. London, Ser. A* **346**, 477–488.
- Moore, J. (1990). "Vibration transmission through frame or beam junctions," *J. Acoust. Soc. Am.* **88**, 2766–2776.
- Morfey, C. L. (1994). "Cylindrical in-plane waves in an elastic plate," *J. Sound Vib.* **173**, 557–560.
- Petersson, B., and Heckl, M. (1996). "Concentrated excitation of structures," *J. Sound Vib.* **196**(3), 295–321.
- Sablik, M. (1982). "Coupling loss factors at a beam L-joint revisited," *J. Acoust. Soc. Am.* **72**, 1285–1288.
- Sarradj, E. (2000). "Calculation of point mobilities at generic multi-plate/beam junctions," *J. Sound Vib.* **229**(2), 411–430.
- Tso, Y., and Norwood, C. (1995). "Vibratory power transmission through three-dimensional beam junctions," *J. Sound Vib.* **185**(4), 595–607.

Investigation into the validity of two-dimensional models for sound radiation from waves in rails

D. J. Thompson,^{a)} C. J. C. Jones, and N. Turner^{b)}

Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton SO17 1BJ, England

(Received 19 July 2002; revised 23 December 2002; accepted 6 January 2003)

The sound radiation from a vibrating rail can be predicted using a two-dimensional model under certain conditions. This paper explores these conditions and shows that, if the decay rate of vibration along the rail becomes large or the wavelength in the rail becomes small, it becomes necessary to allow for three-dimensional radiation characteristics. In practice, however, noise from a rail can be predicted using a two-dimensional model for frequencies above about 250 Hz, and even where three-dimensional effects become important, these can be allowed for by simple correction terms. When the wavelength in the rail approaches that of acoustic waves in air, the angle between the direction of sound radiation from the rail and the normal to the rail increases, in some cases to more than 45°. This must be accounted for if the performance of noise barriers is to be calculated using a two-dimensional approach. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1555612]

PACS numbers: 43.40.Rj, 43.20.Rz, 43.50.Lj [EGW]

I. INTRODUCTION

Railway rolling noise is produced by vibrations of the wheels and the rails, caused by roughness excitation at their point of contact.^{1–3} Theoretical models have been developed for predicting this noise for particular designs of wheel and track.^{1,4} These consist of separate models for the prediction of vibration from the roughness input and for the prediction of radiated sound from the vibration. In the TWINS software⁴ the rolling noise can be calculated in terms of the sound power radiated from the wheel, the rail and the sleeper or, alternatively, as the average sound pressure level resulting at a particular “microphone” position near the track. This paper describes an approach for calculating the radiation from the rail.

The theoretical models are intended to allow various noise reduction technologies to be assessed. One of the noise reduction techniques that has been considered recently is a low barrier placed close to the rail.^{5–8} The distance of the barriers from the rail and their size may be small in comparison with the wavelength of sound in much of the frequency range of interest. It is therefore necessary to use methods such as the boundary element method to evaluate their performance so that the effects of diffraction can be taken into account reliably.

Since the rail and barrier effectively have a two-dimensional geometry, it is advantageous to use a two-dimensional boundary element model, or equivalent, to study rail radiation and barrier performance. For example, a two-dimensional equivalent source model of the rail radiation has been developed⁹ and is implemented in the TWINS model.⁴ The alternative, a full three-dimensional model, would become very large especially at high frequencies, as the rail

vibrates over a large length. Another approach would be to decompose the results according to the acoustic wave number in the direction of the track axis and to construct three-dimensional results using a Fourier transform.¹⁰ This would also involve considerable additional computation compared with a purely two-dimensional approach. This paper therefore examines the extent to which two-dimensional models for the sound radiation from a rail can be used, and identifies where a three-dimensional model is necessary.

It is clear that a two-dimensional model should be appropriate provided that (i) the structural wavelength in the rail is much longer than the wavelength of sound in air at the same frequency, and (ii) the decay of vibration with distance along the rail is small. If condition (ii) is not satisfied, the rail vibration will be localized at the forcing point, that is the wheel/rail contact position, and will tend to form a point source rather than a line source. These two requirements will be studied in more detail to provide quantitative conditions for the validity of a two-dimensional approach.

In Sec. II the form of rail vibration is reviewed. Section III discusses typical results from a two-dimensional boundary element calculation. An equivalent simple source model is introduced in Sec. IV and used in Sec. V to study the two conditions mentioned above.

II. RAIL VIBRATION

A. Track vibration models

Rail vibration consists of the superposition of various types of propagating structural waves.^{11,12} At low frequencies, the track support stiffness inhibits the propagation of these waves and hence their decay rate along the rail is high. Above the corresponding “bounce” resonance of the rail on its supports, waves in the rail “cut on,” that is, their decay rate falls and is governed only by damping effects. In order to account for the various types of wave in the rail, a model is considered here in which the cross-section is represented

^{a)}Author to whom correspondence should be addressed; electronic mail: djt@isvr.soton.ac.uk

^{b)}Present address: BOC Edwards, Dolphin Road, Shoreham-by-Sea, West Sussex BN43 6PY, England.

TABLE I. Parameters used to represent the two tracks considered.

	Track with soft pads		Track with stiff pads	
	Vertical	Lateral	Vertical	Lateral
Rail	UIC60		UIC54	
Rail loss factor	5×10^{-4}		5×10^{-4}	
Pad stiffness, N/m	8×10^7	1.6×10^7	1.3×10^9	1×10^8
Pad loss factor	0.16	0.16	0.25	0.25
Sleeper mass, kg	140		122	
Sleeper spacing, m	0.6		0.6	
Ballast stiffness, N/m	8×10^7	4×10^7	1.5×10^8	7.5×10^7
Ballast loss factor	1.0	1.0	0.5	0.5

by finite element matrices.¹² The support is assumed to be continuous, comprising damped springs to represent the rail pads, a mass to represent the sleepers, and another layer of springs to represent the ballast. The effect of the discrete nature of the support is neglected here. Its main effect is around 1000 Hz where the wavelength of vertical bending waves equals twice the sleeper spacing. For lateral vibration the effect is much smaller.¹¹

The rail response due to the wheel/rail interaction forces can be evaluated in terms of these propagating/decaying waves. Expressing their complex propagation constants as $s_i = \beta_i + jk_i$, where k_i represents the propagating part of the wave number and β_i represents the decay with distance, the response u in wave i at position z along the rail is given by

$$u_i(z) = u_i(0)e^{-s_i|z|} = u_i(0)e^{-(\beta_i + jk_i)|z|}. \quad (1)$$

The decay rate of a wave i in dB/m is given by $20 \log_{10}(e) \text{Re}\{s_i\} = 8.686 \beta_i$. The response consists of a sum of many waves, although only a small number of them, generally those with the lowest decay rates, make a significant contribution to the sound radiation.

The total sound power from the rail is predicted as the sum of the powers from each wave. This relies on the decay rates being relatively low and the wavelengths of each wave being different. Under such circumstances the relative phase of the respective wave motions can be ignored. For waves that have a high decay rate, which therefore form a localized point source, the relative phase of their motion is important. However, these waves usually have a negligible contribution to the total power.

B. Wave numbers of typical tracks

The most important parameter of a track that affects the wave propagation along the rail is the dynamic stiffness of the pad that is usually placed beneath the rail.¹³ This can vary in a large range. The model of Ref. 12 has been used to calculate the propagation constants (complex wave numbers) in the rail for two tracks with pad stiffnesses at opposite ends of the range of stiffness values found in practice, namely, a soft 10 mm thick studded pad and a stiff 5 mm thick pad. The parameters used to represent these two tracks, including the pads, sleepers and ballast, are listed in Table I.

Figures 1 and 2 show the propagating wave number, $k_i = \text{Im}\{s_i\}$ for these two tracks. The wave number is expressed in terms of the ratio k_i/k_{air} , where $k_{\text{air}} = \omega/c_0$ is the wave number of sound in air at the corresponding frequency. It can

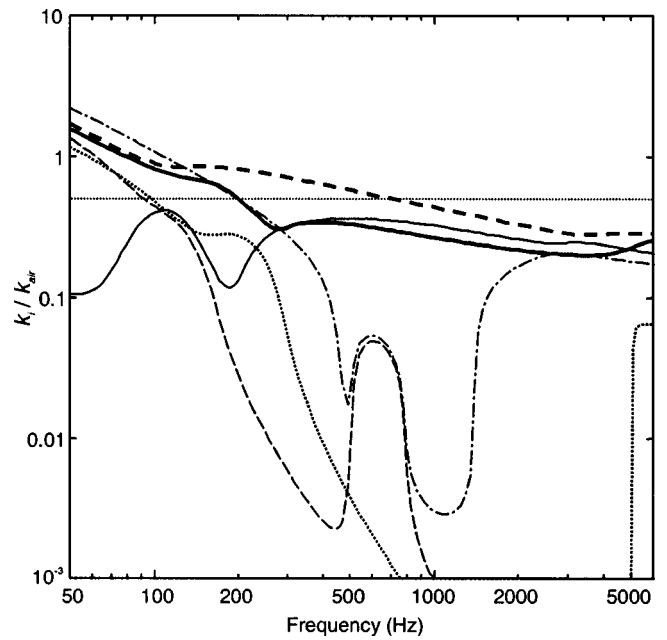


FIG. 1. Propagating part of normalized wave number calculated for track with soft pads. — vertical bending wave, vertical near-field wave, --- lateral bending wave, — · — torsional wave, — · — web bending wave, — — — lateral near-field wave.

be seen that above 1000 Hz, in both cases, all the wave types shown have a ratio smaller than 0.5. This means that condition (i) stated in the Introduction is satisfied here. The various waves are identified according to their low wave number behavior.¹² At least below 4 kHz, these are the only waves that contribute significantly to the noise radiated by a conventional rail. Apart from longitudinal waves, which are not excited, other “wave” solutions are evanescent, having high decay rates.

Figures 3 and 4 show the equivalent results for the decaying part of the wave number. This is, again, expressed in

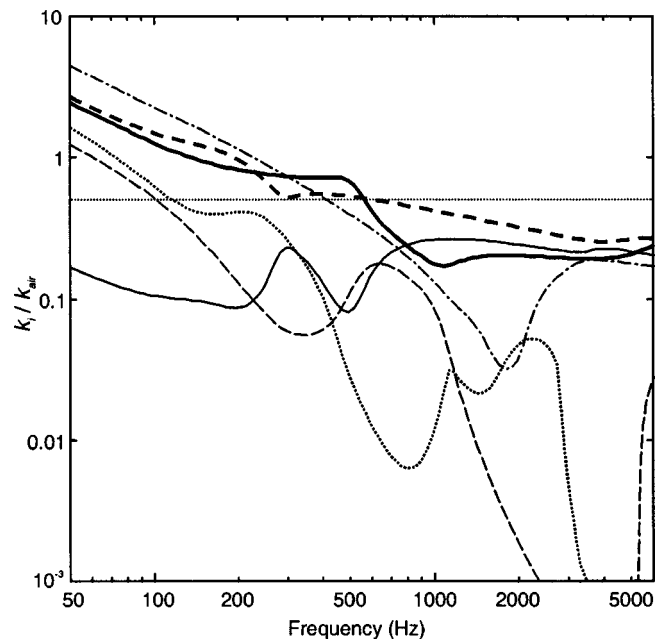


FIG. 2. Propagating part of normalized wave number calculated for track with stiff pads. Key as Fig. 1.

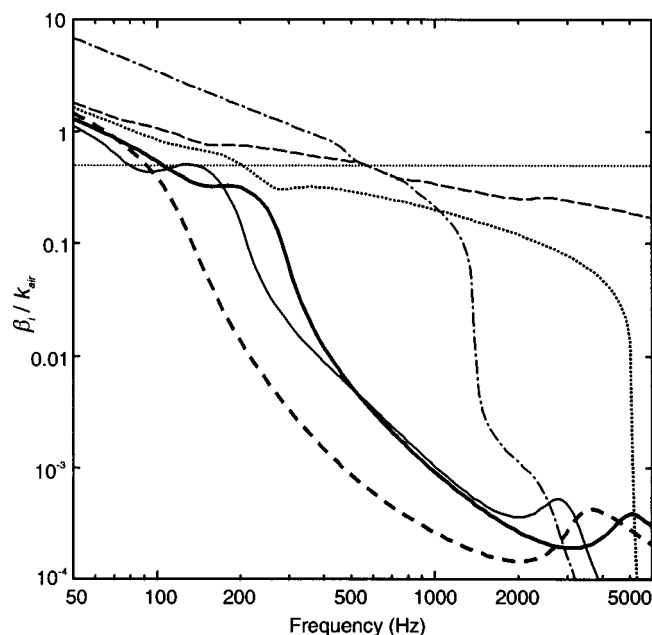


FIG. 3. Decaying part of normalized wave number calculated for track with soft rail pads. Key as Fig. 1.

terms of the ratio to the acoustic wave number, β_i / k_{air} . In this case also, it can be seen that above 1000 Hz for both tracks all wave types have a ratio smaller than 0.5. This means that condition (ii) stated in the Introduction is satisfied here. The waves that produce the largest part of the noise radiated by a rail are generally those with the lowest decay rates, i.e., the lowest values of β_i , since this implies the longest vibrating length of rail. For these waves $\beta_i / k_{air} < 1$ even at frequencies as low as 100 Hz.

From this examination of the wave numbers of two tracks at opposite extremes of the pad stiffness range, it may be concluded that two-dimensional models for the noise radiation from the rail should always be appropriate for fre-

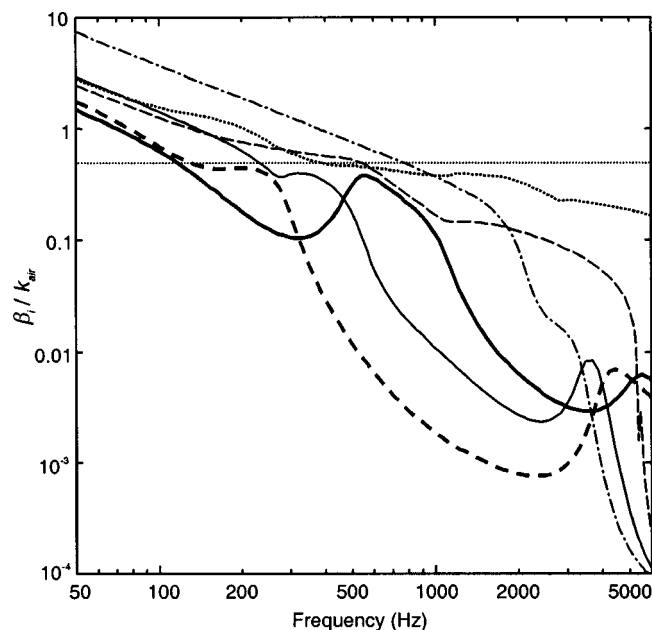


FIG. 4. Decaying part of normalized wave number calculated for track with stiff rail pads. Key as Fig. 1.

quencies above 1000 Hz. Any requirement to use three-dimensional models is likely to be restricted to frequencies below 1000 Hz. The remainder of this paper therefore concentrates on frequencies below 1000 Hz. In this frequency region, the web bending wave and other higher order waves do not propagate. Deformation of the rail cross-section can be neglected and the only waves of significance are vertical bending, lateral bending, and torsion.

III. TWO-DIMENSIONAL BOUNDARY ELEMENT MODELING

For a structure which extends infinitely in one direction (z) with a constant cross-section and which has the same velocity distribution on every cross-section, the sound power per unit length, W' , can be written as

$$W' = \rho_0 c_0 \sigma_{2D} D \langle \bar{v}^2 \rangle, \quad (2)$$

where ρ_0 is the density of air, c_0 is the speed of sound in air, D is the perimeter length of the cross-section, and $\langle \bar{v}^2 \rangle$ is the averaged mean-square normal velocity on the perimeter. The factor σ_{2D} is the radiation efficiency, here relating to the two-dimensional case.

An alternative formulation of Eq. (2) can be used if the velocity field represents rigid motion of the cross-section, e.g., vertically or laterally. The product $D \langle \bar{v}^2 \rangle$ can be replaced by $2h \bar{v}^2$, where \bar{v}^2 is the mean square velocity amplitude of the whole section in the appropriate direction and h is a projection of the perimeter into a plane normal to the direction of vibration (e.g., the height for a lateral motion).

A two-dimensional model of a UIC 60 rail cross-section has been constructed using a proprietary boundary element package. The model makes use of the vertical line of symmetry of the rail so that half the cross-section is represented using 32 line elements. The CHIEF method of specifying extra nodes on the interior of the rail section was used to overcome the nonuniqueness of the solution due to modes of the equivalent internal problem. The model represents a freely suspended rail, whereas in practice the ballast and sleepers are in the immediate vicinity of the rail. However, neglect of these objects does not affect the test of the validity of a two-dimensional model for a three-dimensional situation, which is the main subject of this paper.

The calculated radiation efficiencies, σ_{2D} , for rigid-body vertical and lateral motions of the rail cross-section are presented in Fig. 5 in decibel form (i.e., radiation index). The rail cross-section is also shown for reference. It can be seen that up to approximately 800 Hz both radiation efficiencies follow a straight line of the form $30 \log_{10}(f/f_0)$, where f is the frequency and f_0 is the frequency at which the line would cross 0 dB. The frequencies f_0 are approximately 870 Hz for vertical motion and 570 Hz for lateral motion. This shape of curve corresponds to the radiation efficiency of a simple line dipole source (oscillating infinite cylinder¹⁴). This is to be expected for frequencies at which the dimensions of the rail are much smaller than the wavelength of sound in air. At higher frequencies, the radiation efficiency tends towards unity (0 dB).

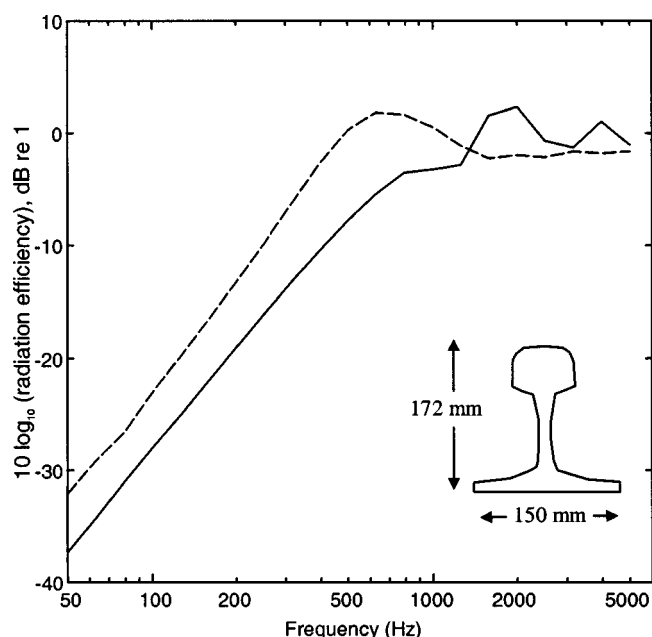


FIG. 5. Radiation efficiencies of UIC 60 rail section modeled in two dimensions. — vertical motion, — — horizontal motion.

Above 1000 Hz there are peaks and troughs in the radiation efficiency curves corresponding to constructive and destructive interference caused by phase differences in the sound from different parts of the rail section. These are represented only approximately in the figure due to the use of one-third octave frequency spacing. These peaks and troughs occur due to the fact that the wavelength of sound in air is similar to the characteristic dimensions of the cross-section for these frequencies. The most significant of these features is the dip at 1000 Hz in the vertical radiation efficiency. At this frequency the wavelength in air is approximately equal to twice the distance between the radiating faces of the top of the rail head and the rail foot. Radiation from these two surfaces in a vertical direction at this frequency thus interferes destructively. At 2000 Hz it would be expected to interfere constructively and a peak is observed.

In the frequency range below 1000 Hz no effects are observed which are due to the cross-sectional shape of the rail, except that the dimensions of the rail section determine the effective source strength for each motion.

IV. A MODEL BASED ON SIMPLE SOURCES

A. Method for studying three-dimensional effects

By inspecting the wave numbers, it has been shown that three-dimensional radiation effects from practical rails are expected to be limited to frequencies well below 1000 Hz. By using two-dimensional boundary element models it has also been shown that the details of the rail cross-section are unimportant for sound radiation in this frequency region. It has also been noted that rail cross-sectional deformation can be neglected in this frequency region. To represent three-dimensional effects, therefore, only variations in the direction along the axis of the rail are important.

The oscillating rail cross-section is replaced by an array of point dipole sources of suitably chosen source strength

arranged along the centerline of the rail. In order for this model to be equivalent to the rail section it is necessary for the sources to be placed close together compared with both the wavelength in air and the wavelength in the structure. The source strengths are chosen allowing for the spacing of the sources, the size of the rail section, and the velocity amplitude and phase on the surface of the rail. In fact, in the results presented here, since the power is normalized to that from an equivalent line source, only the relative magnitude and phase of the source strengths is of importance. These are chosen according to Eq. (1) to represent the radiation from a single structural wave. The length of the source region that is modelled depends on the real part of the propagation constant. It is taken as $6.5/\beta_i$, which ensures that the amplitude has reduced by 56 dB by the end of the source region.

B. Numerical study

Using this method, the sound radiation is studied from a single decaying travelling wave in the rail, emanating in two directions symmetrically from a forcing point. The results are obtained for a range of values of k_i and β_i , that is the imaginary and real parts of the complex propagation constant s_i which are varied independently. Although frequency is another variable, by expressing the results in nondimensional form with both k_i and β_i normalized to k_{air} , results are obtained which are in fact independent of frequency. This has also been verified by performing some calculations at different frequencies. The range of values taken for $\xi = k_i/k_{\text{air}}$ and $\gamma = \beta_i/k_{\text{air}}$ is chosen to represent values occurring in track in the frequency range of interest. However, small values of β_i bring a penalty in terms of computation time as the length of the source region becomes very large. Results are presented in terms of three parameters representing the overall power, the average direction of radiation and the ratio of average squared pressure, at a "trackside" location, to the power.

V. RESULTS

A. Sound power

The sound power radiated by a vibrating body can be expressed as

$$W = \rho_0 c_0 \sigma S \langle \overline{v^2} \rangle, \quad (3)$$

where σ is the radiation efficiency, S is the surface area and $\langle \overline{v^2} \rangle$ is the surface averaged mean square normal velocity. In the case of a rail, the surface area S is effectively infinite and has to be chosen to be long enough to capture the dominant part of the decaying wave. In the current study, for reasons given in Sec. II, the velocity is assumed constant on any cross-section, i.e., it is a function only of the length coordinate z . Thus Eq. (3) can be written as

$$W = \rho_0 c_0 \sigma 2h \int_{-L/2}^{L/2} v_0^2(z) dz, \quad (4)$$

where h is the equivalent height of the section, and L is the length over which the average is taken.

The equivalent two-dimensional result is given in Eq. (2), which can be written as

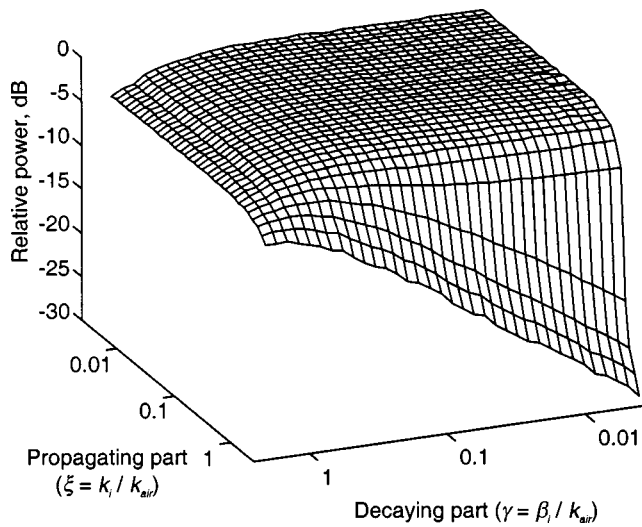


FIG. 6. Variation of sound power W/W_{2D} with rail nondimensional wave number.

$$W'_{2D} = W_{2D}/L = \rho_0 c_0 \sigma_{2D} 2h \bar{v}_0^2, \quad (5)$$

where σ_{2D} is the radiation efficiency obtained from a two-dimensional calculation (e.g., Fig. 5), and \bar{v}_0^2 is the mean square velocity assumed to act at all positions z . As an approximation of a three-dimensional case, Eq. (5) can be used with \bar{v}_0^2 given by the integral in Eq. (4), i.e. representing the average of the mean square velocity over the length L . From Eqs. (4) and (5), the ratio of the sound power radiated in the true three-dimensional situation and the two-dimensional approximation is thus:

$$W/W_{2D} = \sigma/\sigma_{2D}. \quad (6)$$

This ratio has been calculated using the simple dipole source model described above for a range of values of the decaying and propagating parts of the wave number. The results are plotted in Fig. 6, which shows the sound power level in the full three-dimensional situation relative to the equivalent two-dimensional case. It can be seen that, over much of the range of complex wave number, the sound power changes little from the two-dimensional approximation. Figure 7 shows the same results in the form of a contour plot. Three characteristic regions can be identified in Figs. 6 and 7:

- (1) The flat “plateau” region in the lower left-hand corner of Fig. 7 (toward the rear of Fig. 6) corresponds to the region where a two-dimensional model is intended to work. Here the decay rate is low, so the rail is effectively infinitely long, and the propagating wave number is also small (long wavelengths). Here $W/W_{2D} = 1$.
- (2) For high decay rates (toward the right of Fig. 7 and the left of Fig. 6) the sound power begins to drop compared to the plateau region. In this region the rail ceases to be a line source and tends to a point source. The radiation efficiency of a point dipole has a frequency dependence of $40 \log_{10}(f/f_0)$ whereas it is $30 \log_{10}(f/f_0)$ for a line dipole; see Fig. 5. This provides an explanation of the reduction in power in this region.

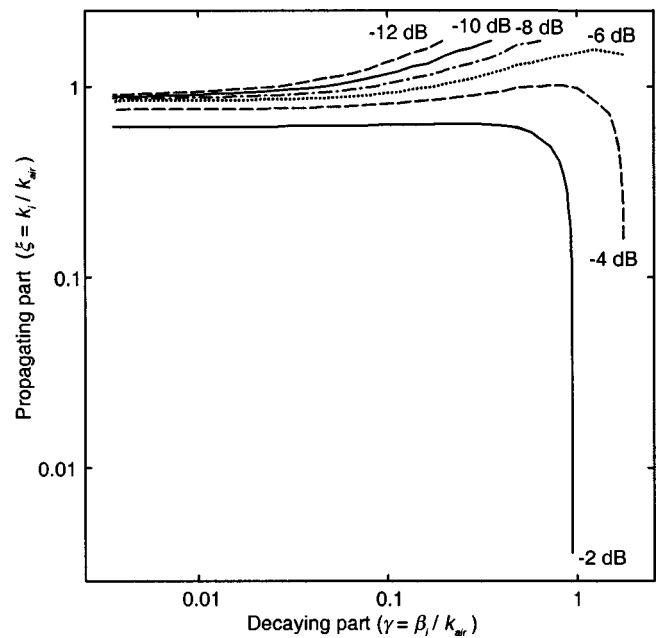


FIG. 7. Contour plot showing variation of sound power W/W_{2D} with decaying and propagating parts of nondimensional wavenumber.

- (3) For low decay rates and large propagating wave numbers or short wavelengths, i.e., towards the top left of Fig. 7 and the lower right of Fig. 6, there is a sharp drop in level (a “precipice”). Where the wavelength in the rail is shorter than that in air and the decay rate is low, acoustic short-circuiting occurs and the radiation is much reduced compared to that for long wavelengths.¹⁵ Some residual radiation occurs due to the discontinuity introduced at the forcing point. Increasing the decay rate leads to a reduction in the effective length over which short-circuiting can occur and an increase in the influence of the discontinuity and hence an increase in the average radiation (i.e., σ).

Figure 7 shows that sound power will be predicted to within 2 dB of the correct value using a two-dimensional model, provided that the propagating wave number, $k_i < 0.65 k_{air}$ and the decaying part of the wave number, $\beta_i < k_{air}$. From Figs. 1–4 it has already been seen that these conditions are easily satisfied for frequencies above 1000 Hz for practical railway tracks. It may be noted that these conclusions are based on the nature of the wave propagation in the z direction and are therefore unaffected by the discrepancies introduced by using simple sources in place of the rail cross-section.

B. Application to track wavenumbers

In order to gain additional insight into the relevance of the phenomena presented above to actual tracks, the real and imaginary parts of the propagation constants of the track are plotted as loci on the complex plane for comparison with Fig. 7. Figure 8 shows these curves for the track with soft rail pads, taken from the data of Figs. 1 and 3. As frequency increases, the curves move from the top right, down and towards the left. The propagation constants of propagating

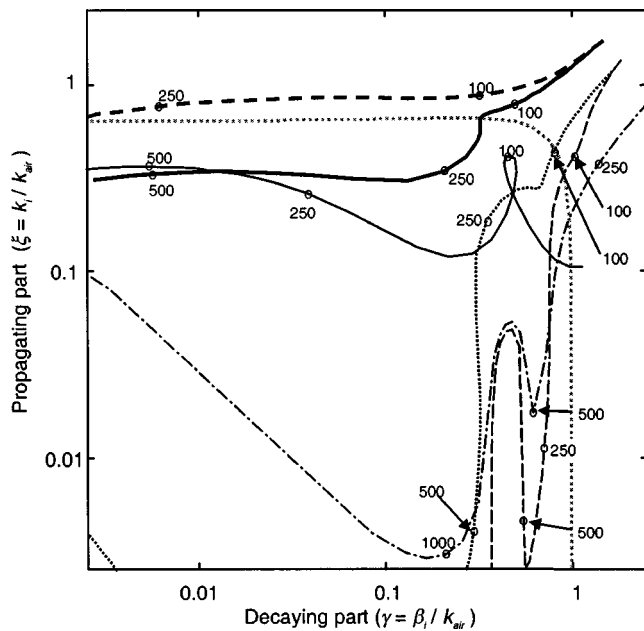


FIG. 8. Propagating versus decaying part of normalized wave number calculated for track with soft pads. The gray dotted line is the -2 dB contour from Fig. 7. Numbers indicate frequency in Hz. — vertical bending wave, vertical near-field wave, --- lateral bending wave, — torsional wave, -.- web bending wave, - - - lateral near-field wave.

waves have a greater imaginary part than real part and therefore tend towards the upper left part of the plot. Near-field waves, on the other hand, tend towards the lower right part of the plot. The web bending wave can be seen to “cut on,” i.e., move from near-field to propagating wave behavior above 1000 Hz. Equivalent results for the track with stiff pads are shown in Fig. 9, from which similar trends can be seen.

The -2 dB contour from Fig. 7 is also shown in these figures. Clearly, most of the waves are within the -2 dB

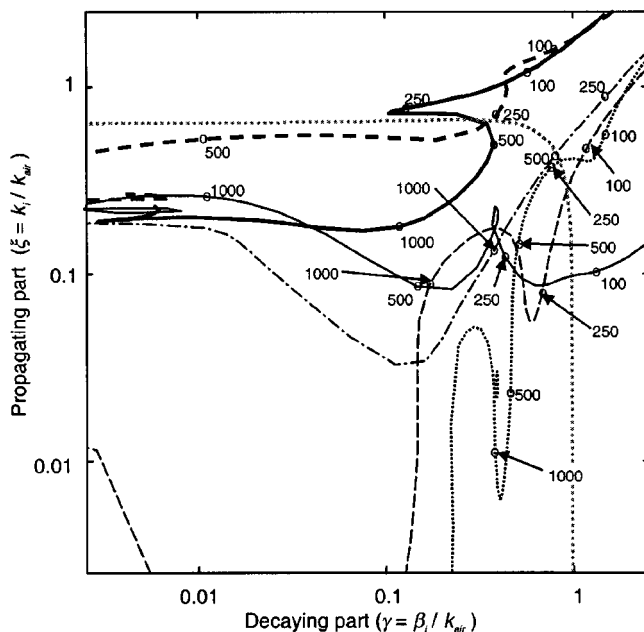


FIG. 9. Propagating versus decaying part of nondimensional wave number calculated for track with stiff pads. Key as for Fig. 8.

contour for most of the frequency range. Below 100 Hz most of the waves have wave numbers with large real and imaginary parts and the curves tend toward the top right of the figure. Also, the lateral bending wave of the track with soft pads is outside the -2 dB contour for frequencies up to about 500 Hz. Apart from this, all other waves of both tracks are within the -2 dB contour above about 250 Hz.

In terms of sound power, therefore, all the waves of importance in a rail, and even the near-field waves, can be considered to radiate sound in a way that is consistent with a two-dimensional model for frequencies above about 250 Hz. The exception is the lateral bending wave for the track with soft pads. Because of the soft pad, this has a particularly low cut-on frequency and also a low decay rate. This wave has a wavelength close to that in air for frequencies between 100 and 250 Hz (see also Fig. 1) which means that the sound radiation is attenuated due to acoustic short-circuiting effects by up to 4 dB in this region.

C. Directivity of sound radiation

A pure line source radiates a two-dimensional sound field, so the component of intensity in the z direction is always zero. A point source, on the other hand, radiates omnidirectionally, so the intensity vector is always orientated radially.

For the present application, the required measure of sound level is the average as a train passes a point at the track-side. This can be seen, equivalently, as the average as the receiver location moves parallel to the rail past the excitation point. The Doppler effect is ignored.

The “directivity” is studied here in terms of the angle of inclination of the intensity vector to the normal to the rail. The angle of the intensity vector at a distance y_0 from the rail is averaged along a line parallel to the rail. This average is weighted according to the magnitude of the intensity by integrating the intensity in two orthogonal directions, I_y in the y direction, normal to the track, and I_z in the z direction parallel to the track:

$$\tan \bar{\theta} = \frac{\int_0^\infty I_z(0, y_0, z) dz}{\int_0^\infty I_y(0, y_0, z) dz}, \text{ for some } y_0 > 0. \quad (7)$$

The integrals are only taken over the range $z \geq 0$. Inclusion of the range $z < 0$ would always lead to $\bar{\theta} = 0$, because of the symmetry of the wave propagation in the rail.

At very low decay rates these results will be independent of distance, y_0 as they represent the angle of the incidence of the radiated sound field to the rail, $\bar{\theta} = \sin^{-1}(k_i/k_{\text{air}})$. Also for a point source, Eq. (7) leads to constant values of $\bar{\theta}$ independent of y_0 . For a point monopole $\bar{\theta} = 45^\circ$, and for a point dipole $\bar{\theta} = \tan^{-1}(0.5) = 26.5^\circ$. These results are demonstrated more fully in Appendix A. For intermediate sources, the distance y_0 may play a role in determining $\bar{\theta}$. The distance used in the present calculations is 3 m.

Figure 10 shows the results as a series of vectors indicating direction. The base of each arrow represents the value of complex wave number at which they are calculated. A

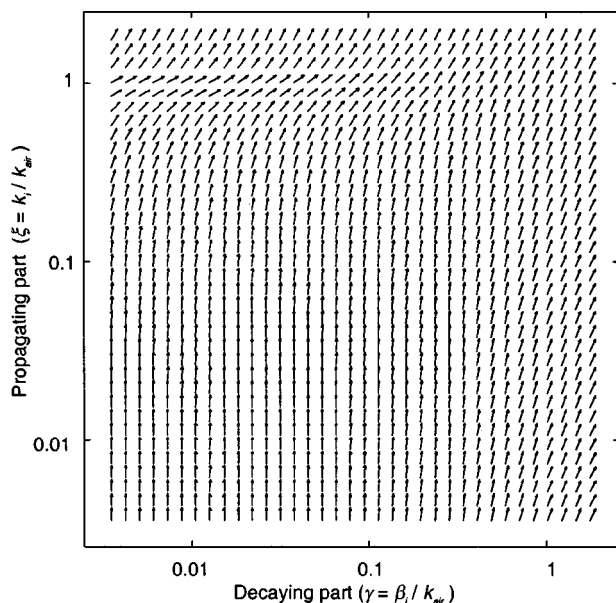


FIG. 10. Average direction of sound intensity vector versus real and imaginary parts of nondimensional wave number.

vertical arrow indicates $\bar{\theta}=0$ (intensity normal to the rail) and a horizontal arrow indicates $\bar{\theta}=90^\circ$ (intensity parallel to the rail). In the bottom left part of the plot, corresponding to the “plateau” region of Figs. 6 and 7, the average direction is close to the normal. In the region along both the upper and right edges of the region shown, the vector is close to the direction associated with a point dipole source (26.5°). Only in the region close to the “precipice” of Figs. 6 and 7 ($k_i \approx k_{\text{air}}$) is the angle greater than this.

These results can be interpreted as follows:

- (1) In the “plateau” region, at the bottom left of Fig. 10, the wavelength in the rail is much longer than that in air and the vibrating length of rail is also long. The sound radiation is therefore two-dimensional and the intensity vector is approximately normal to the rail.
- (2) At the right-hand edge of Fig. 10 the length of vibrating rail is very short and the rail acts as a point dipole source.
- (3) In the region where $k_i \approx k_{\text{air}}$ and the decay rate is low, the rail radiates sound with a significant wave number component in the z direction, $k_z = k_i$ and a smaller component in the y direction that satisfies $k_z^2 + k_y^2 = k_{\text{air}}^2$. The average angle tends to $\sin^{-1}(k_i/k_{\text{air}})$, as shown in Fig. 11. As the decay rate increases, this is partly modified by the radiation from the discontinuity close to the driving point.
- (4) In the region where $k_i > k_{\text{air}}$ and the decay rate is low (the top of Fig. 10) the far field radiation from the free wave propagation at $z \gg 0$ is greatly attenuated by acoustic short-circuiting. Only the discontinuity in the wave field near the driving point radiates significant net power. The average directivity is therefore similar to that of a point dipole.

The plot in Fig. 10 can be compared with the loci of wave numbers of actual tracks (Figs. 8 and 9). It is found

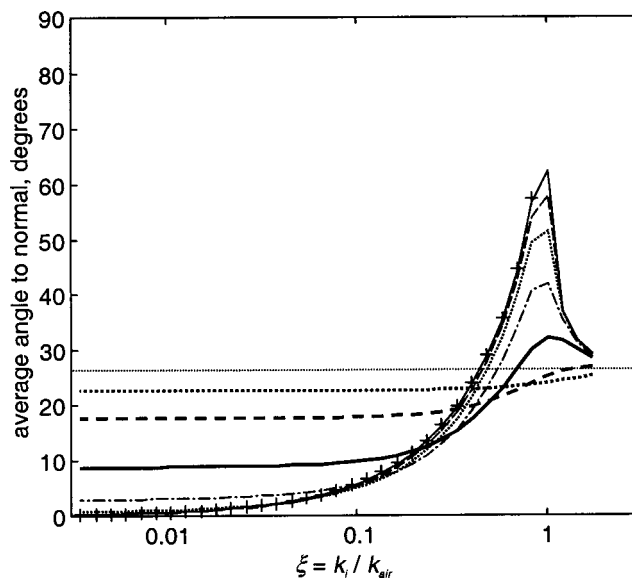


FIG. 11. Average direction of sound intensity vector versus imaginary parts of nondimensional wave number. +: $\sin^{-1}(k_i/k_{\text{air}})$. Curves for different values of $\gamma = \beta_i/k_{\text{air}}$: — 0.0035, - - - 0.022, 0.055, - · - · 0.14, — — — 0.34, - - - 0.84, 1.74. Grid line at 26.5° : result for point dipole.

that, apart from low frequencies where it has already been observed that a point source is more appropriate than a line source, only the lateral bending wave for the track with the soft pad has a direction of radiation significantly different from the normal. The angle in this case has a maximum of about 45° around 150–200 Hz. At higher frequencies the angle reduces as k_i/k_{air} reduces.

The angle associated with each wave has been determined by interpolating the field in Fig. 10 to the wave numbers of the tracks shown in Figs. 8 and 9. The results are shown as graphs of angle against frequency in Figs. 12 and 13 for frequencies where the wave number is within the range considered in Fig. 10. There are some gaps in these curves where the wave numbers go out of the range covered by Fig. 10.

From these graphs it can be seen that the various curves tend toward the point dipole value of 26.5° at low frequencies and generally tend to values close to 0° at high frequencies. For the track with soft pads, the lateral bending wave has a high angle of incidence, with a maximum value of 54° , around 200 Hz. This decreases toward higher frequency but unfortunately the decay rate becomes too low to allow comparison with the data of Fig. 10. For the track with stiff pads, both the vertical and lateral bending waves have average angles of incidence that rise above the point source value; in the case of the vertical wave the maximum angle is 38° around 250 Hz. At higher frequencies the angles reduce as the waves enter the ‘plateau’ region of Fig. 6.

D. Sound pressure estimates

It is shown in Appendix B that the average sound pressure along a line a distance away from a line of incoherent point dipole sources is 0.7 dB higher than at the same distance from a coherent dipole line source of the same overall power. For monopoles the corresponding difference is 2 dB.

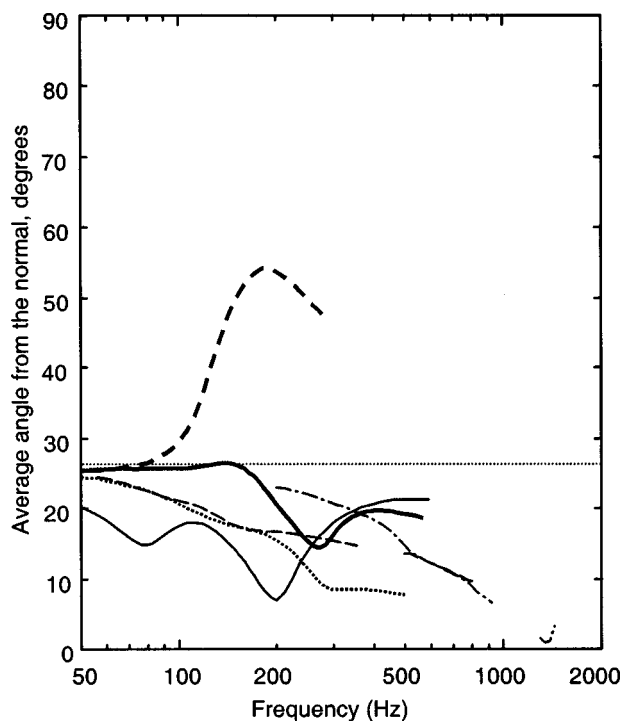


FIG. 12. Average direction of sound intensity vector for track with soft pads. — vertical bending wave, vertical near-field wave, --- lateral bending wave, — torsional wave, - - - web bending wave, - - - lateral near-field wave. Grid line at 26.5°: result for point dipole.

Since the time-averaged pressure due to a moving point source can also be represented as the pressure at a single point due to a line of incoherent point sources, it can be expected that, in the “point source regions” of high real or imaginary parts of wave number, the average pressure will be higher than in the “line source region” for a given power.

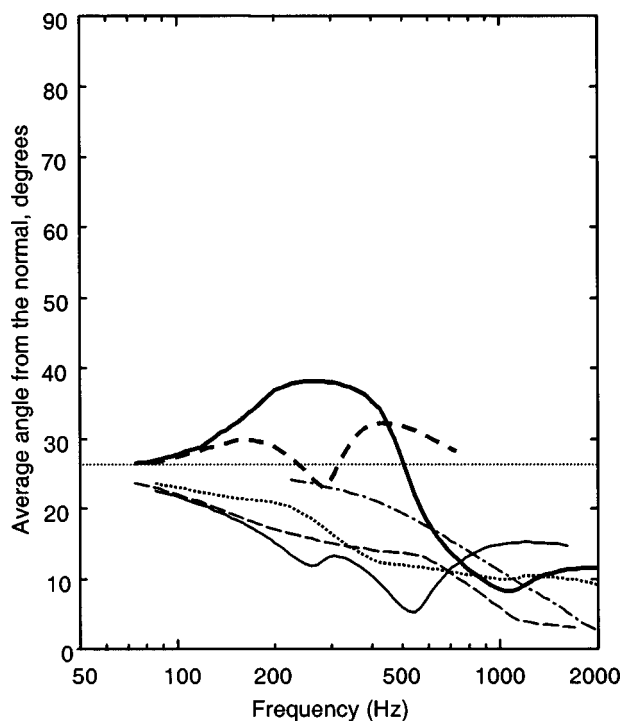


FIG. 13. Average direction of sound intensity vector for track with stiff pads. Key as Fig. 12.

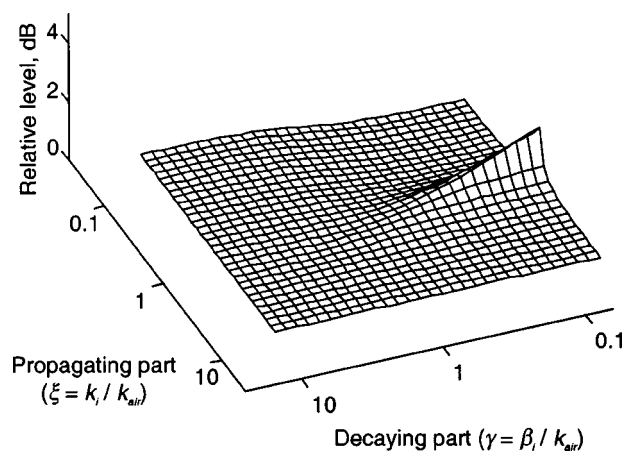


FIG. 14. Sound pressure level relative to that for a coherent line source of the same power.

Figure 14 shows the average sound pressure level at 3 m relative to that for a coherent line source of the same power. Note that the range of values of complex wave number differs from the earlier figures. In the region with low values of ξ and γ this tends to 0 dB, whereas for high values of ξ and γ it tends to 0.7 dB as expected. In the coincidence region, close to $\xi=1$ there is a higher level, rising to +2.5 dB at $\gamma=0.1$. This corresponds to the region of the power results where the level is part way down the “precipice,” and will partly compensate the reduction in power observed in this region, for example for the lateral wave of the track with the soft rail pad.

VI. CONCLUDING REMARKS

The extent to which three-dimensional effects need to be considered in the radiation from a rail has been studied using a line of dipole sources with their source strengths adjusted to represent wave motion in the rail. At long wavelengths and low decay rates a two-dimensional approach is shown to be valid in predicting sound power, sound pressure and the direction of sound propagation. Provided that the propagating part of the rail wave number, $k_i < 0.65k_{\text{air}}$ and the decaying part of the rail wave number, $\beta_i < k_{\text{air}}$ a two-dimensional approach yields results within 2 dB of the correct value.

For typical tracks with both soft and stiff pads it is shown that three-dimensional effects need only be considered for frequencies below about 250 Hz. It is thus seen that the two-dimensional model gives satisfactory results for most practical cases. Where this is not the case, notably at low frequencies, it is possible to use the two-dimensional model with some corrections. These should allow for the reduction in radiation efficiency from a $30 \log_{10}(f)$ dependence to a $40 \log_{10}(f)$ dependence when the decay rate is high. Allowance for acoustic short-circuiting is also required when the wavelength is short. Simple formulas can be fitted to the results presented here and used in combination with a two-dimensional model.

To predict the effect of low barriers close to the rail it is also necessary to allow for the direction of radiation, which can be at a significant angle to the normal for radiation from the vertical or lateral bending waves in the rail. Below 250

Hz, a three-dimensional boundary element calculation or equivalent is required. However, in practice, the low frequency performance of barriers is not important for the overall sound reduction and a two-dimensional calculation is likely to be sufficient.

ACKNOWLEDGMENTS

This work was undertaken as part of the Silent Track project funded by the European Union (Brite Euram Project No. BE 96-3017, Contract No. BRPR-CT 96-0258) and coordinated by the European Rail Research Institute (ERRI).

APPENDIX A: AVERAGE ANGLE OF SOUND RADIATION TO THE NORMAL FOR A MOVING POINT SOURCE

Monopole

Consider a point monopole located at the point $x=y=z=0$ emitting sound power W . A line of receivers is located on the line $y=y_0$ parallel to the z axis. This is equivalent to a point monopole which moves past a fixed microphone (apart from the Doppler effect which is ignored). The sound intensity at $(0, y_0, z)$ is given by:

$$I = \frac{W}{4\pi(y_0^2 + z^2)}. \quad (A1)$$

This has a component in the y direction of $I_y = I \cos \theta$ and a component in the z direction of $I_z = I \sin \theta$, where $\tan \theta = z/y_0$. Evaluating the integrals of Eq. (6),

$$\bar{I}_z = \int_0^\infty I_z(y_0, z) dz = \frac{W}{4\pi} \int_0^\infty \frac{\sin \theta}{(y_0^2 + z^2)} dz, \quad (A2)$$

$$\bar{I}_y = \int_0^\infty I_y(y_0, z) dz = \frac{W}{4\pi} \int_0^\infty \frac{\cos \theta}{(y_0^2 + z^2)} dz, \quad (A3)$$

and substituting $\tan \theta = z/y_0$, so $dz = y_0 \sec^2 \theta d\theta$ (and $\sec^2 \theta = 1 + z^2/y_0^2$)

$$\bar{I}_z = \frac{W}{4\pi y_0}; \quad \bar{I}_y = \frac{W}{4\pi y_0}. \quad (A4)$$

The “average” angle is thus given by $\tan \bar{\theta} = \bar{I}_z/\bar{I}_y = 1$, i.e., $\bar{\theta} = 45^\circ$, which is independent of y_0 .

Dipole

Similarly for a point dipole located at the point $x=y=z=0$ orientated with its maximum intensity in the direction of the y axis, the sound intensity at $(0, y_0, z)$ is given by:

$$I = \frac{I_0 y_0^2 \cos^2 \theta}{(y_0^2 + z^2)}, \quad (A5)$$

where I_0 is the intensity at the position $y=y_0, z=0$. Evaluating the integrals of Eq. (6),

$$\bar{I}_z = \int_0^\infty I_z(y_0, z) dz = I_0 \int_0^\infty \frac{\cos^2 \theta \sin \theta}{(1 + (z/y_0)^2)} dz, \quad (A6)$$

$$\bar{I}_y = \int_0^\infty I_y(y_0, z) dz = I_0 \int_0^\infty \frac{\cos^3 \theta}{(1 + (z/y_0)^2)} dz, \quad (A7)$$

and substituting $\tan \theta = z/y_0$, so $dz = y_0 \sec^2 \theta d\theta$

$$\bar{I}_z = \frac{I_0 y_0}{3}; \quad \bar{I}_y = \frac{2I_0 y_0}{3}. \quad (A8)$$

The “average” angle is thus given by $\tan \bar{\theta} = \bar{I}_z/\bar{I}_y = 0.5$, i.e., $\bar{\theta} = 26.5^\circ$, which again is independent of y_0 .

APPENDIX B: AVERAGE SOUND PRESSURE FROM COHERENT LINE SOURCE AND MOVING POINT SOURCE

Line monopole

Consider a line monopole source (i.e., a pulsating cylinder) on the z axis emitting sound power W' per unit length. The sound intensity vector is always normal to the z axis and has amplitude $I(r)$ dependent only on the distance r ,

$$I(r) = \frac{W'}{2\pi r}. \quad (B1)$$

This represents a cylindrical wave front. In the far field it is close to a plane wave and the mean-square pressure can be approximated by

$$\bar{p}^2 = \frac{\rho_0 c_0 W'}{2\pi r}. \quad (B2)$$

Line dipole

A line dipole can be represented by an oscillating cylindrical source, the solution for which is an intensity distribution with a $\cos^2 \varphi$ angular dependence in the x - y plane. Again the intensity vector lies in the x - y plane. Integrating over a cylindrical surface,

$$W' = \int_0^{2\pi} I_0(r) \cos^2 \varphi r d\varphi = \pi r I_0(r), \quad (B3)$$

where $I_0(r)$ is the intensity at radius r on the axis of the dipole. Hence, the mean-square pressure in the far field on the axis of the dipole is

$$\bar{p}^2 = \frac{\rho_0 c_0 W'}{\pi r}. \quad (B4)$$

Point monopole

Consider a point monopole at $x=y=z=0$, with a line of receivers located on the line $y=y_0$. From Eq. (A1), in the far field, the average mean-square pressure on the line $(-L/2, L/2)$ is

$$\begin{aligned} \langle \bar{p}^2 \rangle &= \frac{\rho_0 c_0}{L} \int_{-L/2}^{L/2} |\bar{I}(y_0, z)| dz \\ &= \frac{\rho_0 c_0 W}{4\pi L} 2 \int_0^{L/2} \frac{1}{(y_0^2 + z^2)} dz. \end{aligned} \quad (B5)$$

Substituting $\tan \theta = z/y_0$,

$$\langle \bar{p}^2 \rangle = \frac{\rho_0 c_0 W}{2\pi L y_0} \tan^{-1}(L/2y_0). \quad (B6)$$

This result also applies to the pressure at a single location due to a line of incoherent monopoles. If their power per unit length is $W' = W/L$, it is possible to let $L \rightarrow \infty$ giving

$$\langle \overline{p^2} \rangle = \frac{\rho_0 c_0 W'}{4y_0}. \quad (\text{B7})$$

Comparing this with Eq. (B2) it can be seen that an array of incoherent point monopoles produces an average sound pressure level which is $10 \log_{10}(\pi/2) = 1.96$ dB higher than a coherent line source of the same power.

Point dipole

Similarly consider a point dipole located at the point $x = y = z = 0$ orientated with its maximum intensity in the direction of the y axis, with a line of receivers located on the line $y = y_0$. From Eq. (A5), in the far field, the average mean-square pressure on the line $(-L/2, L/2)$ is

$$\begin{aligned} \langle \overline{p^2} \rangle &= \frac{\rho_0 c_0}{L} \int_{-L/2}^{L/2} |\tilde{I}(y_0, z)| dz \\ &= \frac{\rho_0 c_0 I_0 y_0^2}{L} 2 \int_0^{L/2} \frac{\cos^2 \theta}{(y_0^2 + z^2)} dz, \end{aligned} \quad (\text{B8})$$

where I_0 is the intensity at $(0, y_0, 0)$. Substituting $\tan \theta = z/y_0$,

$$\langle \overline{p^2} \rangle = \frac{\rho_0 c_0 I_0 y_0}{L} (\chi + \sin 2\chi), \quad (\text{B9})$$

where $\chi = \tan^{-1}(L/2y_0)$. From an integral over a spherical surface, the total power W from a point dipole is related to the maximum intensity, I_0 , by $W = 4\pi I_0(r)r^2/3$. Substituting this into Eq. (B9) gives,

$$\langle \overline{p^2} \rangle = \frac{3\rho_0 c_0 W}{4\pi L y_0} (\chi + \sin 2\chi). \quad (\text{B10})$$

This result also holds for the pressure at a single point due to a line of incoherent dipoles. If their power per unit length is $W' = W/L$, it is possible to let $L \rightarrow \infty$, $\chi \rightarrow \pi/2$, giving the average pressure on the line $(0, y_0, z)$ as

$$\langle \overline{p^2} \rangle = \frac{3\rho_0 c_0 W'}{8y_0}. \quad (\text{B11})$$

Comparing this with Eq. (B4) it can be seen that an array of incoherent point dipoles produces an average sound pressure level which is $10 \log_{10}(3\pi/8) = 0.71$ dB higher than a coherent line dipole source of the same power.

- ¹P. J. Remington, "Wheel/rail rolling noise, I: Theoretical analysis," *J. Acoust. Soc. Am.* **81**, 1805–1823 (1987).
- ²P. J. Remington, "Wheel/rail rolling noise: What do we know? What don't we know? Where do we go from here?," *J. Sound Vib.* **120**, 203–226 (1988).
- ³D. J. Thompson and C. J. C. Jones, "A review of the modeling of wheel/rail noise generation," *J. Sound Vib.* **231**, 519–536 (2000).
- ⁴D. J. Thompson, B. Hemsworth, and N. Vincent, "Experimental validation of the TWINS prediction program for rolling noise, Part 1: Description of the model and method," *J. Sound Vib.* **193**, 123–135 (1996).
- ⁵P. J. Remington, M. J. Rudd, and I. L. V  r, "Wheel/rail noise and vibration. Vol. 2 Applications to control wheel/rail noise," US Department of Transportation Report UMTA-06-0025-75-11 (1975).
- ⁶R. R. K. Jones, "Bogie shrouds and low barriers could significantly reduce wheel/rail noise," *Railway Gazette International* **150**, 459–462 (1994).
- ⁷H. A. Holties, "Railway noise reduction using porous surfaces and small barriers," *Proc. Euronoise 98*, M  nchen, Germany, pp. 301–305 (1998).
- ⁸C. J. C. Jones, D. J. Thompson, and T. P. Waters, "Application of numerical models to a system of train- and track-mounted acoustic shields," *Int. J. Acoust. Vib.* **6**, 185–192 (2001).
- ⁹M.-F. Petit, M. Heckl, J. Bergemann, and J. Baee, "Berechnung des Abstrahlma  es von Eisenbahnschienen anhand der Multipolsynthese" (Calculation of the radiation efficiency of railway rails using multipole synthesis), *Proc. DAGA 1992*, pp. 997–1000 (1992).
- ¹⁰E. M. Salomons, A. C. Geerlings, and D. Duhamel, "Comparison of a ray model and a Fourier-boundary element method for traffic noise situations with multiple diffractions and reflections," *Acust. Acta Acust.* **83**, 35–47 (1997).
- ¹¹D. J. Thompson and N. Vincent, "Track dynamic behaviour at high frequencies. Part 1: Theoretical models and laboratory measurements," *Veh. Syst. Dyn. Suppl.* **24**, 86–99 (1995).
- ¹²D. J. Thompson, "Wheel-rail noise generation, Part III: Rail vibration," *J. Sound Vib.* **161**, 421–446 (1993).
- ¹³N. Vincent, P. Bouvet, D. J. Thompson, and P. E. Gautier, "Theoretical optimization of track components to reduce rolling noise," *J. Sound Vib.* **193**, 161–171 (1996).
- ¹⁴F. J. Fahy, *Sound and Structural Vibration* (Academic, London, 1985), pp. 103–104.
- ¹⁵G. Maidanik, "Response of ribbed panels to reverberant acoustic fields," *J. Acoust. Soc. Am.* **34**, 809–826 (1962).

Helmholtz resonator with extended neck^{a)}

Ahmet Selamet^{b)} and Iljae Lee

Department of Mechanical Engineering and The Center for Automotive Research,
The Ohio State University, Columbus, Ohio 43210-1107

(Received 23 August 2002; revised 3 January 2003; accepted 13 January 2003)

Acoustic performance of a concentric circular Helmholtz resonator with an extended neck is investigated theoretically, numerically, and experimentally. The effect of length and shape of, and the perforations on the neck extension is examined on the resonance frequency and the transmission loss. A two-dimensional analytical method is developed for an extended neck with constant cross-sectional area, while a three-dimensional boundary element method is applied for the variable area and perforated extension. Lumped and one-dimensional approaches are also included to illustrate the effect of the higher order modes. For a piston-driven model, predicted resonance frequencies using lumped, one-dimensional, and two-dimensional analytical methods are compared with those from multidimensional boundary element method. Analytical and computational transmission loss predictions for pipe-mounted model are compared to the experimental data obtained from an impedance tube setup. It is shown that the resonance frequency may be controlled by the length, shape, and perforation porosity of the extended neck without changing the cavity volume. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1558379]

PACS numbers: 43.50.Gf [ANN]

LIST OF SYMBOLS

a_1	Neck radius
a_2	Cavity radius
A_n	Modal amplitudes in domain I
B_n	Modal amplitudes in domain II
C_n	Modal amplitudes in domain III
c_0	Speed of sound
f_r	Resonance frequency
G	Green's function
J_m	Bessel function of the first kind and order m
k	Wave number
k_0	Wave number of mode (0,0)
k_n	Wave number of mode (0, n)
ℓ_1	Base neck length
ℓ_2	Neck extension length
ℓ_n	Total neck length
ℓ_c	Cavity length
ℓ_3	$\ell_c - \ell_2$
\mathbf{n}	Unit normal vector in the outward direction
P	Acoustic pressure
q_1, q_2	Points in domains
r	Coordinate
S	Cross-sectional area of acoustic domains
S_n	Neck cross-sectional area
S_p	Pipe cross-sectional area
T	Impedance matrix
T_{ij}	Transfer matrix elements
TL	Transmission loss
u	Acoustic velocity

u_p	Piston velocity
V_c	Cavity volume
x_1, x_2, x_3	Coordinates
x, y, z	Cartesian coordinates
Y_m	Bessel function of the second kind and order m
Z_H	Acoustic impedance of a Helmholtz resonator
α_n	Roots of $J_1(\alpha_n) = 0$
β_n	Roots of Eq. (9)
Γ	Acoustic domain boundary
δ_n	End correction
ρ_0	Density of air
σ	Duct porosity
ψ_n	Eigenfunctions
ω	Angular velocity

Superscript

+	Traveling in the positive direction
−	Traveling in the negative direction
i	Inlet
o	Outlet
r	Rigid wall

Subscript

A, I	Domain I
B, II	Domain II
C, III	Domain III
c	Cavity
M	Main duct

I. INTRODUCTION

Helmholtz resonator is an effective acoustic attenuation device at low frequencies with its resonance dictated by the

combination of cavity and neck and their relative orientation. The classical lumped analysis of this attenuator gives the resonance frequency as $f_r = (c_0/2\pi) \sqrt{S_n / \{V_c(\ell_n + \delta_n)\}}$, where c_0 is speed of sound, S_n the neck cross-sectional area, V_c the resonator volume, ℓ_n the neck length, and δ_n the end correction to account for higher modes excited at the discontinuities, which can be determined by the geometry and

^{a)}A preliminary version of this manuscript has been presented at the 142nd meeting of the Acoustical Society of America at Fort Lauderdale, FL, 3–7 December 2001.

^{b)}Electronic mail: Selamet.1@osu.edu

location of the neck relative to the volume and main duct.^{1,2} Ingard¹ investigated the effect of neck geometry, such as cross-sectional area shape, location, and size, on the resonance frequency of a Helmholtz resonator with circular or rectangular cross-sectional area for volume. He developed end corrections for both single and double holes to account for the higher order modes at the interface between neck and cavity. Chanaud² examined the effect of both orifice and cavity geometry on the resonance frequency of a Helmholtz resonator using Ingard's end correction. The effect of the depth and width for variable and fixed volumes was studied. He presented the limitations of simple lumped and transcendental models based on the predictions, and concluded that for a fixed volume and orifice size, the orifice position changed the resonance frequency substantially, while the orifice shape was not significant. References 1 and 2 considered only very short neck length compared to wavelength. Tang and Sirignano³ studied a Helmholtz resonator with a neck comparable to wavelength as an application to reducing combustion instability. They developed a general formulation and applied it to quarter wave resonators and Helmholtz resonators with various neck lengths.

Recently, Selamet and co-workers⁴⁻⁶ employed several approaches to examine the effect of cavity volume and neck locations. They illustrated the effect of length-to-diameter ratio of the volume on the resonance frequency and transmission loss characteristics using lumped and one-dimensional (1D) approaches.⁴ Earlier works have been extended further by studying a number of circular concentric configurations with lumped, one-dimensional radial and axial, two-dimensional (2D) analytical approaches, and three-dimensional (3D) boundary element method (BEM).⁵ They also developed a 3D analytical approach to investigate the effect of neck offset on the behavior of circular asymmetric Helmholtz resonator.⁶

While the Helmholtz resonator is known to be an effective acoustic reflector at low frequencies, the use of it may, at times, be limited due to volume constraints. Thus, it is important to lower the resonance frequency without increasing volume or reduce volume without increasing resonance frequency. While a wealth of literature exists on the effect of neck cross-sectional area and opening location to volume, the impact of neck extension remains to be investigated. The present work therefore concentrates on the effect of neck extension geometry (Fig. 1) on the Helmholtz resonator behavior. The neck extension length will shift the resonance frequency without changing the volume. Such an extension also acts like a quarter wave resonator in the cavity resulting in additional transmission loss peaks at higher frequencies compared to a resonator without extension. Furthermore, the neck shape involving either variable cross-sectional area or perforations can readily change the acoustic characteristics of the resonator. Such an ability to modify the resonance frequency without changing the volume may be desirable in a variety of applications, including adaptive passive-noise control.⁷⁻⁹ Among other designs, Cheng and co-workers,⁸ for example, suggested two concentric cylindrical neck extensions with preferably parabolic cut-outs. The rotation of the

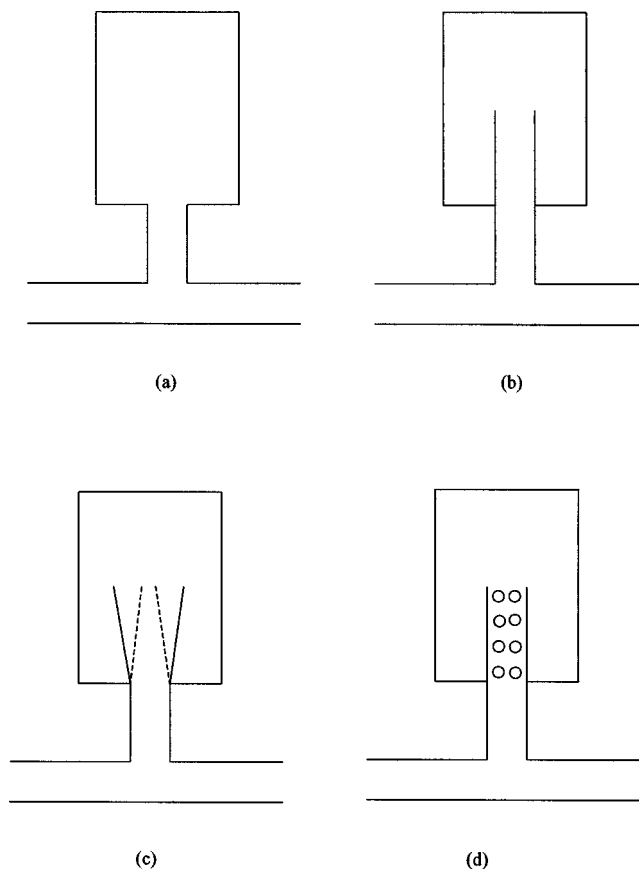


FIG. 1. Helmholtz resonators without (a) and with extended necks (b)–(d).

cylindrical walls varies the alignment of cut-outs, therefore effectively changing the neck length.

The objective of the present study is, in the absence of mean flow, to (1) investigate theoretically, numerically, and experimentally the acoustic attenuation performance of a concentric circular Helmholtz resonator with an *extended neck*; and (2) examine the effects of *length*, *shape*, and *perforation* of the neck extension on the resonator behavior. A two-dimensional analytical method is developed for an extended neck with constant cross-sectional area, while a three-dimensional boundary element method is applied for the variable-area and perforated extension. Lumped and one-dimensional approaches are also presented to illustrate the effect of the higher order modes. For a piston-driven model, predicted resonance frequencies using lumped, 1D, and 2D analytical methods are compared with those from 3D BEM. Analytical and computational transmission loss predictions for pipe-mounted model are compared to the experimental data obtained from an impedance tube setup.

Following this Introduction, Sec. II develops a two-dimensional analytical approach and Sec. III summarizes a 3D BEM. Section IV compares the analytical and computational predictions with experiments, and discusses the effect of geometry of the neck extension on the resonance frequency and transmission loss characteristics. Section V concludes the study with final remarks.

II. TWO-DIMENSIONAL ANALYTICAL APPROACH

A two-dimensional analytical approach is introduced next to determine the acoustic characteristics of a piston-

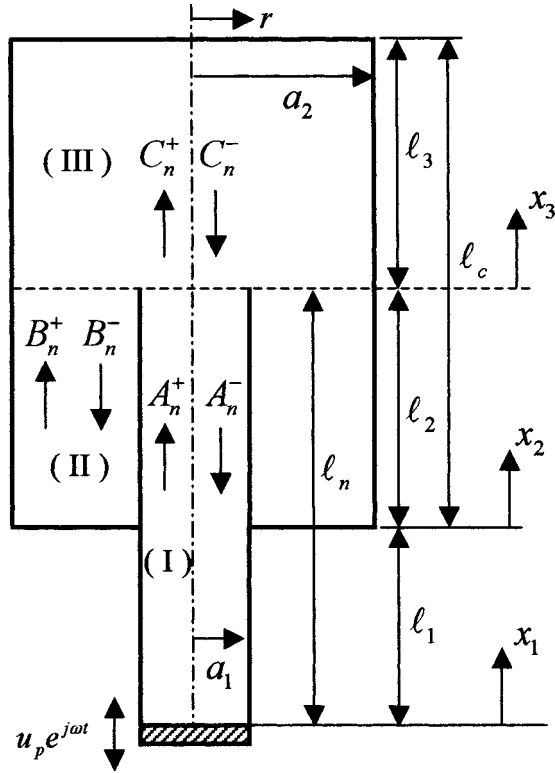


FIG. 2. Wave propagation in a piston-driven Helmholtz resonator with a neck extended into the cavity volume.

driven circular Helmholtz resonator with an extended neck, which consists of circular and annular ducts (Fig. 2). For a two-dimensional axisymmetric wave propagation in a circular or annular duct, the Helmholtz equation is given, in cylindrical coordinates (r, x) , by

$$\nabla^2 P(r, x) + k^2 P(r, x) = 0, \quad (1)$$

where P and k are the acoustic pressure and the wave number, respectively. The solution to Eq. (1) is obtained next in three domains: I—neck, II—annular, and III—circular in the volume.

A. Solution for circular neck (domain I)

The solution to Eq. (1) in domain I or circular neck (Fig. 2) can be written as¹⁰

$$P_A(r, x_1) = \sum_{n=0}^{\infty} (A_n^+ e^{-jk_{A,n}x_1} + A_n^- e^{jk_{A,n}x_1}) \psi_{A,n}(r), \quad (2)$$

where P_A is the acoustic pressure, A_n^+ and A_n^- the modal amplitudes corresponding to components traveling in the positive and negative x_1 directions in domain I, respectively, $k_{A,n}$ the wave number, and $\psi_{A,n}(r)$ the eigenfunctions. For this circular duct, the eigenfunctions are given by

$$\psi_{A,n}(r) = J_0\left(\alpha_n \frac{r}{a_1}\right), \quad (3)$$

where J_m is the Bessel function of the first kind and order m , a_1 the duct diameter, and α_n the roots satisfying the rigid wall boundary condition of

$$J'_0(\alpha_n) = J_1(\alpha_n) = 0, \quad (4)$$

and

$$k_{A,n} = \begin{cases} \sqrt{k_0^2 - \left(\frac{\alpha_n}{a_1}\right)^2}, & k_0 > \frac{\alpha_n}{a_1} \\ -\sqrt{k_0^2 - \left(\frac{\alpha_n}{a_1}\right)^2}, & k_0 < \frac{\alpha_n}{a_1} \end{cases} \quad (5)$$

the axial wave number of the mode $(0, n)$, and k_0 wave number of the mode $(0, 0)$. The negative sign in Eq. (5) is assigned so that $e^{-jk_{A,n}x_1}$ decays exponentially in x_1 .

The particle velocity may then be written, in terms of the linearized momentum equation, as

$$u_A(r, x_1) = \frac{1}{\rho_0 \omega} \sum_{n=0}^{\infty} k_{A,n} [A_n^+ e^{-jk_{A,n}x_1} - A_n^- e^{jk_{A,n}x_1}] \psi_{A,n}(r), \quad (6)$$

where ρ_0 is the medium density and ω the angular velocity.

B. Solution for annular volume (domain II)

For a concentric annular duct with inner and outer duct diameters of a_1 and a_2 , the solution to Eq. (1) can be written in domain II as

$$P_B(r, x_2) = \sum_{n=0}^{\infty} (B_n^+ e^{-jk_{B,n}x_2} + B_n^- e^{jk_{B,n}x_2}) \psi_{B,n}(r) \quad (7)$$

with

$$\psi_{B,n}(r) = J_0\left(\beta_n \frac{r}{a_2}\right) - \frac{J_1(\beta_n)}{Y_1(\beta_n)} Y_0\left(\beta_n \frac{r}{a_2}\right), \quad (8)$$

where Y_m is the Bessel function of the second kind and order m , β_n the root satisfying the boundary condition of

$$J_1\left(\beta_n \frac{a_1}{a_2}\right) - \frac{J_1(\beta_n)}{Y_1(\beta_n)} Y_1\left(\beta_n \frac{a_1}{a_2}\right) = 0 \quad (9)$$

and

$$k_{B,n} = \begin{cases} \sqrt{k_0^2 - \left(\frac{\beta_n}{a_2}\right)^2}, & k_0 > \frac{\beta_n}{a_2} \\ -\sqrt{k_0^2 - \left(\frac{\beta_n}{a_2}\right)^2}, & k_0 < \frac{\beta_n}{a_2} \end{cases} \quad (10)$$

Again the negative sign in Eq. (10) is assigned so that $e^{-jk_{B,n}x_2}$ decays exponentially in x_2 . The particle velocity is then

$$u_B(r, x_2) = \frac{1}{\rho_0 \omega} \sum_{n=0}^{\infty} k_{B,n} [B_n^+ e^{-jk_{B,n}x_2} - B_n^- e^{jk_{B,n}x_2}] \psi_{B,n}(r). \quad (11)$$

C. Solution for circular volume (domain III)

The solution to Eq. (1) in domain III or circular volume (Fig. 2) can be written as

$$P_C(r, x_3) = \sum_{n=0}^{\infty} (C_n^+ e^{-jk_{C,n}x_3} + C_n^- e^{jk_{C,n}x_3}) \psi_{C,n}(r) \quad (12)$$

with

$$\psi_{C,n}(r) = J_0\left(\alpha_n \frac{r}{a_2}\right), \quad (13)$$

where α_n are the roots satisfying the rigid wall boundary condition of Eq. (4), and

$$k_{C,n} = \begin{cases} \sqrt{k_0^2 - \left(\frac{\alpha_n}{a_2}\right)^2}, & k_0 > \frac{\alpha_n}{a_2} \\ -\sqrt{k_0^2 - \left(\frac{\alpha_n}{a_2}\right)^2}, & k_0 < \frac{\alpha_n}{a_2}. \end{cases} \quad (14)$$

The particle velocity is then

$$u_C(r, x_3) = \frac{1}{\rho_0 \omega} \sum_{n=0}^{\infty} k_{C,n} [C_n^+ e^{-jk_{C,n}x_3} - C_n^- e^{jk_{C,n}x_3}] \psi_{C,n}(r). \quad (15)$$

D. Boundary conditions

The coefficients in Eqs. (2), (7), and (12) are determined next by the boundary conditions. At $x_2=0$ and $x_3=\ell_3$, the rigid boundary conditions, $u_B=0$ and $u_C=0$, give, in view of Eqs. (11) and (15),

$$B_n^+ = B_n^- \quad (16)$$

and

$$C_n^- = C_n^+ e^{-2jk_{C,n}\ell_3}. \quad (17)$$

At $x_1=\ell_n$ or $x_3=0$, the pressure continuity condition between domains I and III, $P_A = P_C$, results in (Miles¹¹ and Selamet and Ji¹²)

$$(A_s^+ e^{-jk_{A,s}\ell_n} + A_s^- e^{jk_{A,s}\ell_n}) \langle \psi_{A,s}, \psi_{A,s} \rangle_A = \sum_{n=0}^{\infty} (C_n^+ + C_n^-) \langle \psi_{C,n}, \psi_{A,s} \rangle_A, \quad (18)$$

where $\langle \rangle_s$ indicates the integration over area S, and $s = 0, 1, 2, \dots, \infty$, which is deferred to Appendix A. Similarly, at $x_2=\ell_2$ or $x_3=0$, the pressure continuity between domains II and III, $P_B = P_C$, leads to

$$(B_s^+ e^{-jk_{B,s}\ell_2} + B_s^- e^{jk_{B,s}\ell_2}) \langle \psi_{B,s}, \psi_{B,s} \rangle_B = \sum_{n=0}^{\infty} (C_n^+ + C_n^-) \langle \psi_{C,n}, \psi_{B,s} \rangle_B. \quad (19)$$

At $x_1=\ell_n$ or $x_2=\ell_2$, the volume velocity continuity, $u_A S_A + u_B S_B = u_C S_C$, gives

$$\begin{aligned} & \sum_{n=0}^{\infty} k_{A,n} (A_n^+ e^{-jk_{A,n}\ell_n} - A_n^- e^{jk_{A,n}\ell_n}) \langle \psi_{A,n}, \psi_{C,s} \rangle_A \\ & + \sum_{n=0}^{\infty} k_{B,n} (B_n^+ e^{-jk_{B,n}\ell_2} - B_n^- e^{jk_{B,n}\ell_2}) \langle \psi_{B,n}, \psi_{C,s} \rangle_B \\ & = k_{C,s} (C_s^+ - C_s^-) \langle \psi_{C,s}, \psi_{C,s} \rangle_C. \end{aligned} \quad (20)$$

E. Resonance frequency and transmission loss

The expressions for the resonance frequency and transmission loss are given next for the piston-driven and pipe-

mounted Helmholtz resonators, respectively. Assuming decay of the higher order modes through neck from $x_1=\ell_n$ to $x_1=0$ (Selamet *et al.*⁵), the piston oscillating with velocity magnitude u_p at $x_1=0$ gives

$$\frac{1}{\rho_0 c_0} (A_0^+ - A_0^-) = u_p. \quad (21)$$

Once Eqs. (16)–(21) are solved using $s=n$ and letting $\rho_0 c_0 u_p = 1$, the acoustic impedance of a Helmholtz resonator at $x_1=0$ is calculated by⁶

$$Z_H = \frac{P_A}{\rho_0 c_0 u_p} = A_0^+ + A_0^-, \quad (22)$$

which then allows the evaluation of resonance frequency.

For a pipe-mounted Helmholtz resonator, transmission loss can be determined, with the assumptions of anechoic termination at the exit of the main duct and plane wave propagation through the main duct, by⁶

$$TL = 10 \log_{10} \left| 1 + \frac{S_n}{2S_p} \left(\frac{1}{Z_H} \right) \right|, \quad (23)$$

where S_n and S_p are the area of the neck and main pipe, respectively.

III. THREE-DIMENSIONAL BOUNDARY ELEMENT METHOD

The wave propagation is governed by the Helmholtz equation in Cartesian coordinates,

$$\nabla^2 P(x, y, z) + k^2 P(x, y, z) = 0, \quad (24)$$

and the Green's theorem results in boundary integral equation as follows:

$$\begin{aligned} C(q_1)P(q_1) &= \int_{\Gamma} \left[G(q_1, q_2) \frac{\partial P(q_2)}{\partial \mathbf{n}} \right. \\ &\quad \left. - P(q_2) \frac{\partial G(q_1, q_2)}{\partial \mathbf{n}} \right] d\Gamma(q_2), \end{aligned} \quad (25)$$

where q_1 and q_2 are points on the boundary surface Γ , $C(q_1)$ coefficient, and $G(q_1, q_2)$ the Green's function or fundamental solution given, for a three-dimensional acoustic domain, by

$$G(q_1, q_2) = \frac{e^{-jk|q_1 - q_2|}}{4\pi|q_1 - q_2|}. \quad (26)$$

A. Piston-driven model

Discretizing the boundary surfaces into a number of elements and applying numerical integration to Eq. (25) yields¹³ for domain I

$$\begin{Bmatrix} P_I(x_1=0) \\ P_I(x_1=\ell_n) \\ P_I^r \end{Bmatrix} = [T_1] \begin{Bmatrix} u_I(x_1=0) \\ u_I(x_1=\ell_n) \\ u_I^r \end{Bmatrix}, \quad (27)$$

and for the domain II+III,

$$\begin{Bmatrix} P_{II+III}(x_3=0) \\ P_{II+III}^r \end{Bmatrix} = [T_2] \begin{Bmatrix} u_{II+III}(x_3=0) \\ u_{II+III}^r \end{Bmatrix}, \quad (28)$$

TABLE I. Helmholtz resonators with various neck lengths [cm] and shapes.

Configuration	Extended neck shape	Base neck length (ℓ_1)	Neck extension length (ℓ_2)	Total neck length ($\ell_n = \ell_1 + \ell_2$)
1	Straight solid	8.5	0–18	8.5–26.5
2	Straight solid	18.5–0.5	0–18	18.5
3	Conical straight solid	8.5	10	18.5
4	Straight perforated	8.5	10	18.5

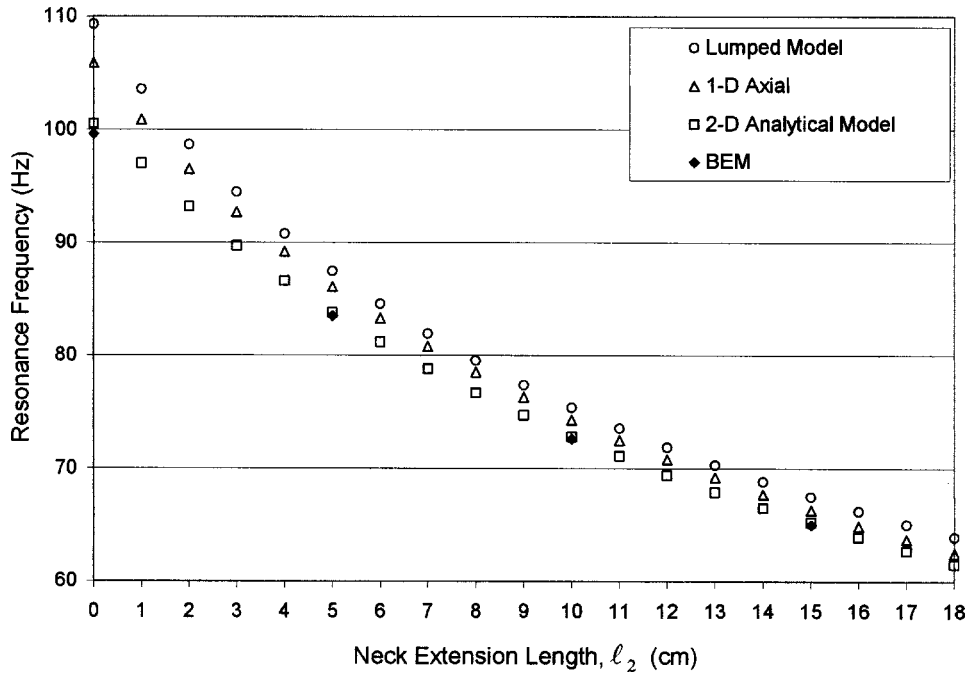


FIG. 5. Resonance frequency predictions for Helmholtz resonators with various neck extension lengths ($\ell_1 = 8.5$ cm).

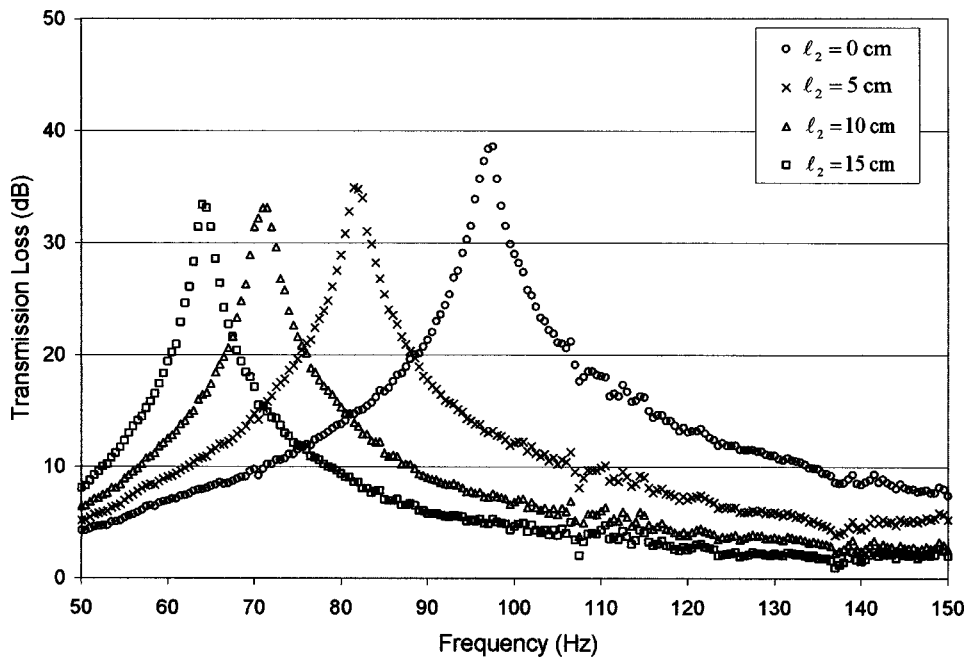


FIG. 6. Measured transmission loss for Helmholtz resonators with neck extension ($\ell_1 = 8.5$ cm).

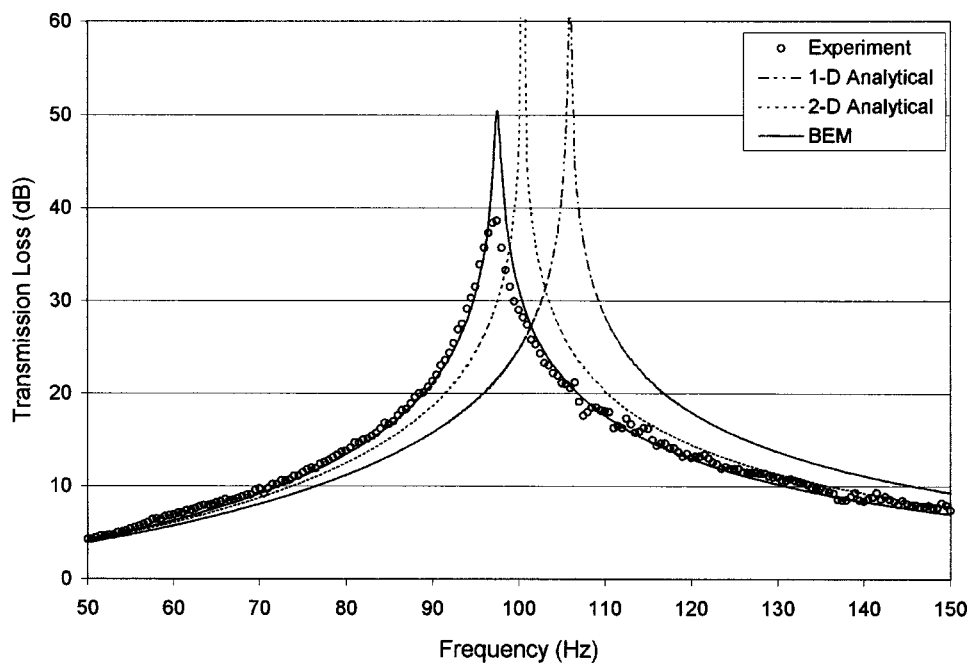


FIG. 7. Comparison of predictions with experiment for a Helmholtz resonator without neck extension ($\ell_1 = 8.5$ cm and $\ell_2 = 0$ cm).

length. The square main duct is then connected to the circular impedance tube with smooth transitions that retain a constant cross-sectional area development.

For a piston-driven Helmholtz resonator of configuration 1, Fig. 5 compares a number of predictions for resonance frequencies. In general, the resonance frequency decreases as the extension length ℓ_2 increases from 0 to 18 cm. 2D analytical and BEM predictions exhibit a good agreement, as expected. The difference between the resonance frequencies using 2D analytical approach with five and ten higher order modes is within 0.3%, thus only the predictions with five modes are presented. Setting $s=0$ in Eqs. (16)–(20) gives 1D axial model predictions, leading to overestimation of resonance frequencies. Lumped model predictions without end correction demonstrate higher resonance frequencies

than the other models. In the lumped model, the volume of a neck extension is subtracted from the total resonator volume. The relatively large difference for short extensions among the analytical methods may be attributed to increasing relative significance of higher order modes (due to decreasing total neck length) generated at the interface between neck and volume.

An impedance tube test setup⁵ is used to obtain transmission loss of pipe-mounted resonators applying the two-microphone technique. Figure 6 shows the measured transmission loss for different neck extension lengths. As the extension length increases, the resonance frequency decreases and the attenuation band becomes narrower. Note that a 15 cm change in the extension length results in a 33 Hz

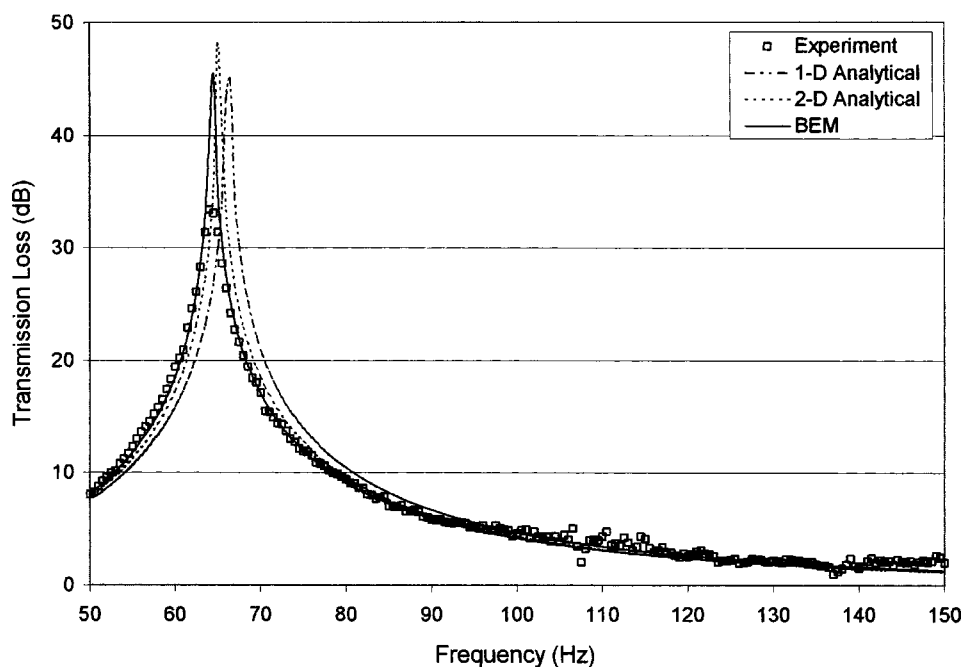


FIG. 8. Comparison of predictions with experiment for a Helmholtz resonator with neck extension ($\ell_1 = 8.5$ cm and $\ell_2 = 15$ cm).

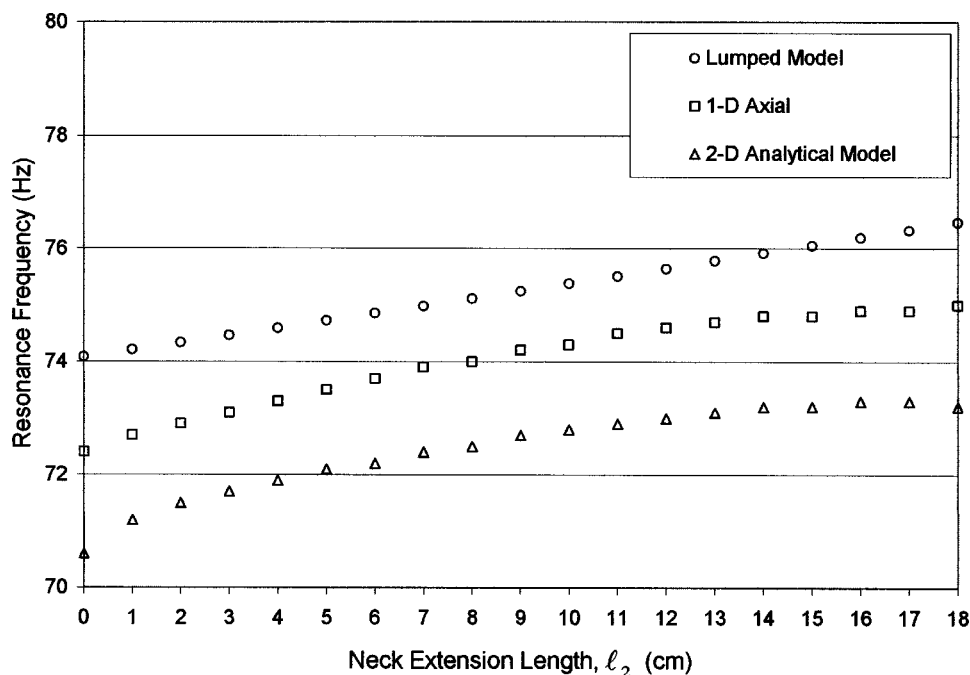


FIG. 9. Predicted resonance frequencies for Helmholtz resonators with fixed total length, $\ell_n = \ell_1 + \ell_2 = 18.5$ cm.

shift in the resonance frequency, which is a significant variation at low frequencies considered here.

In addition to the resonance frequency, transmission loss of the pipe-mounted Helmholtz resonator can be predicted using Eqs. (23) and (32). For a Helmholtz resonator without and with ($\ell_2 = 15$ cm) neck extension [Figs. 1(a) and (b), configuration 1], Figs. 7 and 8 compare such transmission loss predictions with experiments. Since the transmission loss behavior for $\ell_2 = 5$ and 10 cm is similar to Figs. 7 and 8 except for the resonance frequencies, the comparisons for $\ell_2 = 5$ and 10 cm are excluded. While the 2D analytical predictions shift to higher frequencies by 0.5–2 Hz, the 3D BEM shows a close agreement with the experiments. The difference between the 2D analytical method and 3D BEM is due to the neglect of higher order modes in the former at the

neck and main duct interface. As the neck extension length ℓ_2 increases, the differences tend to diminish since the total neck length increases. Both 2D analytical and BEM have peaks higher than the experiment since damping or dissipation effects are neglected in the predictions.

For a fixed total neck length (configuration 2) $\ell_n = 18.5$ cm, the predicted resonance frequency shifts as a function of the extension length ℓ_2 as shown in Fig. 9, which include lumped, 1D, and 2D analytical models. For a given geometry, the variation in resonance frequency remains within 3 Hz for all analytical approaches.

Figure 10 considers a 10-cm-long conical neck extension [Fig. 1(c), configuration 3] and compares the BEM predictions with experiments. Due to the complexity of the geometry, only the BEM predictions are presented here.

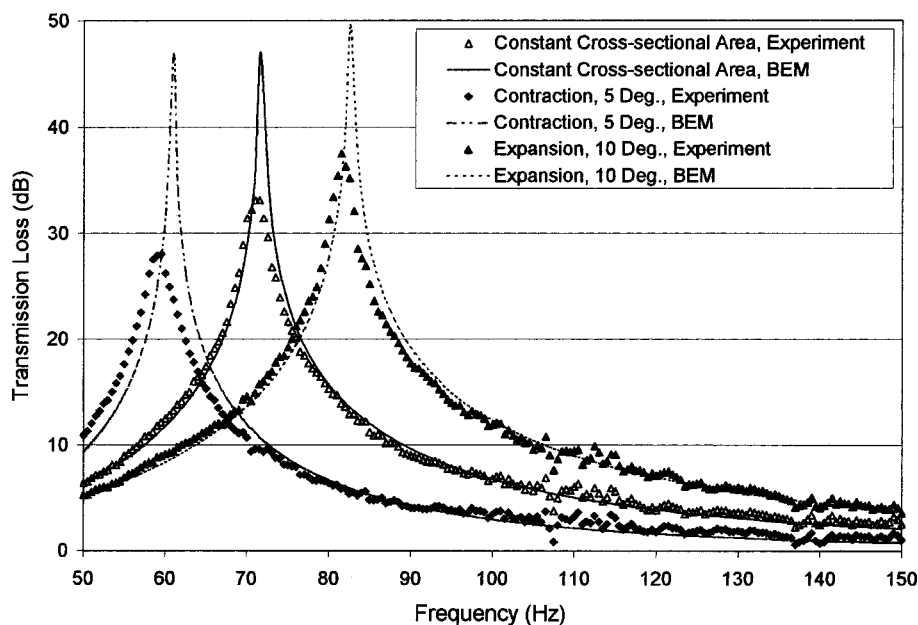


FIG. 10. Comparison of BEM predictions with experiment for a Helmholtz resonator with conical neck extension ($\ell_1 = 8.5$ cm and $\ell_2 = 10$ cm).

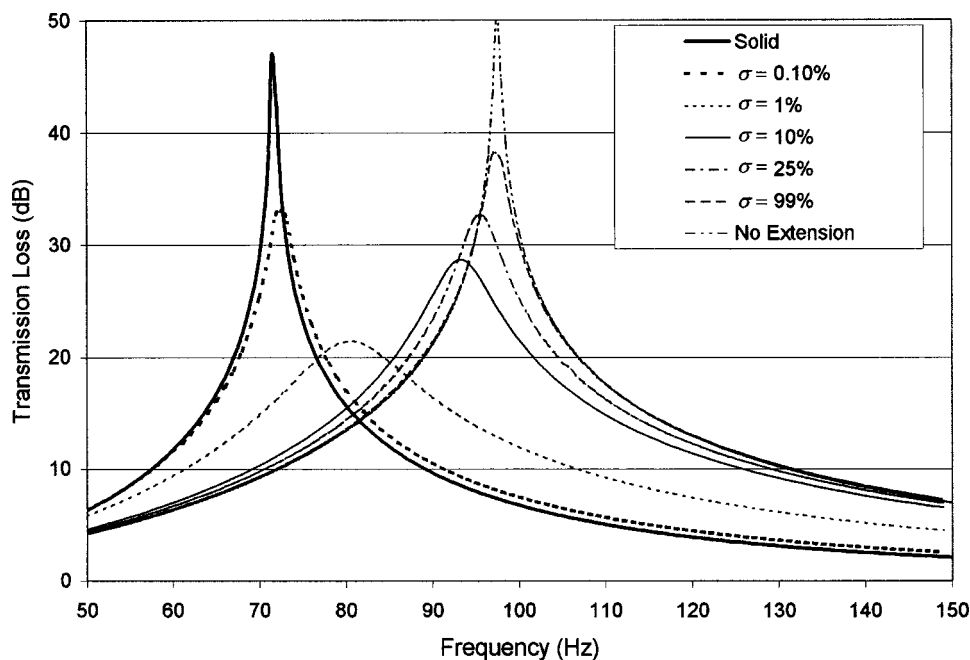


FIG. 11. Comparison of BEM predictions with experiment for a Helmholtz resonator with perforated neck extension ($\ell_1 = 8.5$ cm and $\ell_2 = 10$ cm).

Compared to the one with constant cross-sectional area extension, the resonator with 10° expansion has increased the resonance frequency and broadened the attenuation band, whereas the 5° contraction has lowered the resonance frequency and narrowed the attenuation band. The directional behavior may be qualitatively interpreted in terms of increasing and decreasing effective cross-sectional areas, respectively. The BEM predictions capture these trends with satisfactory accuracy.

As demonstrated in Figs. 6–10, the resonance frequency and transmission loss behavior can be modified by changing the neck length or shape. The resonance frequency can be further controlled by the perforation of the neck extension. Figure 11 shows transmission loss predicted by BEM for Helmholtz resonators with perforated neck extension [Fig. 1(d), configuration 4]. The top of the extension is fully open to the cavity volume and the cylindrical extension with $\ell_2 = 10$ cm is now perforated with porosity σ . As σ increases, the resonance frequency increases, and the magnitude of the transmission first decreases followed by an increase. Sullivan's empirical expression¹⁵ is used for perforation impedance in these predictions. As σ increases, the extension length effectively varies from $\ell_2 = 10$ cm to $\ell_2 \approx 0$ cm at high porosities, therefore the corresponding resonance frequency varies from the limit of solid extension to the no-extension case with accompanying increase in frequency. Also, as σ varies from 1 to 99%, the resistance or damping of the perforation impedance decreases, leading to an increase in transmission loss.

V. CONCLUDING REMARKS

Acoustic characteristics of a Helmholtz resonator with extended neck into cavity have been studied analytically, numerically, and experimentally. Predictions by lumped, 1D axial, 2D analytical, and 3D BEM models are compared to each other and the experimental results. Extension of neck into the cavity lowers the resonance frequency substantially

and narrows the transmission loss band. With the total neck lengths retained, the shape of the neck extension, e.g., conical contraction and expansion, also changes the resonance frequency and transmission loss. Thus, modification of the neck extension length or shape may be an effective method to control the resonance frequency of a Helmholtz resonator without changing the cavity volume. Also adding perforation to the neck extension can change the resonance frequency and transmission loss behavior.

For a piston-driven Helmholtz resonator, the 2D analytical method shows a good agreement with BEM, while simplified lumped and 1D models overpredict the resonance frequencies. For a pipe-mounted resonator, transmission loss predictions by 2D analytical model show reasonable agreement with experiments. The deviation of 2D analytical transmission loss prediction from experiment is mainly due to the neglect of higher order modes in the vicinity of neck and main duct interface. As the total neck length increases, the difference between 2D analytical method and experiments decreases.

The extension of neck into the cavity can shift the resonance frequency down without increasing the volume. Changing the length and shape of or adding perforations to the neck extension can move the resonance frequency and modify the transmission loss behavior. Such trends can be appealing in noise control applications where designers may be limited by the space constraints.

APPENDIX A: INTEGRATION IN EQS. (18)–(20)

The following integral relations for Bessel functions are considered (Abramowitz and Stegun¹⁶):

$$\int x J_0(\alpha x) J_0(\beta x) dx = \frac{x \{ \alpha J_1(\alpha x) J_0(\beta x) - \beta J_0(\alpha x) J_1(\beta x) \}}{\alpha^2 - \beta^2}, \quad (\text{A1})$$

$$\begin{aligned} \int x J_0^2(\alpha x) dx &= \frac{x^2}{2} \{J_0'(\alpha x)\}^2 + \frac{x^2}{2} \{J_0(\alpha x)\}^2 \\ &= \frac{x^2}{2} \{J_1^2(\alpha x) + J_0^2(\alpha x)\}, \end{aligned} \quad (\text{A2})$$

$$\begin{aligned} \int x J_0(\alpha x) Y_0(\alpha x) dr \\ = \frac{1}{2} x^2 \{J_0(\alpha x) Y_0(\alpha x) + J_1(\alpha x) Y_1(\alpha x)\}, \end{aligned} \quad (\text{A3})$$

where J_n and Y_n are the first and second kind Bessel functions, respectively. The integration denoted by $\langle \rangle_s$ in Eqs. (18)–(20) can be calculated from Eqs. (A1)–(A3) and the orthogonality of Bessel functions as follows:

$$\langle \psi_{A,s}, \psi_{A,s} \rangle_A = \int_0^{a_1} r J_0^2\left(\alpha_s \frac{r}{a_1}\right) dr = \frac{a_1^2}{2} J_0^2(\alpha_s), \quad (\text{A4})$$

$$\langle \psi_{C,n}, \psi_{A,s} \rangle_A = \begin{cases} \int_0^{a_1} r J_0\left(\alpha_n \frac{r}{a_2}\right) J_0\left(\alpha_s \frac{r}{a_1}\right) dr = \frac{a_1 \frac{\alpha_n}{a_2} J_0(\alpha_s) J_1\left(\alpha_n \frac{a_1}{a_2}\right)}{\left(\frac{\alpha_s}{a_1}\right)^2 - \left(\frac{\alpha_n}{a_2}\right)^2} & \left(\frac{\alpha_n}{a_2} \neq \frac{\alpha_s}{a_1}\right) \\ \int_0^{a_1} r J_0^2\left(\alpha_s \frac{r}{a_1}\right) dr = \frac{1}{2} a_1^2 J_0^2(\alpha_s) & \left(\frac{\alpha_n}{a_2} = \frac{\alpha_s}{a_1}\right), \end{cases} \quad (\text{A5})$$

$$\langle \psi_{B,s}, \psi_{B,s} \rangle_B = \int_{a_1}^{a_2} r \left\{ J_0\left(\beta_s \frac{r}{a_2}\right) - \frac{J_1(\beta_s)}{Y_1(\beta_s)} Y_0\left(\beta_s \frac{r}{a_2}\right) \right\}^2 dr = \frac{1}{2} a_2^2 \psi_{B,s}^2(a_2) - \frac{1}{2} a_1^2 \psi_{B,s}^2(a_1), \quad (\text{A6})$$

$$\begin{aligned} \langle \psi_{C,n}, \psi_{B,s} \rangle_B \\ = \begin{cases} \int_{a_1}^{a_2} r J_0\left(\alpha_n \frac{r}{a_2}\right) \left\{ J_0\left(\beta_s \frac{r}{a_2}\right) - \frac{J_1(\beta_s)}{Y_1(\beta_s)} Y_0\left(\beta_s \frac{r}{a_2}\right) \right\} dr = \frac{\alpha_n \frac{a_1}{a_2} J_1\left(\alpha_n \frac{a_1}{a_2}\right) \psi_{B,s}(a_1)}{\left(\frac{\beta_s}{a_2}\right)^2 - \left(\frac{\alpha_n}{a_2}\right)^2} & (\alpha_n \neq \beta_s) \\ \int_{a_1}^{a_2} r J_0\left(\alpha_n \frac{r}{a_2}\right) \left\{ J_0\left(\beta_s \frac{r}{a_2}\right) - \frac{J_1(\beta_s)}{Y_1(\beta_s)} Y_0\left(\beta_s \frac{r}{a_2}\right) \right\} dr = \frac{1}{2} a_2^2 J_0(\alpha_n) \psi_{B,s}(a_2) - \frac{1}{2} a_1^2 J_0\left(\alpha_n \frac{a_1}{a_2}\right) \psi_{B,s}(a_1) & (\alpha_n = \beta_s), \end{cases} \end{aligned} \quad (\text{A7})$$

$$\langle \psi_{A,n}, \psi_{C,s} \rangle_A = \begin{cases} \int_0^{a_1} r J_0\left(\alpha_n \frac{r}{a_1}\right) J_0\left(\alpha_s \frac{r}{a_2}\right) dr = \frac{\alpha_s \frac{a_1}{a_2} J_0(\alpha_n) J_1\left(\alpha_s \frac{a_1}{a_2}\right)}{\left(\frac{\alpha_s}{a_2}\right)^2 - \left(\frac{\alpha_n}{a_1}\right)^2} & \left(\frac{\alpha_s}{a_2} \neq \frac{\alpha_n}{a_1}\right) \\ \int_0^{a_1} r J_0^2\left(\alpha_n \frac{r}{a_1}\right) dr = \frac{1}{2} a_1^2 J_0^2(\alpha_n) & \left(\frac{\alpha_s}{a_2} = \frac{\alpha_n}{a_1}\right), \end{cases} \quad (\text{A8})$$

$$\begin{aligned} \langle \psi_{B,n}, \psi_{C,s} \rangle_B \\ = \begin{cases} \int_{a_1}^{a_2} r \left\{ J_0\left(\beta_n \frac{r}{a_2}\right) - \frac{J_1(\beta_n)}{Y_1(\beta_n)} Y_0\left(\beta_n \frac{r}{a_2}\right) \right\} J_0\left(\alpha_s \frac{r}{a_2}\right) dr = \frac{\alpha_s \frac{a_1}{a_2} J_1\left(\alpha_s \frac{a_1}{a_2}\right) \psi_{B,n}(a_1)}{\left(\frac{\beta_n}{a_2}\right)^2 - \left(\frac{\alpha_s}{a_2}\right)^2} & (\alpha_s \neq \beta_n) \\ \int_{a_1}^{a_2} r \left\{ J_0\left(\beta_n \frac{r}{a_2}\right) - \frac{J_1(\beta_n)}{Y_1(\beta_n)} Y_0\left(\beta_n \frac{r}{a_2}\right) \right\} J_0\left(\alpha_s \frac{r}{a_2}\right) dr = \frac{1}{2} a_2^2 J_0(\alpha_s) \psi_{B,n}(a_2) - \frac{1}{2} a_1^2 J_0\left(\alpha_s \frac{a_1}{a_2}\right) \psi_{B,n}(a_1) & (\alpha_s = \beta_n), \end{cases} \end{aligned} \quad (\text{A9})$$

$$\langle \psi_{C,s}, \psi_{C,s} \rangle_C = \int_0^{a_2} r J_0^2\left(\alpha_s \frac{r}{a_2}\right) dr = \frac{a_2^2}{2} J_0^2(\alpha_s). \quad (\text{A10})$$

- ¹U. Ingard, "On the theory and design of acoustic resonators," *J. Acoust. Soc. Am.* **25**, 1037–1061 (1953).
- ²R. C. Chanaud, "Effects of geometry on the resonance frequency of Helmholtz resonators," *J. Sound Vib.* **178**, 337–348 (1994).
- ³P. K. Tang and W. A. Sirignano, "Theory of a generalized Helmholtz resonator," *J. Sound Vib.* **26**, 247–262 (1973).
- ⁴N. S. Dickey and A. Selamet, "Helmholtz resonators: One-dimensional limit for small cavity length-to-diameter ratios," *J. Sound Vib.* **195**, 512–517 (1996).
- ⁵A. Selamet, P. M. Radavich, N. S. Dickey, and J. M. Novak, "Circular concentric Helmholtz resonator," *J. Acoust. Soc. Am.* **101**, 41–51 (1997).
- ⁶A. Selamet and Z. L. Ji, "Circular asymmetric Helmholtz resonators," *J. Acoust. Soc. Am.* **107**, 2360–2369 (2000).
- ⁷K. Nagaya, Y. Hano, and A. Suda, "Silencer consisting of two-stage Helmholtz resonator with auto-tuning control," *J. Acoust. Soc. Am.* **110**, 289–295 (2001).
- ⁸C. R. Cheng, J. D. McIntosh, M. T. Zuroski, and L. J. Eriksson, "Tunable acoustic system," United States Patent 5,930,371, 1999.
- ⁹J. M. de Bedout, M. A. Franchek, R. J. Bernhard, and L. Mongeau, "Adaptive-passive noise control with self-tuning Helmholtz resonators," *J. Sound Vib.* **202**, 109–123 (1997).
- ¹⁰M. L. Munjal, *Acoustics of Ducts and Mufflers* (Wiley, New York, 1987).
- ¹¹J. Miles, "The reflection of sound due to a change in cross section of a circular tube," *J. Acoust. Soc. Am.* **16**, 14–19 (1944).
- ¹²A. Selamet and Z. L. Ji, "Acoustic attenuation performance of circular expansion chamber with extended inlet/outlet," *J. Sound Vib.* **223**, 197–212 (1999).
- ¹³Z. Ji, Q. Ma, and Z. Zhang, "Application of the boundary element method to predicting acoustic performance of expansion chamber mufflers with mean flow," *J. Sound Vib.* **173**, 57–71 (1994).
- ¹⁴M. H. Lawry, *IDEAS Master Series Student Guide* (Structure Dynamics Research Corporation, Milford, OH, 1999).
- ¹⁵J. W. Sullivan and M. J. Crocker, "Analysis of concentric-tube resonators having unpartitioned cavities," *J. Acoust. Soc. Am.* **64**, 207–215 (1978).
- ¹⁶M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).

Noise reduction in a launch vehicle fairing using actively tuned loudspeakers^{a)}

Jonathan D. Kemp

Acentech, Inc., Cambridge, Massachusetts 02138-1118

Robert L. Clark

Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708-0300

(Received 17 April 2001; revised 19 November 2002; accepted 8 January 2003)

Loudspeakers tuned as optimal acoustic absorbers can significantly reduce damaging, low frequency, reverberant noise in a full-scale launch vehicle fairing. Irregular geometry, changing payloads, and the compliant nature of the fairing hinder effective implementation of a passively tuned loudspeaker. A method of tuning the loudspeaker dynamics in real time is required to meet the application requirements. Through system identification, the dynamics of the enclosure can be identified and used to tune the dynamics of the loudspeaker for reduction of targeted, high intensity, low-frequency modes that dominate the acoustic response in the fairing. A loudspeaker model with desired dynamics serves as the reference model in a control law designed to tune the dynamics of a non-ideal loudspeaker to act as an optimal tuned absorber. Experimental results indicate that a tuned loudspeaker placed in the nose cone of the fairing significantly reduces acoustic energy and verifies results calculated from the simulation. © 2003 Acoustical Society of America.

[DOI: 10.1121/1.1558371]

PACS numbers: 43.50.Ki [MRS]

I. INTRODUCTION

The high-intensity, low frequency acoustic excitation that occurs in a rocket fairing at launch can induce structural-acoustic vibrations that damage the payload. As engineers endeavor to create larger, lightweight, cost-effective fairings, acoustic excitation becomes a more critical factor in payload launch survivability. Previous research in this area has investigated both active and passive methods for acoustic attenuation,^{1–3} in addition to vibration control of the fairing itself.^{4,5} At low frequencies, lightly damped modes dominate the interior acoustic response of the fairing, and the ineffectiveness of passive methods such as acoustic blankets, fiberglass, or acoustic foam impels consideration of active and hybrid control methods. This study presents the experimental results of an actively controlled, optimally tuned acoustic absorber implemented in a full-scale fairing testbed.

Previous work has shown that dissipative control can be achieved in acoustic enclosures using a collocated pressure sensor and a constant volume-velocity source.^{6,7} Extension of this principle yielded significant attenuation of low frequency acoustic modes in a launch vehicle fairing by utilizing arrays of sensors and actuators working cooperatively to achieve global control. The transducer arrays were spatially weighted in order to selectively couple to low-frequency modes, which reduced the order of the controller and yielded significant modal control with minimal control spillover.⁸ Global peak reductions (8–10 dB) of targeted acoustic modes required an extensive array of control sources, condi-

tioning and amplification measures, microphones, control hardware, and wiring. Further review of these requirements has prompted investigation of less equipment intensive, “hybrid” active/passive controllers.

Previous research involving hybrid systems for sound absorption demonstrated the effectiveness of surface impedance control in rectangular enclosures^{9,10} and impedance tubes¹¹ with low modal density. Both loudspeaker diaphragms and panels with tuned dynamics demonstrated attenuation through active impedance changes at enclosure boundaries, but none demonstrated acoustic absorption from within the enclosure. More relevant to this work, the authors joined an effort to develop a passive vibroacoustic attenuator based on a hybrid of a proof-mass actuator (PMA) and a shunted loudspeaker.¹² That study highlighted the possibility of combining a tuned mass damper for structural damping and a tuned diaphragm for acoustic damping. The attenuator was dynamically modeled and placed within a fully coupled structural-acoustic finite element model of the fairing. Results predicted a broadband noise reduction of 3 dB, but difficulties in the experimental development and application of such an attenuator left room for improvement.

As an extension of the hybrid principle, this work applies tuned mass damping theory to sound absorption applications in a flexible, irregular cavity that encloses changing payloads and demonstrates significant, three-dimensional modal complexity. This study proposes to attenuate low frequency modes in a launch vehicle fairing with loudspeakers tuned as optimal acoustic absorbers. The absorbers promote efficiency of the control system while demonstrating control results comparable to previous active studies. As a primary advantage, the significant logistical issues of typical active systems have been eliminated. The control law discussed herein actively tunes a loudspeaker to passively reduce the

^{a)}Portions of this work were presented in “Noise reduction in a launch vehicle fairing using actively tuned loudspeakers,” Proceedings of Joint 140th Meeting of the Acoustical Society of America and NOISE-CON 2000, Newport Beach, CA, 2000.

targeted mode(s), removing typical acoustic spillover in the control bandwidth, and allowing global attenuation with a single actuator. A parametric optimization scheme selects an optimal tuned absorber model which is compared to an identified model of the non-ideal loudspeaker in the fairing; the difference between them is minimized by a model reference control law, based on H_2 design. Loudspeaker tuning requires only low-order control in a limited bandwidth. These achievements increase the performance and stability of the controller, while reducing control energy and hardware by more than 90%.

Preliminary investigation of actively tuned acoustic absorber design facilitates application of that technology for noise control in acoustic enclosures, and more specifically, in an experimental, full-scale, composite launch vehicle fairing model at Duke University. In support of continuing efforts to improve payload survivability, the research sought definitive control results to promote further study of the effectiveness and efficiency of adaptive acoustic control within the fairing volume.

II. THEORY

A. Fairing acoustics

A rocket fairing is essentially a flexible-walled structure surrounding an acoustic enclosure that protects the payload during launch. Fairings have typically been constructed of metal, such as aluminum, but composites have been used more recently. Although composites offer many advantages, they are less massive and therefore allow increased transmission of exterior disturbances. Vibration of the fairing excites acoustic resonances within the enclosure, which detrimentally couple to the payload. Data from previous satellite launches indicate that the overall acoustic levels inside the fairing during launch can exceed 140 dB. At low frequencies (10 to 200 Hz), the levels reach approximately 120 to 130 dB.¹³

The fairing considered in this work is based on the Orbital Sub-orbital Program Space Launch Vehicle (OSPSLV) minotaur. The shroud is 5.3 meters long with a maximum diameter of approximately 1.3 meters. Therefore, the low-frequency modes are longitudinal rather than radial. The fairing is excited by a number of disturbances, including aerodynamic buffeting during flight,¹⁴ structural vibrations induced by the rocket motors, and pyrotechnic shocks during stage separations. These sources act in addition to the explosive noise produced by the motors themselves.

The most harmful vibration occurs in the first 30–45 seconds after ignition, as both acoustic transmission and aerodynamic buffeting decrease with decreasing atmospheric pressure and density. While on the launch pad, the earth acts as an acoustic baffle, reflecting enormous waves of energy at the payload. As the vehicle lifts away from earth, increased speed causes significant aerodynamically induced noise. When the payload exhibits a structural resonance near the acoustic resonances of the fairing, the acoustic loading can result in appreciable damage to the payload. The Department of Defense reports that at least 40% of first-day satellite failures result from vibration damage incurred at launch.¹⁵ Flat

surfaces of solar panels and light-weight structures such as thin films, membranes, and precision optics are particularly susceptible to damage from the low frequency excitation. The presence of multiple, random, low-frequency disturbance sources precludes passive and feedforward control, compelling the use of an innovative feedback control scheme.

B. Tuned acoustic absorbers

The concept of loudspeakers tuned as optimal acoustic absorbers grew from the familiar use of tuned mass-spring systems for vibration absorption in structures. Here, active tuning controls enclosed acoustics. Departing from previous acoustic approaches which attempted to change the enclosure boundary, this work develops the ability of production loudspeakers to act in reverse, or as self-contained tunable absorbers. Essentially, a loudspeaker is actively tuned to respond passively as though it were an ideal absorber for a targeted modal frequency.

Realization of the ideal absorber begins with development of a theoretical model that couples the response of the fairing enclosure to a disturbance source and a typical loudspeaker. An optimization scheme then selects an ideal absorber for reduction of global acoustic energy at a targeted modal frequency. A model reference control law can then minimize the response difference between the ideal model and the actual loudspeaker, forcing that loudspeaker to act as an ideal absorber.

III. MODELING, DESIGN, AND OPTIMIZATION

The combined behavior of the fairing disturbances is unpredictable, but the modal acoustic response in the enclosure due to the vibration of the fairing structure can be easily modeled. A rigid walled cylinder model of equal dimension provides a strong prediction of the magnitude, phase, and frequency of the lowest modal frequencies, and is constructed here for the purposes of absorber design and optimization. The development follows Morse and Ingard¹⁶ and Cheng and Nicolas.¹⁷ Application of the appropriate boundary conditions allows solution of the homogeneous wave equation in a cylinder of length L with a Hankel function of order m :

$$p(r, \theta, z, t) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \sum_{\mu=1}^{\infty} A_{mn\mu} \sin\left(m\theta + \gamma \frac{\pi}{2}\right) \times J_m(k_r r) \cos\left(\frac{\mu\pi}{L} z\right) e^{i(\omega t - k_\theta \theta - k_{z_{mn\mu}} z)}, \quad (1)$$

where p represents the pressure in the fluid, c is the local speed of sound, and r , z , and θ are the cylindrical coordinates within the cavity. Here, m , n , and μ , and γ are the azimuthal, radial, longitudinal, and symmetric (rotational) modal indices, respectively. Here, k_θ is the azimuthal wave number and k_z is the longitudinal wave number. The radial wave number k_r is the n th root of the derivative of the Bessel function of the first kind evaluated at the cavity wall.

The model includes a vibrating rectangular panel, curved and placed at the enclosure wall to simulate the trans-

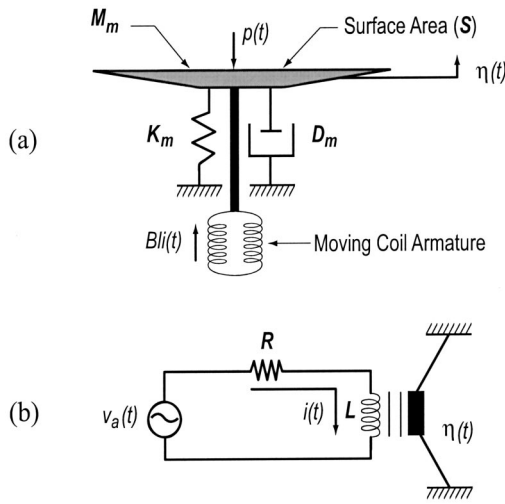


FIG. 1. (a) Mechanical schematic of loudspeaker system. (b) Electrical schematic of loudspeaker system [after Clark *et al.* (Ref. 20)].

mission path in the actual fairing. Following fluid-structure coupling techniques, the modal interaction model by Fahy,¹⁸ and recent work by Henry and Clark,¹⁹ the model results. From the inhomogeneous wave equation for the pressure in the cavity, the equations of motion for the enclosure as disturbed by the curved panel can be written as

$$\Lambda_{n_p} \ddot{v}_{n_p}(t) + \Lambda_{n_p} \omega_{n_p}^2 v_{n_p}(t) = -\rho_0 c^2 S_0 \sum_{k=1}^K C_{n_p k} \ddot{w}_k(t), \quad (2)$$

where ρ_0 is the density of the fluid in the cavity, and $v_{n_p}(t)$ is a generalized modal coordinate. The panel undergoing out-of-plane displacement, w , is modeled as a distributed volume velocity source just inside the enclosure boundary. The modal volume, Λ_{n_p} , is represented by the integral of spatial mode shape functions, and the subscripts m , n , and μ have been replaced in by the generalized modal index n_p . The coupling coefficient $C_{n_p k}$ is the surface integral of the product of the mode shape functions of the panel and the enclosure.

A. Tuned absorber coupling

Modeling and design of the tunable acoustic absorber requires coupling the equations of motion for a typical loudspeaker to the acoustics of the enclosure model developed above. The passive interaction of the tuned loudspeaker with the fairing acoustics forms the tunable boundary condition for control application. In preliminary tests, the nose cone seemed ideal for installation of the tuned absorber: the primary active control targets are the modes of lowest frequency, so the absorber should interact with longitudinal, not radial or azimuthal, wave forms. These wavelengths are most easily targeted spatially by observation and control at the end caps of the fairing. Applied damping at the boundary condition responsible for longitudinal modal response will reduce acoustic energy at the modal frequency, thereby reducing overall acoustic energy propagation within. Additionally, placement of the tuned absorber in the nose cone minimizes the control system impact on available payload volume.

A model of a loudspeaker, radiating in its piston mode, was therefore incorporated in one end cap of the enclosure. The electromechanical coupling and structural acoustic coupling are described by assimilating the typical loudspeaker equations of motion²⁰ with the enclosure model above. Figure 1(a) provides a schematic representation of the mechanical system, and the respective equation of motion can be written as

$$M_m \ddot{w}(t) + D_m \dot{w}(t) + K_m w(t) = Bli(t) - S_0 p(t), \quad (3)$$

where M_m , D_m , and K_m are the mechanical mass, damping, and stiffness of the speaker, and $w(t)$ represents the displacement of the speaker coil. The force input has two components: electromotive force is the product of the field strength of the inductor B , the conductor length l , and the current $i(t)$; and the pressure input is the product of the surface area of the loudspeaker diaphragm, S_0 , and the acoustic pressure, $p(t)$. This pressure input, or radiation resistance, is the most crucial coupling term for this application. Figure 1(b) provides a schematic representation of the electrical system, governed by the following equation:

$$L \frac{di(t)}{dt} + R_s i(t) = v_a(t) - Bl\dot{w}(t). \quad (4)$$

Note that the voice coil (Bl) couples the electrical and mechanical systems in the speaker. These equations of motion can now be coupled through diaphragm displacement w and acoustic pressure p to the acoustic model of the enclosure for full system modeling and optimal design.

B. Optimization

A constrained optimization scheme determines the optimal mechanical and electrical parameters of the loudspeaker for acoustic energy absorption at a targeted mode in the enclosure. Within reasonable constraints, the routine could select candidate speakers by arbitrary variation of each significant property of the loudspeaker system: B , l , R , M_m , K_m , D_m , and S_0 . However, to enhance the feasibility of actual loudspeaker construction and tuning, the parameter field was narrowed to those available for physical tuning: the resistance R of the circuit, and the apparent spring stiffness, K_m , of the mechanical system. The cost functional was calculated as the H_2 norm of the targeted modal acoustic output v_n of the coupled loudspeaker/enclosure system. The routine utilized a constrained optimization function provided in the optimization toolbox for MATLAB. The function, *fmincon.m*, follows a sequential quadratic programming method in which a quadratic programming subproblem with linearized constraints is solved at each iteration as the routine searches for the minimum of the given cost functional. A positive definite quasi-Newton estimate of the Hessian of the Lagrangian function is updated at each iteration to provide the direction and magnitude of the next iteration. This estimate is calculated with the BFGS (Broyden, Fletcher, Goldfarb, and Shanno) formula which enhances the efficiency of the scheme.²¹

To assure the accuracy of the optimization results, several sets of initial conditions were passed to the routine for

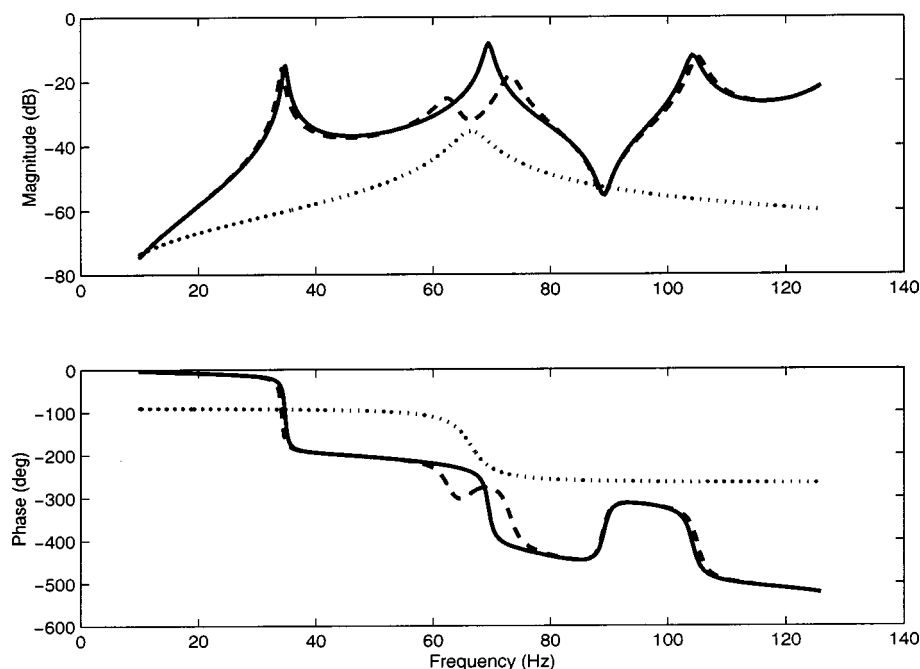


FIG. 2. Frequency response of enclosure without (—) and with (— — —) reduction due to coupled tuned absorber (···).

each targeted mode. The routine repeatedly selected similar models within one hundred iterations. The dynamic response of a typical optimal absorber demonstrates increased response at a frequency slightly above or below the targeted modal frequency. This observation confirmed initial hypotheses about the parallels between optimal acoustic absorbers and tuned vibration control devices. Figure 2 displays the predicted acoustic response of an optimal loudspeaker coupled to the acoustic enclosure developed above. Note the significant reduction in response at the second mode. This pressure response, taken in the acoustic far field of the tuned loudspeaker, demonstrates significant global reduction from a single tuned actuator. This ideal actuator serves as the reference model in a control scheme designed to tune the actual loudspeaker to act as the ideal absorber.

IV. EXPERIMENTAL IMPLEMENTATION

Development of the tuned absorber continues with application of model reference control to the loudspeaker in the fairing. Based upon work by Ogata,²² the development of a model reference plant involves evaluation of a performance

metric calculated as the response difference between the actual system and the model reference system, given identical inputs.

An approach based on the eigensystem realization algorithm^{23,24} identifies low-order, discrete, state-space realizations of the untuned speaker for comparison with the optimal loudspeaker model developed above. A common source then disturbs both the realized system and the optimal model through the respective voltage inputs, ensuring that the controller affects the loudspeaker regardless of acoustic pressure input. The characteristics of the disturbance and its coupling to the fairing are irrelevant to the controller, broadening the scope of possible application. The outputs of the two systems are compared and the difference between them serves as the performance variable in the control law.

The augmented system^{25,26} was then assembled and modeled using Simulink,²⁷ and H_2 control laws were computed using MATLAB. Filters designed to shape the controller are added for stability but increase the order (number of states) of the augmented plant and the resulting control law. The controller, now of unnecessarily large order, is reduced

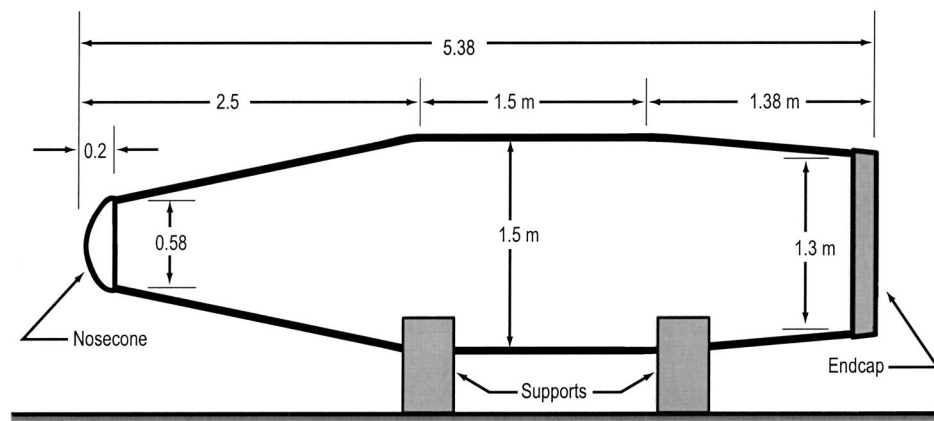


FIG. 3. Illustration of fairing testbed.

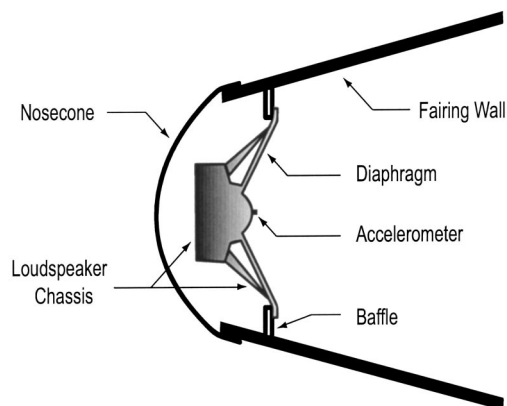


FIG. 4. Schematic of absorber system mounted in nose cone.

through model reduction techniques, such as balanced residualization or truncation.^{26,28} After computation, a Tustin routine transforms the reduced-order control law into the discrete domain and the control law is downloaded to a digital signal processor (DSP) for implementation.

Comparison of the open-loop and closed-loop frequency response functions from a disturbance input (band-limited random noise) to an array of performance microphones yields a controller performance metric. Averaging these frequency response measurements provides representative open-loop and closed-loop frequency responses and facilitates evaluation of the overall effects of the controller as a function of frequency. From this evaluation, the occurrence of spillover, the degree of coupling to the target modes, and the average amount of local and global attenuation is observed.

A. Fairing testbed

Experiments are conducted on a full scale composite fairing model containing the loudspeaker/absorber, the accelerometer sensor, the acoustic disturbance source, and sixteen arbitrarily positioned performance microphones. The fairing (shown schematically in Fig. 3) is approximately 5.3 meters in length, 1.3 meters in diameter (maximum), and tapered at both ends. A plywood end cap is attached to the base of the fairing and sealed appropriately. A hemispherical aluminum end cap completes the nose cone and seals the acoustic

chamber behind the loudspeaker, as shown in Fig. 4. The loudspeaker and baffle are rigidly attached to the nose cone. The loudspeaker is a manufactured 18 inch subwoofer, occupying roughly 60% of the area of the nose cone baffle.

The accelerometer is fixed at the diaphragm center for maximum signal and focus on the piston mode of the speaker. A small conditioning device located within the fairing amplifies the accelerometer signal. A disturbance loudspeaker in the corner of the fairing near the base end cap excites the interior cavity modes. The performance microphones were distributed throughout the interior at arbitrary positions to measure the overall controller effects. All cabling was connected through a panel at the base of the fairing. The controller, spectrum analyzers, power amplifiers, microphone conditioners, and other required hardware were housed external to the fairing. A block diagram of the setup is presented in Fig. 5.

V. RESULTS

A. Tuned absorber

The experiment begins with an investigation of the coupling effects between the loudspeaker and the fairing acoustics. Figure 6 supports the predicted effects of this coupling. The frequency response of the loudspeaker is measured both in open air and after installation in the fairing nose cone baffle. Both were measured as frequency response functions between a band-limited random disturbance applied to the voltage input of the loudspeaker and the voltage signal from the accelerometer mounted on the loudspeaker diaphragm. Note the strong modal response when the loudspeaker is coupled to the fairing acoustics. The modes apparent here are coupled mechanoacoustic modes of the loudspeaker and fairing, not merely a reflection of the fairing acoustics. This coupling is crucial to the success of the tuned absorber, as discussed in Sec. III, and verifies conjecture about this pressure input to the dynamic system modeled herein.

Modeling of the loudspeaker dynamics for the purpose of control is accomplished through system identification techniques. Following identification, the control law is applied to tune the loudspeaker as an optimal absorber. The controller effectively changes the dynamic response of the loudspeaker, as shown in Fig. 7. The loudspeaker response is

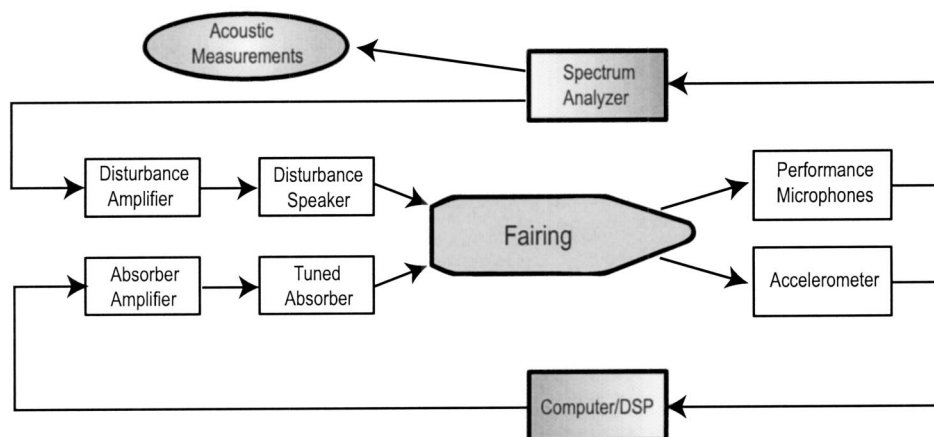


FIG. 5. Schematic of experimental setup.

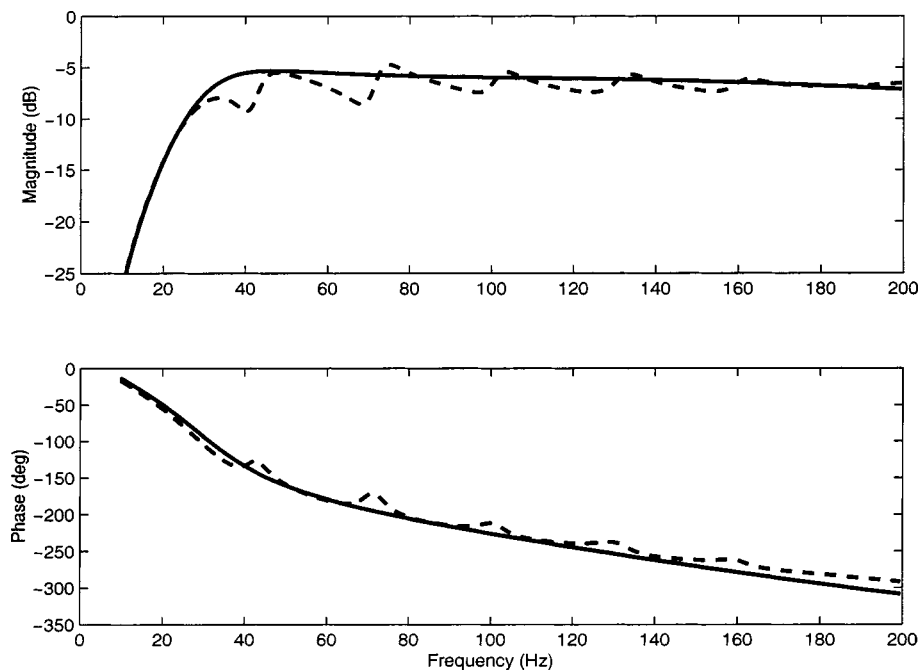


FIG. 6. Acoustic coupling demonstrated by loudspeaker responses measured from voltage input to accelerometer in open air (—) and after mounting in nose cone baffle (---).

measured by the accelerometer relative to an acoustic disturbance applied to the fairing volume. Note the significant peak introduced at the targeted frequency of 65 Hz. Figure 7 details the effects of the controller when the disturbance is purely acoustic; the modes of the fairing are apparent. The controller does not retard response of the loudspeaker at the non-targeted first acoustic mode, so the response of the speaker at that frequency remains unchanged.

Model reference control demonstrates authority over the dynamics of an actual loudspeaker. By design, the control law can tune the loudspeaker to absorb at any frequency, though some frequencies are more conducive to tuning than others. The efficiency of the controller and the magnitude of resulting effects are dependent on the dynamics of the speaker, as expected. If the desired tuning frequency is proximal to a coupled mode of the loudspeaker and fairing, then the controller is more successful, as the effort required to

move a system pole is significantly less than that required to introduce a new pole. Additionally, the modal targeting introduced by the noise filters in the control law prevents placement of control effort out of the most efficient bandwidth. An inherently efficient controller results from this combination of control strategies.

B. Acoustic control in the fairing

Demonstration of a stable tuned loudspeaker allows implementation in the fairing volume. Key to the stability and minimal energy consumption of the controller is the separation of the controller from the acoustics of the fairing. The control law seeks only to tune the mechanics of the loudspeaker, not to actively control the fairing acoustics. This separation precludes the development of typical acoustic instabilities found in most active acoustic control systems,

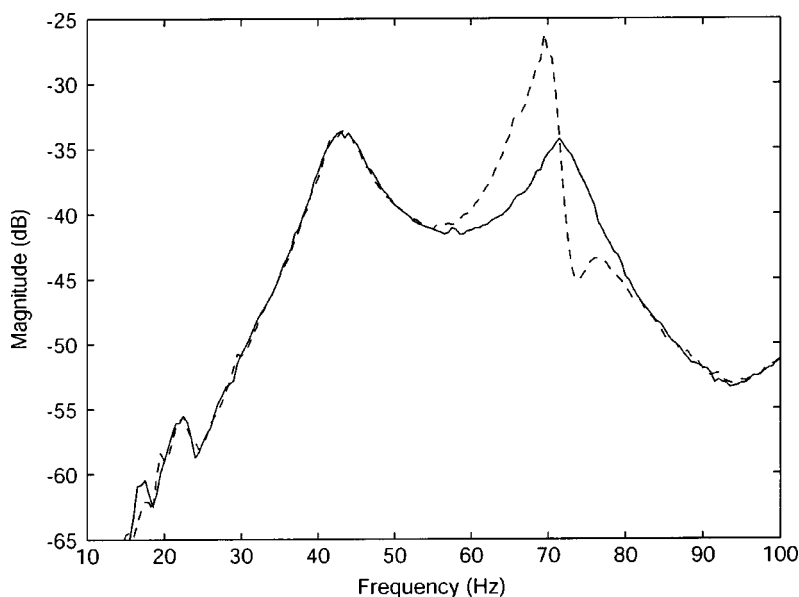


FIG. 7. Loudspeaker acceleration response measured from acoustic disturbance to accelerometer without (—) and with active tuning (---).

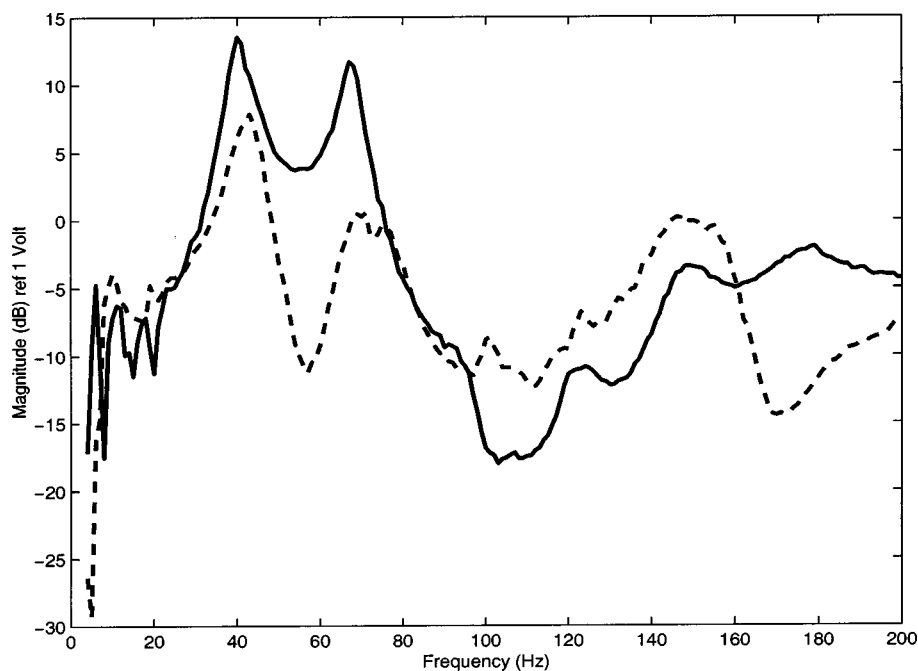


FIG. 8. Local acoustic pressure response measured from acoustic disturbance to a single microphone (No. 3) without (—) and with active tuned absorber (---) targeting second acoustic mode (69 Hz).

allowing significant control signal gain without fear of unstable acoustic feedback. In other words, the controller actively tunes the loudspeaker to act as a passive acoustic absorber, as discussed in the preceding sections.

Optimization efforts revealed significant control capability at the second mode of the fairing, 69 Hz. Proximity to an actual coupled modal frequency of the loudspeaker and fairing promotes effective control of loudspeaker dynamics. Candidate absorber models with natural frequencies slightly above or below the targeted 69 Hz were chosen by the optimization routine. The best results, both predicted and actual, utilized an absorber tuned to 80 Hz. The optimal tuning frequency is roughly 15% above the targeted modal frequency, a result typical of tuned mass absorber design. Figure 8 displays the local pressure response for this control configura-

tion, measured at a microphone located in the cross-sectional center of the fairing, approximately 0.8 meters from the tuned diaphragm. The 12 dB reduction at the second mode (69 Hz) is apparent; the resonant peak is split and of greatly reduced magnitude. This microphone represents the response of the fairing in the acoustic near field of the tuned loudspeaker.

An evaluation of global performance in the fairing yields the acoustic response shown in Fig. 9. This response was measured as the simple average of the pressures recorded at 16 arbitrary positions within the fairing. Significant global peak reduction (4 dB) is apparent; control is achieved.

Note that the global reduction is less significant than the local control shown in Fig. 8, as expected. To investigate the performance at individual microphones throughout the space,

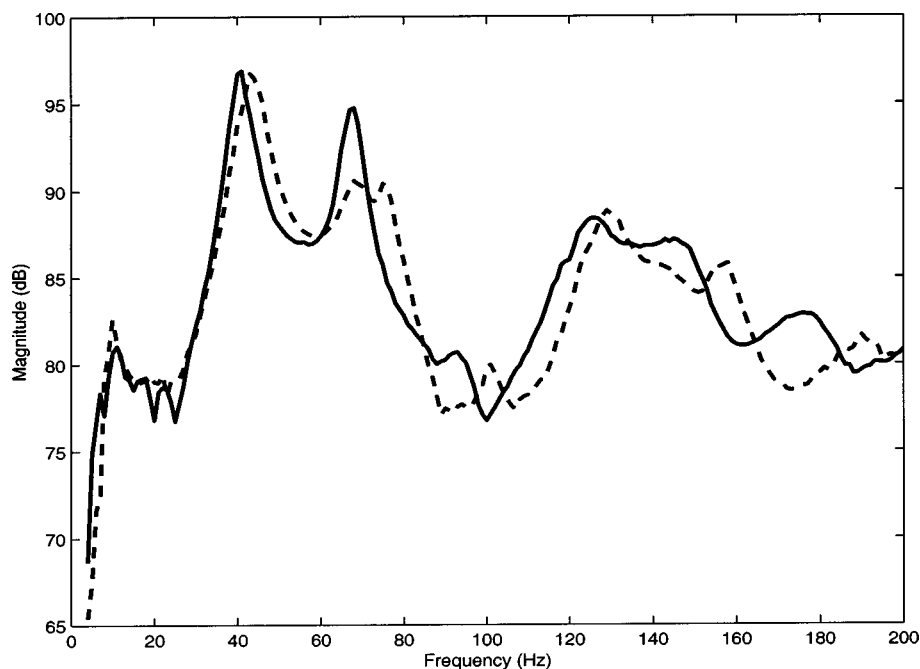


FIG. 9. Global acoustic pressure response, as averaged from acoustic disturbance to 16 microphones, without (—) and with active tuned absorber (---) targeting second acoustic mode (69 Hz).

TABLE I. 69 Hz mode peak pressure reductions achieved with an actively tuned loudspeaker, calculated from measurements of frequency responses between acoustic disturbance and each microphone.

Microphone ID	Peak reduction (69 Hz)
1	7 dB
2	8 dB
3	12 dB
4	8 dB
5	0 dB
6	1 dB
7	0 dB
8	6 dB
9	6 dB
10	5 dB
11	12 dB
12	4 dB
13	-3 dB
14	3 dB
15	5 dB
16	5 dB

Table I details the second mode (69 Hz) peak reduction measured at each location. The microphone locations are numbered consecutively from forward to aft. The local frequency response plotted in Fig. 8 corresponds to microphone number 3, as the first two positions were closer to the loudspeaker but not in the center of the fairing. Note that attenuation is achieved at 13 of the 16 microphone locations, the exceptions being two microphones where no attenuation was noted and a 3 dB increase measured at microphone number 13. However, the overall RMS pressure level in the 200 Hz bandwidth was still reduced by 0.61 dB at microphone No. 13. Additionally, two separate locations, one in the acoustic near field (No. 3) and one in the far field (No. 11) measured peak reductions of 12 dB. The previous active system requiring 16 actuators did not achieve any peak reduction greater than 12 dB.⁸

A promising observation throughout the tests is the voltage necessary to achieve this level of control. A frequency spectrum of the control voltage is provided in Fig. 10. This controller, by design as merely an enabling device for pas-

sive absorption, required 0.216 Volts RMS (1.180 V peak) for the duration of the testing. These levels are at least an order of magnitude less than the voltage required per actuator by active systems achieving similar levels of control with multiple actuators.⁸ Power requirements are also appropriately reduced. Each of the above results demonstrates the effectiveness of a low-order model-reference controller used to tune a loudspeaker to act as an optimal tuned acoustic absorber in a full-scale payload fairing.

VI. CONCLUSION

Preliminary results presented here demonstrate the effectiveness of a tuned loudspeaker acting as a self-contained optimal acoustic absorber in a launch vehicle fairing. Application of structural tuned mass absorber theory to the realm of acoustics has allowed efficient acoustic reduction with active technology but without the logistical cost of previous active and passive control apparatus.

Following development of a theoretical model of the enclosure and of a typical loudspeaker, an optimization scheme selects the dynamic properties of an ideal absorber for a chosen acoustic mode of the enclosure. This optimal absorber then serves as the reference model in a control law designed to actively tune an off-the-shelf loudspeaker to act as an optimal tuned absorber. Placement of this tuned loudspeaker in the nose cone of a full-scale launch vehicle fairing allows experimental verification of the tuned absorber concept. A tuned loudspeaker acting passively in the fairing enclosure absorbs significant acoustic energy at targeted modal frequencies.

The feasibility of a launchable application remains the quest of this technological development. The advantages here are simple: the actively tuned absorber minimizes controller impact on payload weight and volume. The system requires only one actuator per acoustic mode to be controlled, rather than the array used for active studies. With each actuator goes an accompanying amplifier, microphone, and signal conditioner. Further, control signal voltage and actuator power requirements are orders of magnitude less

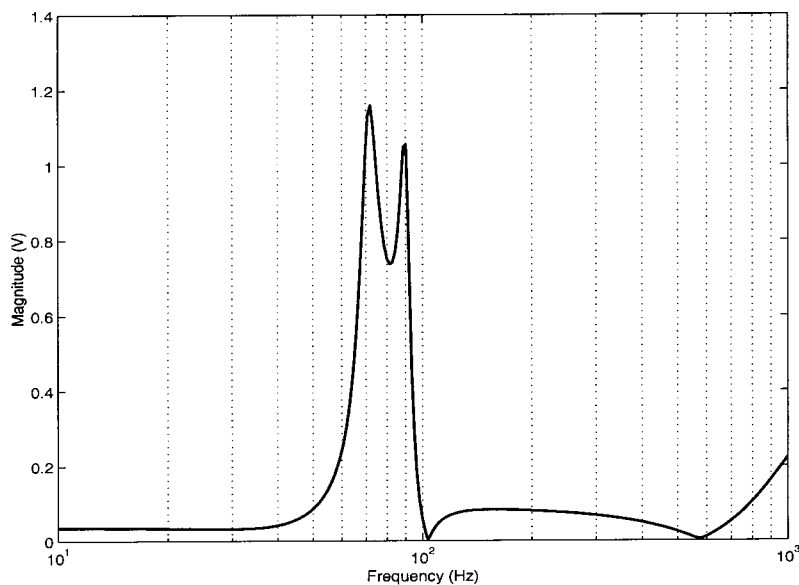


FIG. 10. Frequency spectrum of the model reference control signal voltage for the actively tuned loudspeaker.

than active counterparts, relieving battery weight. This hybrid controller accomplishes efficient, lightweight noise attenuation in a frequency range where both active and passive control have significant logistical and performance problems.

Future work will address challenges that arise in the application of any controller to a launchable enclosure. Acoustic response of the fairing, and therefore optimization and modal targeting parameters, will change with payload and atmospheric conditions, so the tuned system must be able to react accordingly. The tuned absorber system offers feasible solutions. Because the tuning parameters are based solely on a targeted frequency, they could be updated without continuous system identification. In fact, these tracking adjustments could be incorporated in the control law. Additionally, further studies should include the use of multiple tuned loudspeakers targeting single and multiple frequencies in the enclosure. Development of these concepts holds promise for future investigations, as the actively tuned loudspeaker has demonstrated acoustic control in a full-scale fairing.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge the partial support for this research provided under the AFOSR Presidential Early Career Award for Scientists and Engineers, Grant No. F49620-98-1-0383, monitored by Major Brian Sanders and Dr. Dan Segalman. We would also like to express our gratitude to the Air Force Research Laboratory for providing the rocket fairing model used in these experiments.

- ¹L. Bradford and J. Manning, "Attenuation of Cassini spacecraft acoustic environment," *Sound Vib.* **30**, 303–337 (1996).
- ²K. Weissman, M. E. McNellis, and W. D. Pordan, "Implementation of acoustic blankets in energy analysis methods with application to the atlas payload fairing," *J. IES* **34**, 32–39 (1994).
- ³C. Niezrecki and H. Cudney, "Preliminary review of active control technology applied to the fairing acoustic problem," Proceedings of the AIAA Adaptive Structures Forum, Salt Lake City, UT, 1996, pp. 101–108.
- ⁴D. Leo and E. Anderson, "Vibroacoustic modeling of a launch vehicle payload fairing for active acoustic control." AIAA 39th Annual SDM Conference, Long Beach, CA, 1998, AIAA-98-2086.
- ⁵F. Shen and D. Pope, "Design and development of composite fairing structures for space launch vehicles," SAE Technical Paper 901836, Aerospace Technology Conference and Exposition, Long Beach, CA, October 1990.
- ⁶S. A. Lane and R. L. Clark, "Dissipative feedback control of a reverberant enclosure using a constant volume velocity source," *J. Vib. Acoust.* **120**, 987–993 (1998).

- ⁷S. A. Lane and R. L. Clark, "Improving loudspeaker performance for active noise control applications," *J. Audio Eng. Soc.* **46**, 508–519 (1998).
- ⁸S. A. Lane, J. D. Kemp, R. L. Clark, and S. Griffin, "Feasibility analysis for active acoustic control of a rocket fairing using spatially weighted transducer arrays," *J. Spacecr. Rockets* **38**, 112–119 (2001).
- ⁹M. Furstoss, D. Thenail, and M. A. Galland, "Surface impedance control for sound absorption," *J. Sound Vib.* **203**, 219–236 (1997).
- ¹⁰O. Lacour, M. A. Galland, and D. Thenail, "Preliminary experiments on noise reduction in cavities using active impedance changes," *J. Sound Vib.* **230**, 69–99 (2000).
- ¹¹J. P. Smith, B. D. Johnson, and R. A. Burdisso, "A broadband passive-active sound absorption system," *J. Acoust. Soc. Am.* **106**, 2646–2652 (1999).
- ¹²S. F. Griffin, J. Gussy, S. A. Lane, J. D. Kemp, and R. L. Clark, "Innovative passive mechanisms for control of sound in a launch vehicle fairing," Proceedings of the 41st Structures, Structural Dynamics, and Materials Conference, Atlanta, GA, 2000.
- ¹³*Multi-Service Launch System Payload Planners Guide* (Martin Marietta Technologies, Inc., Denver, CO, 1994). Prepared for the USAF Space and Missile Command/CUBM.
- ¹⁴K. Dotson, R. L. Baker, and B. H. Sako, "Launch vehicle buffeting with aeroelastic coupling effects," *J. Fluids Struct.* **14**, 1145–1171 (2000).
- ¹⁵Calming bad vibes in launch vehicles, Ballistic Missile Defense Organization Update Newsletter, Summer 1999.
- ¹⁶P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
- ¹⁷L. Cheng and J. Nicolas, "Radiation of sound into a cylindrical enclosure from a point-driven end plate with general boundary conditions," *J. Acoust. Soc. Am.* **91**, 1504 (1992).
- ¹⁸F. Fahy, *Sound and Structural Vibration* (Academic, London, 1985).
- ¹⁹J. K. Henry and Robert L. Clark, "Noise transmission from a curved panel into a cylindrical enclosure: Analysis of structural-acoustic coupling," *J. Acoust. Soc. Am.* **109**, 1456–1463.
- ²⁰R. L. Clark, K. D. Frampton, and D. G. Cole, "Phase compensation for feedback control of enclosed sound fields," *J. Sound Vib.* **195**, 710–718 (1996).
- ²¹T. Coleman, M. A. Branch, and A. Grace, *Optimization Toolbox User's Guide* (The Mathworks, Inc., Natick, MA, 1999), pp. 2-1–2-44. Third printing Revised for Version 2(R11).
- ²²K. Ogata, *Modern Control Engineering* (Prentice-Hall, Upper Saddle River, NJ, 1997), pp. 912–915.
- ²³J. Juang, *Applied System Identification* (Prentice-Hall, Englewood Cliffs, NJ, 1994).
- ²⁴J. Juang and R. S. Pappa, "An eigensystem realization algorithm for modal parameter identification and model reduction," *J. Guid. Control* **8**, 620–627 (1985).
- ²⁵R. L. Clark, G. P. Gibbs, and W. R. Saunders, *Adaptive Structures, Dynamics and Control* (Wiley, New York, 1998).
- ²⁶S. Skogestad and I. Postlewaite, *Multivariable Feedback Control* (Wiley, Chichester, England, 1996).
- ²⁷*Simulink, Dynamic Simulation for Matlab* (The Mathworks, Inc., Natick, MA, 1999). Revised for Version 2(R11).
- ²⁸*Robust Control Toolbox User's Guide* (The Mathworks, Inc., Natick, MA, 1999). Revised for Version 2(R11).

Experimental investigation of spatial correlation in broadband reverberant sound fields

Ingyu Chun,^{a)} Boaz Rafaely, and Phillip Joseph

Institute of Sound and Vibration Research, University of Southampton, Southampton, Hampshire SO17 1BJ, United Kingdom

(Received 31 July 2001; revised 13 September 2002; accepted 13 January 2003)

The spatial correlation has previously been investigated for tonal and narrow-band sound fields. This letter presents an experimental investigation of the spatial correlation coefficients in a reverberation chamber driven by broadband signals. The main objective is to verify recent theoretical results for broadband spatial correlation in diffuse sound fields. Experimental results show good agreement with theoretical predictions when the frequency band of the sound field is entirely above the Schroeder frequency. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1558375]

PACS numbers: 43.55.Br, 43.55.Cs [MK]

I. INTRODUCTION

The spatial correlation coefficient is an important quantity for applications in which the diffuse field assumption is made. For example, it can give an indication of the minimum permissible spacing of microphones necessary to obtain independent measurements of mean-square sound pressure for the determination of sound power levels in reverberation chambers.¹ Moreover, the analysis of zones of quiet for local active sound control in enclosed sound fields uses the spatial correlation coefficients of diffuse sound fields.²

Many authors have previously studied theoretically and experimentally the spatial correlation coefficients for pure-tone³ and narrow-band diffuse fields.⁴ Rafaely⁵ has recently studied the spatial-temporal correlation function theoretically in a diffuse field excited by broadband random signals with prescribed power spectral density. This letter presents experimental verification of these theoretical results. Section II discusses the spatial correlation coefficient; Sec. III describes the setup, procedure, and results of the experiment, and Sec. IV provides conclusion.

II. SPATIAL CORRELATION COEFFICIENT

The spatial-temporal correlation coefficient ρ between two position vectors \mathbf{x} and \mathbf{y} at times t_1 and t_2 , respectively, is defined as³

$$\rho(\mathbf{x}, t_1, \mathbf{y}, t_2) = \frac{E\{p(\mathbf{x}, t_1)p(\mathbf{y}, t_2)\}}{\sqrt{E\{p^2(\mathbf{x}, t_1)\}E\{p^2(\mathbf{y}, t_2)\}}}, \quad (1)$$

where $E\{\}$ denotes the expectation operation, and p is the instantaneous acoustic pressure.

A diffuse sound field consists of an infinite number of statistically uncorrelated plane progressive waves.⁴ If a sound field is assumed to be perfectly diffuse, and stationary and ergodic with respect to time and space,⁴ the spatial correlation coefficient $\rho(r)$ with zero time delay can be written as⁵

$$\rho(r) = \frac{1}{\sigma^2} \int_{-\infty}^{\infty} S(f) \frac{\sin(kr)}{kr} df, \quad (2)$$

where $S(f)$ is the power spectral density of the acoustic signal in a perfect diffuse field, f is the frequency in Hertz, k is the wave number, r is the separation distance between two position vectors \mathbf{x} and \mathbf{y} ($r = |\mathbf{y} - \mathbf{x}|$), and σ^2 is the variance of the excitation signal. Equation (2) shows that the broadband spatial correlation coefficient is the integration over frequency of the pure-tone spatial correlation coefficients,³ i.e., a sinc function $[\sin(kr)/kr]$, weighted by the normalized power spectral density of the excitation signal.

When the broadband spatial correlation coefficients are predicted in practice using Eq. (2), this theoretical result has two main sources of estimation errors originating from the validity of the two terms in Eq. (2): the sinc function and the power spectral density. The pure-tone spatial correlation coefficient curve can deviate substantially from the sinc function in reverberant sound fields with the degree of deviation depending upon the degree of diffuseness at the corresponding frequency.³ The spectrum $S(f)$ in Eq. (2) also includes the frequency response of the transducer and associated electrical equipment, such as loudspeaker and microphone, and the irregular frequency response of the room, which could degrade the estimation.

The Schroeder frequency⁶ is the crossover frequency that marks the transition from individual distinct resonances to the case where there are many overlapping modes. It is expected that the estimation errors originating from both terms in Eq. (2) in the frequency range below the Schroeder frequency are generally higher than the corresponding errors in the frequency range above the Schroeder frequency, under same conditions. This is due to the weak modal overlap of the room and lack of diffuseness at low frequencies.

III. EXPERIMENT

A. Experimental setup and procedure

Equation (1) has been used to determine the spatial correlation coefficients from acoustic pressure measurements in a rectangular reverberation chamber of 348 m³ volume. The Schroeder frequency of the reverberation chamber is calcu-

^{a)} Author to whom correspondence should be addressed. Electronic mail: ic@isvr.soton.ac.uk

lated as 350 Hz, obtained as a average value of the Schroeder frequencies calculated from measured reverberation times in one-third octave bands with center frequencies from 100 Hz to 1000 Hz, assuming a speed of sound of 346.4 ms^{-1} . Standard deviation of these Schroeder frequencies is 36 Hz and the average value of the reverberation times is 9.5 s. Broadband random test signals with three different frequency bandwidths were used in this experiment, which were generated by filtering white noise signals through high order digital butterworth filters.⁷ The first test signal of bandwidth between 350 Hz and 1 kHz was chosen to investigate a broadband sound field of strong modal overlap at frequencies entirely above the Schroeder frequency. The second signal of bandwidth between 0 Hz and 350 Hz was chosen to investigate a broadband sound field of weak modal overlap at frequencies entirely below the Schroeder frequency. The third signal of bandwidth between 0 Hz and 1 kHz was chosen to investigate a broadband sound field at frequencies both below and above the Schroeder frequency. The orders of the digital butterworth filters are 4 for the first signal, 8 for the second signal, and 16 for the third signal. Each solid line in Fig. 1 shows the power spectral density of each test signal measured at one typical position in the reverberation chamber. Each dashed line in Fig. 1 shows the theoretical values of power spectral density of each test signal at the output of the digital bandpass or lowpass filter.⁷ The measured power spectral density is different from the theoretical spectrum because the frequency response of the transducer and associated electrical equipments, the irregular frequency response of the room, and finite averaging modifies the theoretical power spectral density. However, this difference can be partially compensated for in the prediction of the spatial correlation coefficient by substituting the measured power spectral density for the theoretical one in Eq. (2). Note that random fluctuation in the measured power spectral density about the mean behavior does not significantly affect the estimate of the spatial correlation in Eq. (2) because this fluctuation becomes smoothed once integration over frequency has been performed.

Figure 2 shows a block diagram of the sound generation system and the data acquisition system.⁷ The loudspeaker of 12 in. diameter was placed in the corner of the reverberation chamber. Two Brüel & Kjær microphones of type 4135 were used in the measurement with phase mismatch of less than 2.5° between 20 Hz and 1 kHz. The digital signals were generated at a sampling rate of 16 384 Hz by a PC, and were then converted to analog signals for driving the loudspeaker. The analog signals acquired by the microphones were converted to digital signals with a sampling rate of 8192 Hz. The acoustic signals at the two microphones were recorded simultaneously for a 10 s duration under steady state conditions. The signals were measured at two microphones with 15 incremental separation distances r . The spatial correlation for each separation distance r was estimated by averaging the correlations spatially at 10 different positions and orientations, chosen randomly in the chamber.

It is important to note that if a sufficiently large number of spatial averages at random positions and orientations are performed, almost any sound field shows the same spatial

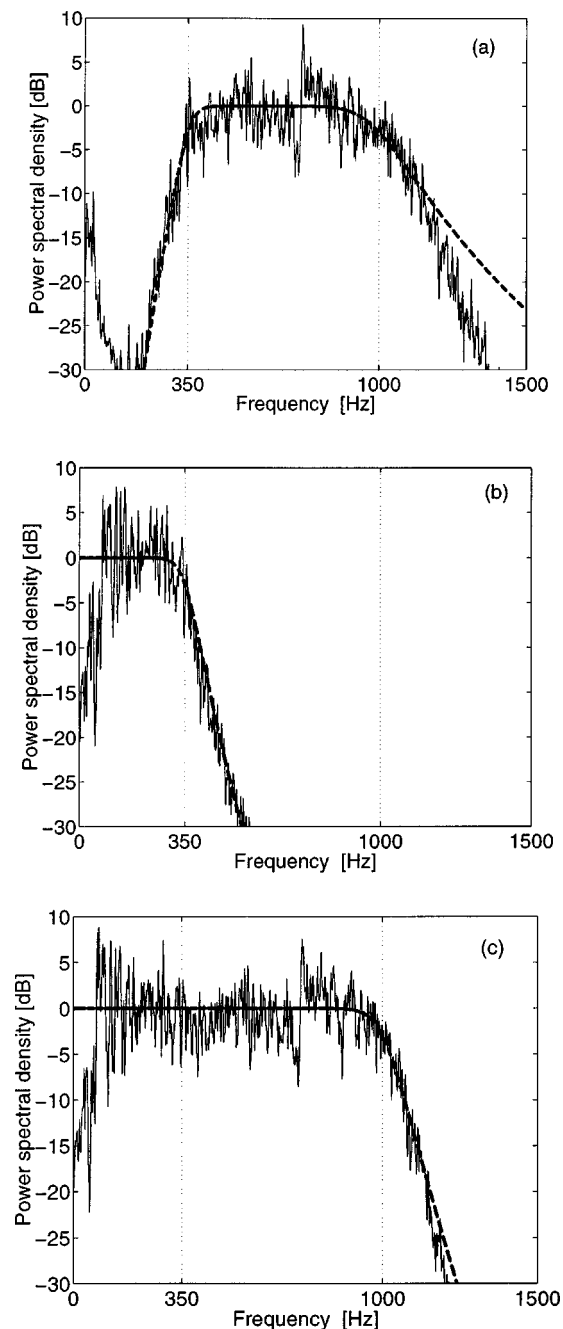


FIG. 1. Power spectral density of the test signals of bandwidths: (a) 350 Hz–1 kHz, (b) 0 Hz–350 Hz, (c) 0 Hz–1 kHz (solid line: measured values, dashed line: theoretical values).

correlation characteristics as a diffuse sound field.⁸ Therefore, the number of spatial averages required in a given sound field depends on the degree of diffuseness and the requirements for accuracy of the specific application. In this experiment, the number of spatial averages was fixed to ten, which was found to be sufficient to achieve a reasonable degree of randomness. Note that a greater number of spatial averages for pure tone sound fields are required to obtain the same degree of diffuseness under the same conditions.⁷ This is because the degree of diffuseness in broadband sound fields is much larger than that in pure-tone or narrow-band sound fields under the same conditions.

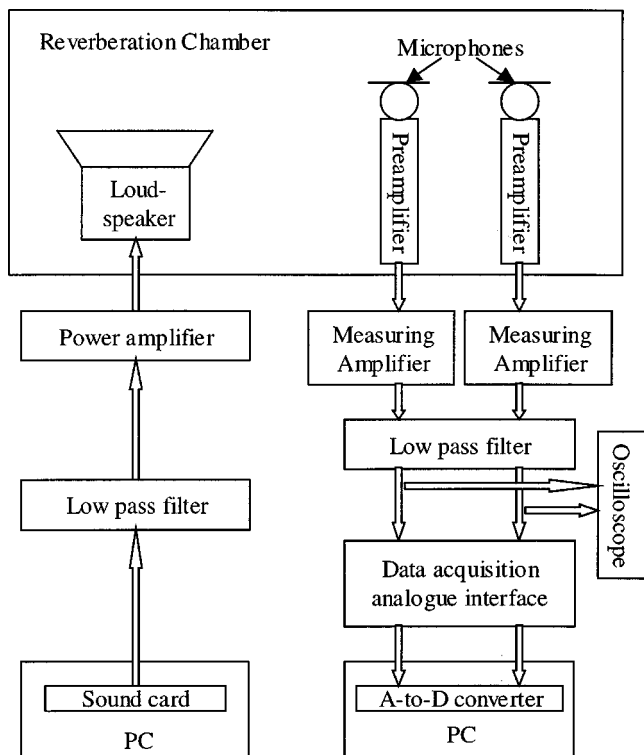


FIG. 2. Block diagram of the measurement system.

B. Results

The measured spatial correlation coefficients for the three test signals are represented by the circles in Fig. 3. The total length of the error bar at each measured value is twice the standard deviation of the estimates of the spatial correlation coefficient at the corresponding separation distance r . The solid lines in Fig. 3 represent the theoretical predictions obtained by using the theoretical values of $S(f)$ in Eq. (2). The dashed lines in Fig. 3 represent the theoretical predictions using the measured values of $S(f)$ in Eq. (2).

Figure 3(a) shows the spatial correlation coefficients of the first test signal of bandwidth between 350 Hz and 1 kHz. The difference between the measured and theoretically predicted spatial correlation coefficient is less than ± 0.03 . Figure 3(b) shows the spatial correlation coefficients of the second signal of bandwidth between 0 Hz and 350 Hz. The difference between the measurement and the theoretical predictions, when the theoretical value of $S(f)$ is used, is larger than that obtained when the measured $S(f)$ is used. Therefore, the use of measured power spectral density gives better agreement between theoretical prediction and measurement in this case. Agreement between measurement and theoretical prediction is poorer than that in case of other test signals. Figure 3(c) shows the spatial correlation coefficients of the third signal of bandwidth between 0 Hz and 1 kHz. Although the signal includes frequency bands below the Schroeder frequency in this case, the difference between measurement and theoretical prediction is typically less than ± 0.065 . The use of measured power spectral density does not improve agreement between measurement and theoretical prediction in this case. A possible explanation is that the sound field in the frequency band above the Schroeder frequency dominates

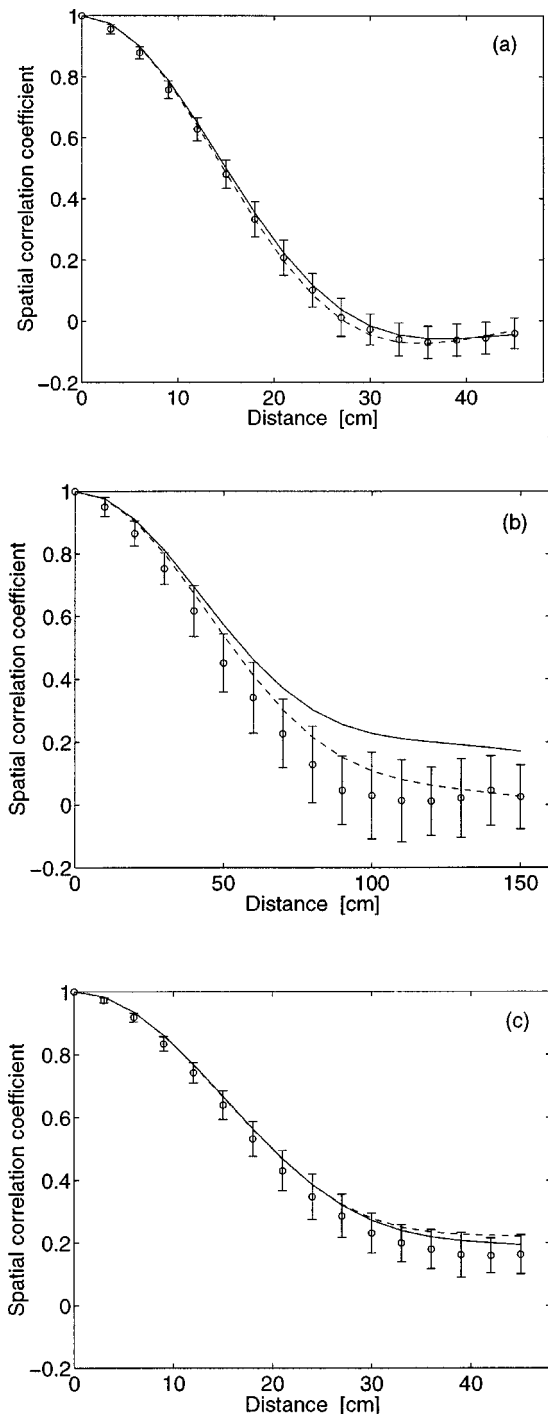


FIG. 3. Spatial correlation coefficient of the test signals of bandwidths: (a) 350 Hz–1 kHz, (b) 0 Hz–350 Hz, (c) 0 Hz–1 kHz (circles: measured values, error bar: standard deviation of the estimate, solid line: theoretical predictions using the theoretical power spectral density, dashed line: theoretical predictions using the measured power spectral density).

the sound field in the frequency band below it, so that the estimation error is relatively small.

IV. CONCLUSIONS

The spatial correlation coefficients of broadband sound fields were measured in a reverberation chamber. Good agreement between theoretical predictions and experimental results was obtained when frequency band of the sound field

was entirely above the Schroeder frequency. Agreement was poorer when the frequency band of the sound field was entirely below the Schroeder frequency. However, agreement can be improved by using the measured power spectral density of the excitation signal in the theoretical predictions in some cases. Reasonably agreement was obtained when the frequency band of the sound field was both above and below the Schroeder frequency.

¹D. Lubman, "Spatial averaging in a diffuse sound field," *J. Sound Vib.* **46**, 532–534 (1969).

²B. Rafaely, "Zones of quiet in a broadband diffuse sound field," *J. Acoust. Soc. Am.* **110**, 296–302 (2001).

³R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thomson, Jr., "Measurements of correlation coefficients in reverberant sound fields," *J. Acoust. Soc. Am.* **27**, 1072–1077 (1955).

⁴F. Jacobsen, "The diffuse sound field," The Acoustics Laboratory, Technical University of Denmark, Report No. 27 (1979).

⁵B. Rafaely, "Spatial-temporal correlation of a diffuse sound field," *J. Acoust. Soc. Am.* **107**, 3254–3258 (2000).

⁶M. R. Schroeder, "On frequency response curves in rooms, comparison of experimental, theoretical, and Monte Carlo results for the average frequency spacing between maxima," *J. Acoust. Soc. Am.* **34**, 76–80 (1962).

⁷I. Chun, "Experimental investigation of spatial correlation of broadband diffuse sound fields," M. S. thesis, ISVR, University of Southampton, 2000.

⁸F. Jacobsen and T. Roisin, "The coherence of reverberant sound fields," *J. Acoust. Soc. Am.* **108**, 204–210 (2000).

Speech transmission index from running speech: A neural network approach

F. F. Li^{a)} and T. J. Cox

School of Acoustics and Electronic Engineering, University of Salford, Salford M5 4WT, United Kingdom

(Received 20 August 2001; revised 30 October 2002; accepted 14 January 2003)

Speech transmission index (STI) is an important objective parameter concerning speech intelligibility for sound transmission channels. It is normally measured with specific test signals to ensure high accuracy and good repeatability. Measurement with running speech was previously proposed, but accuracy is compromised and hence applications limited. A new approach that uses artificial neural networks to accurately extract the STI from received running speech is developed in this paper. Neural networks are trained on a large set of transmitted speech examples with prior knowledge of the transmission channels' STIs. The networks perform complicated nonlinear function mappings and spectral feature memorization to enable accurate objective parameter extraction from transmitted speech. Validations via simulations demonstrate the feasibility of this new method on a one-net-one-speech extract basis. In this case, accuracy is comparable with normal measurement methods. This provides an alternative to standard measurement techniques, and it is intended that the neural network method can facilitate occupied room acoustic measurements.

© 2003 Acoustical Society of America. [DOI: 10.1121/1.1558373]

PACS numbers: 43.55.Mc, 43.58.Gn, 43.60.Lq [MK]

I. INTRODUCTION

Traditionally, two different approaches are used to quantify room acoustics: subjective and objective assessments. Subjective measurements are based on human perception and so usually use music or speech signals. Objective measurements, on the other hand, use artificial test signals such as noise, to ensure reproducibility and repeatability. Much room acoustics research concerns quantifying acoustic quality in terms of objective parameters and this enables designs to be readily made and evaluated. Objective measurements, however, use high sound pressure levels that are usually unacceptable to audiences. This hinders objective measurements under occupied in-use conditions. Occupied measurements are important because it is well established that occupancy affects the acoustic, especially the absorption and background noise levels. It is suggested that many of the problems encountered in occupied objective parameter measurements could be overcome if the naturally occurring signals in a space, such as music or speech, were used as test signals. A technique to achieve this for speech is given in this paper.

Speech Transmission Index (STI) is a common objective parameters used to assess speech intelligibility of spaces and other transmission channels, such as classrooms, theaters, public address and telephony systems.^{1–4} STI combines two major phenomena that affect speech intelligibility—reverberation and noise—to extract a single index that gives good correlation with subjective perception.⁵ Moreover, STI method utilizes simple artificial test signals and enables portable instrumentation to be implemented.⁶ Consequently, STI has been adopted as a mainstream objective parameter for speech intelligibility⁷ and is included in standards and per-

formance specifications. Nevertheless, it is known that in certain circumstances STI is not completely successful.⁸

The normal STI measurement method uses artificial test signals and so is not particularly well suited to occupied measurement. The standard method takes about 10–15 minutes to perform,⁹ which is a long time to expect occupants to listen to noise and yet continue with their normal activities. For this reason, RaSTI was developed in the mid 1980s, and more recently, another new technique, STI-PA has been proposed.⁹ Both RaSTI and STI-PA have reduced measurement time, but the test signal is still noise, and consequently neither are true noninvasive test techniques. Steeneken and Houtgast did, however, propose a method to estimate the modulation transfer function (MTF) and in turn STI from running speech.¹⁰ It was suggested that the MTF is roughly estimated by comparing the envelope spectra of source and received speech signals. This technique works, but at a cost of compromised accuracy. For this reason, practical measurements of STI are rarely made with running speech, but still rely predominantly on artificial test signals.

Inspired by the fact that humans can sensitively differentiate reverberation times, artificial intelligence methods have been developed as a means of extracting objective parameters from speech. Previously, a time domain approach was applied to extract reverberation parameters from separate monosyllable word utterances.¹¹ The time domain method, however, is not applicable to running speech and has signal to noise ratio problems when estimating octave band parameters. A neural network method to estimate STI from running speech was proposed and a few pilot results published.¹² Since then, many refinements have been made, especially to the preprocessor, to form a more accurate and robust method. This paper will present the details of this new method to accurately estimate STI from running speech excerpts using artificial neural networks (ANNs). ANNs are

^{a)}Current address: Department of Computing and Mathematics, Manchester Metropolitan University, Manchester M1 5GD, UK.

systems that can perform nonlinear mapping, in this case from running speech to objective parameters. ANNs learn a mapping through experience, in this case by being exposed to many examples of running speech in rooms and the corresponding objective parameter values of the room. The inputs to the neural network are 60 s speech excerpts, preprocessed with an envelope spectrum estimator. This preprocessing exploits and extends the underlying mechanism of the envelope spectrum technique developed for the standard STI method. Refinements are needed to allow better resolutions and maintain a reasonable number of input neurons to facilitate machine learning. As it is the basis for many aspects of the project, the paper starts by discussing the envelope spectrum and STI method.

II. STI METHOD AND ENVELOPE SPECTRA

The STI method is closely linked to the envelope of running speech. Speech needs to retain its original envelope to be intelligible, the more the envelope is modified the poorer the intelligibility. A room acts as a low pass filter of speech envelopes, smoothing the envelope and hence degrading intelligibility. Moreover, ambient noise disturbs speech signals and reduces intelligibility. Both reverberation and noise cause the normalized low frequency envelope spectrum of speech to decrease.¹³ Consequently, the MTF was introduced to quantify the combined effect of reverberation and ambient noise by means of modulation index reductions.² By properly combining and processing selected frequencies in the MTF, a single index STI is formed.³ The measurement of the MTF is the core process in determining STI, and three different methods exist. Once MTFs are obtained, the STI can be calculated from 98 MTF data points via a series of weighting, limiting and averaging processes.^{3–7} The three methods for obtaining the MTF are as follows.

A. Standard method

The envelope shaping effect of a transmission channel is determined by measuring the MTFs in seven octave bands (125 Hz–8 kHz) using sine-wave modulated noise with its spectrum shaped to be the same as typical long-term speech.⁷ The modulation transfer function MTF is determined by the ratio of modulation index of output to input intensity. This process needs to be carried out for each of the 14 modulation frequencies, and hence the process is relatively slow. The advantage of this method is that it works with nonlinear systems, such as many public address systems.

B. Impulse response method

Schroeder¹⁴ systematically reviewed and discussed MTF measurement methods from a signal and system perspective, and reconfirmed a relationship between MTF and room impulse response $h(t)$:

$$\text{MTF}(F) = \frac{|\int_0^\infty h^2(t) e^{-2\pi j F t} dt|}{\int_0^\infty h^2(t) dt} \frac{1}{(1 + 10^{(-s/n)/10})}, \quad (1)$$

where s/n is the signal to noise ratio, and allows the effect of ambient noise interference to be included. This part of Eq. (1) was added by Steeneken and Houtgast.⁴ Therefore, an MTF can be obtained by first measuring the impulse response. This is often done using a maximum length sequence signal or a swept sine wave. This approach becomes invalid when a system is nonlinear as the impulse response in Eq. (1) should be linear. This can be a serious limitation. For example, many speech reinforcement and public address systems use compressors to improve intelligibility.

C. Using the speech envelope spectrum

A short rectangular window is moved along a running speech signal, typically 40–60 s long. The square of the windowed portion is divided by the average value of the squared long-time speech signal; this gives the intensity function.¹³ An average of the intensity function is taken, and the low frequency envelope spectra of the function found. The part of the envelope spectra important to speech intelligibility lies in frequency band 0–20 Hz. It was suggested that MTFs could be roughly estimated from the envelope spectra of original and transmitted speech.¹⁰ Let $E_X(F)$ be the envelope of original speech and $E_Y(F)$ be the envelope spectrum of received speech, then $\text{MTF}(F)$ is estimated by

$$\text{MTF}(F) \approx E_Y(F)/E_X(F). \quad (2)$$

This method was validated by empirical results showing that the STI obtained using this approach and measured through standard methods have a reasonably good agreement (a correlation coefficient of 0.971).¹⁰ In both the original paper and the standards,^{5,10} however, it is pointed out that MTFs obtained from speech envelopes using Eq. (2) have compromised accuracy.

III. NEURAL NETWORK METHOD

A. Rationale

Room effects are contained in the difference between received, $Y(\omega)$, and transmitted, $X(\omega)$, speech signals. The use of input and output envelope to gain the MTF as suggested above, can be regarded as a linear approximation of a squared linear time invariant filter:

$$\begin{aligned} Y(\omega) &= H(\omega)X(\omega), \\ E_Y(F) &\approx \text{MTF}(F)E_X(F), \end{aligned} \quad (3)$$

where $H(\omega)$ is the transfer function of the room, which precisely describes the input–output relationship of the signals. The MTF, on the other hand, approximately relates the envelopes of the input and output signals.

Such a relation described by MTFs would be accurate if (a) the envelope of speech were periodic and (b) the spectrum of speech were white with constant power per unit bandwidth.¹⁴ Unfortunately, running speech is a complicated, nonstationary stochastic process, only approximately conforming to these criteria. This makes the mapping relationship between the envelope spectra and MTFs imprecise. Artificial neural networks are therefore considered to perform

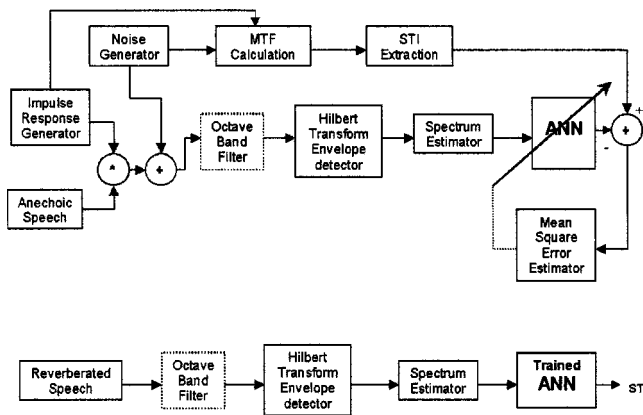


FIG. 1. Block diagram of the training (top) and retrieval (bottom) phases of the system.

this nonlinear mapping. The mapping is likely to be more accurate, because ANNs are inherently powerful nonlinear mapping engines.^{15,16}

A speech excerpt can be regarded as a speech envelope modulated noise, once it is chosen, its envelope spectrum and the spectrum of the carrier is fixed and the difference between these and the standard test signal can be obtained. The ANN algorithm is used to perform nonmodel based regression to memorize features of speech samples and compensate the errors found in the estimated MTFs. From MTF to STI is a deterministic nonlinear limiting and linear combination process. ANNs are known to be able to calculate any computable nonlinear and linear functions and therefore should be able to perform these calculations.¹⁷ As one step further, it is possible to combine the accurate MTF estimation and the subsequent calculation of STI value into one neural network system, i.e., use the neural algorithm to map received envelope signals directly onto STI values. This one stage mapping is useful as it simplifies implementation. Such an ANN network method is illustrated as a block diagram in Fig. 1.

Two phases, training and retrieval are involved in the development of the ANNs and applying them to real world problems. In the training phase, the ANNs learn from examples, memorize related information and generalize from a closed set of training data to a class of cases including those not in the training set. This is achieved by presenting a large number of examples to the ANN and utilizing multivariable optimization techniques to minimize the total errors between the actual STI value and the output of the ANN. Examples used to training the ANN are generated using simulation techniques. Convolutions of anechoic speech and simulated impulse responses are used as transmitted speech examples. The expected STI values are obtained from the impulse responses and additional ambient noise. The knowledge of the original envelope spectrum is implicitly built into the ANN as part of training. As a result, there is no need to monitor the original speech in the retrieve phase. This reduces two channel measurement to one channel. However, as shall be discussed later, the drawback is that it is limited to a one-speech-one-net scheme, i.e., a particular neural network

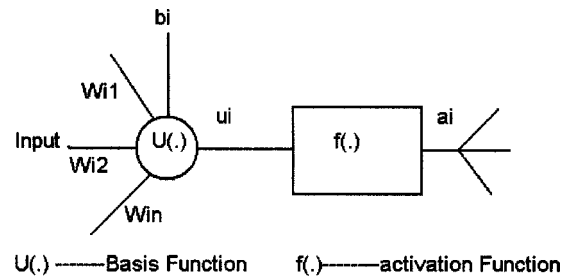


FIG. 2. The neuron model.

learns to memorize features of a particular speech excerpt and works with that excerpt only.

B. Neural network architecture

A nonlinear, multilayer, feed forward network trained by the back-propagation algorithm¹⁷⁻¹⁹ is chosen for this work. Backpropagation is a successful algorithm; the penalty of using this type of neural network can be in excessive training times. The fundamental building blocks of the neural network is the cellules processing neuron unit as depicted in Fig. 2. Typically, it comprises two functions: a linear basis function used to gather input signals, and an activation function $f(\cdot)$ to nonlinearly process the information. The basis function used is

$$u_i = \sum_{j=1}^n w_{ij}x_j + b_i, \quad (4)$$

where w_{ij} is the weight connecting the j th neuron to i th neuron, and x_j is the output of j th neuron. The activation function is

$$a_i = f(u_i) = \frac{1}{1 + e^{-u_i}}. \quad (5)$$

The activation function is used for the two hidden layers to provide nonlinear mapping capability. The neural network is constructed in a feed forward fashion by interconnecting a large number of these simple neurons as shown in Fig. 3. The leftmost input layer takes signals from the preprocessor and distributes them to subsequent layers without processing the signals. There are two nonlinear hidden layers. The STI is a normalized index from 0 to 1. As a common ANN design

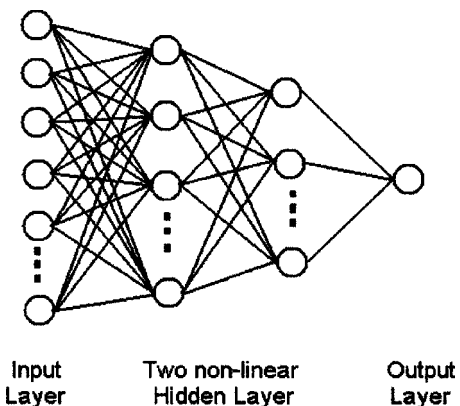


FIG. 3. Multilayer feed forward network architecture.

consideration, a hard limiting nonlinear neuron at the output layer could be used to clamp the output values so that no output can possibly go beyond the interval of 0 to 1; however, it is found in this application that hard limiting the output reduces the back propagation of errors to the hidden layers in the early stages of training, and hence inhibits convergence. So, a linear summation function without a nonlinear activation function is adopted as the output neuron.

The neural network is trained under a supervised model as an approximator. The training, as shown in Fig. 1, is to iteratively apply the preprocessed speech examples to the input of the neural network and minimize the mean square errors between the teachers (known STIs) and the output of the neural network over all examples in the training set. The optimization is done by iteratively updating the connection weights (w_{ij} and b_i) within the neural network using the well-known back propagation algorithms.^{15,17-19} Training is achieved using the delta learning rule:^{15,16}

$$w_{ij}^{(m+1)} = w_{ij}^{(m)} + \eta \Delta w_{ij}^{(m)}, \quad (6)$$

where $w_{ij}^{(m+1)}$ is the new connection weights, $w_{ij}^{(m)}$ is the previous weights, $\Delta w_{ij}^{(m)}$ is the estimated maximum gradient according to backpropagation algorithm and η is the learning rate. Learning with too large a learning rate can cause the algorithm to diverge and can also mean important minima are missed, but too small a learning rate results in slow learning and the network is prone to being trapped in local minima. The standard backpropagation algorithm employs a constant learning rate, which is empirically determined. Experiments with the STI problem showed a constant learning rate, when the value is suitable, converges steadily, but is very slow. A modified learning rule, with variable learning rate, is used to speed up the training. The training phase is divided into three periods: output clamping, intermediate training and fine tuning. In the early stage of training, the ANN tends to gradually converge to output values within the [0 1] interval. In the intermediate stage of training, the ANN fits the details from training set. In the final stage, the ANN fine tunes itself to give the best performance for generalization. Different learning rates are used in first and last period. When the output of the ANN is beyond the [0 1] interval, larger steps are used to quickly drive the ANN to produce outputs in the [0,1] region. In the final fine tuning period, smaller steps are used. This modified learning thus can be expressed as

$$\eta = \begin{cases} (1.2-1.3) \eta & \text{when output} \notin [0,1], \\ \eta & \text{others,} \\ (0.3-0.5) \eta & \text{when error reduction} \rightarrow 0. \end{cases} \quad (7)$$

Such a modification to standard training method is found effective in speeding up the early phase of training and is numerically robust. Using a variable learning rate is common practice in neural network applications²⁰ and has been shown to speed training for a wide range of applications. Figure 4 shows the typical error reduction found in the early period of training. A rapid drop of ensemble errors can normally be achieved due to the enhanced error back propagation when output is beyond the interval of 0 to 1. Incidentally, If initial

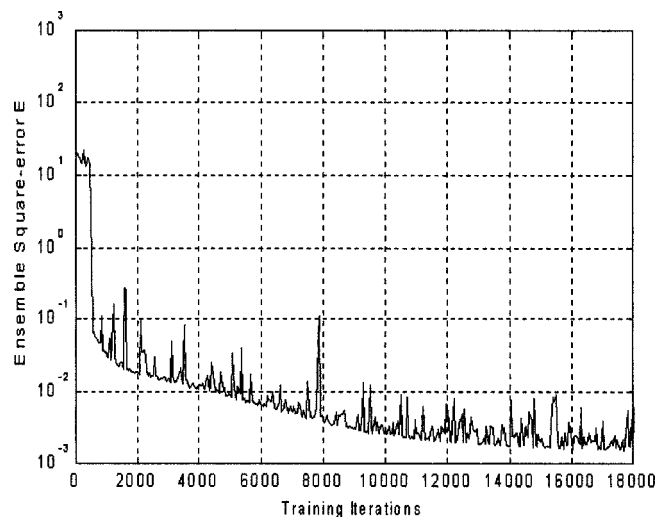


FIG. 4. Typical error reduction found in early phases of training. Each training iteration comprises a block of 50 examples. y axis is the sum of squared errors tested by randomly chosen blocks comprising 50 examples.

convergence is slow, the weights of the neural network are randomized and the training process restarted.

The termination criteria for neural networks are crucial. The requirement for generalization, i.e., that the network is successful with data not seen in training, indicates that seeking global minimum mean square error for the training set is neither necessary, nor the best solution. Tuning the ANNs to overly fit the training data set does not necessarily guarantee that the best generalization results. Therefore, the training is continuously monitored by the examination of the neural networks' response to a small set of validation data. The training stops when, either the predefined prediction accuracy is achieved or a sign of over-fitting occurs (consistently increasing mean-square-error), whichever the first. It is known that a 0.02 standard deviation is typical when using the STI standard method,⁶ and so this is used as a termination criterion.

There are three other important aspects to specify in the ANNs: size, structures, and preprocessors. In theory, the more neurons and hidden layers a neural network possesses the better its function mapping capability will be. Nevertheless, excessive number of neurons and too many hidden layers cause problems in practice,¹⁵ as back propagation tends to be slow and learning becomes inefficient and extremely time consuming. As a practical rule of thumb, small sized ANNs are preferable if they suffice. The suitable size for the network is determined through trial and error. Training tends to be more efficient when input information is coded in a suitable format for the ANN.^{15,17} As a common practice, a preprocessor, which functions to perform data reduction and signal conditioning is used to bridge the real world signals and the input layer of the ANN. In this application, an envelope spectrum estimator is the core of the preprocessor.

C. Preprocessor

There are three key issues to be considered in designing the preprocessor for STI extraction. First, useful information should be retained while redundant information should be reduced—only about 0.01% of the original data in the speech

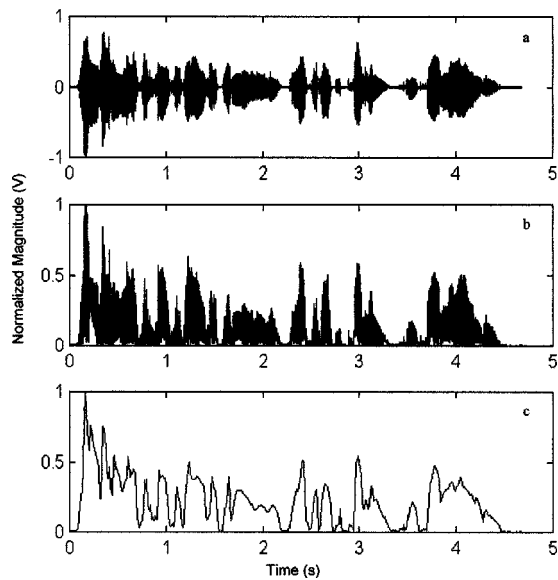


FIG. 5. Envelopes detected by Hilbert transform. Top is speech signal, middle is envelope, and bottom is 80 Hz low pass filtered envelope. Graphs normalized to maximum magnitude.

can be sent to the neural network otherwise the number of input neurons becomes too large. Second, existing knowledge should be used, otherwise the ANN has to model functions that could have been more efficiently processed using traditional means—this minimizes the load on the ANN. Third, the input vectors to the ANN should be normalized to between -1 and $+1$ as this speeds the training. The envelope spectrum estimator is naturally considered as the preprocessor. It can significantly reduce the amount of data in long speech excerpts, but maintains decisive information for STI values.

The best estimate of the long term envelope spectrum needs to be obtained. Traditional methods applied repetitions of the speech to spectral analyzers,^{9,13} but nowadays more sophisticated algorithms are available. A Hilbert transform is used as it gives a better estimation of the time signal envelopes.²¹ Accordingly, the envelope $ev(t)$ is

$$ev(t) = \sqrt{s^2(t) + s_h^2(t)}, \quad (8)$$

where $s_h(t)$ is the Hilbert transform of speech signal $s(t)$ defined by

$$s_h(t) = H[s(t)] \equiv \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(t-t')}{t'} dt'. \quad (9)$$

Figure 5 demonstrates a detected envelope using this method. Such a detector gives an unambiguous definition of signal envelopes; it is superior to techniques where window width and overlap have to be chosen.²¹ Another attractive feature of the Hilbert transform based detector is that it can be performed via a fast Fourier transform (FFT) enabling a quicker implementation.

Figure 6 is a block diagram of the proposed preprocessor. One minute of speech is used. Octave bandpass filters are inserted when needed. Only low frequency contents found in the envelope spectra are of interest. Consequently, envelope signals are low pass filtered by a fourth order But-

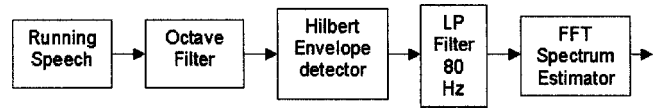


FIG. 6. Block diagram of envelope spectrum preprocessor.

terworth filter with a -3 dB point of 75 Hz and the signal is then resampled at 160 Hz. The decimated envelope signals are then passed onto the power spectrum estimator to obtain the envelope spectrum. This is done using an overlapped Hanning window technique, and the windows are typically 2.5 s long. Envelope spectra are normalized to the average energy of speech signal excerpts (when a sine wave having a RMS value equal to the mean intensity of the speech signal passes through, 0 dB is obtained). The normalization has important practical and physical meaning.

- (1) Ensuring envelope spectra are not input signal level dependent.
- (2) Expressing the frequency components of speech envelope with respect to the total energy.
- (3) Including both speech envelope fluctuation and interference noise levels in the envelope spectra.¹³

Envelope spectra are frequency domain sampled and fed into the input layer of the neural networks. Not surprisingly, it is found that the window width and FFT length of the spectrum estimator has a significant impact on obtaining accurate results. According to the standard STI method, 14 data points at central frequencies of 1/3-octave bands from 0.63 to 12.5 Hz are used. To achieve these sample frequencies requires some zero padding of the time windows. This frequency domain sampling is found adequate in training ANN on single speech excerpt in octave bands. Using a digital implementation of the envelope spectrum estimator enables higher resolution frequency sampling to be achieved, giving a more detailed representation of the envelope spectrum. This is found particularly useful in training ANNs on broadband unfiltered speech and multiple speech examples as discussed in Secs. IVD and IVE.

STI values are determined from MTFs in octave bands of speech interest. However, as speech signals have limited bandwidth, very little energy is found in frequency band above 6.3 kHz. Figure 7 shows the spectrum of a typical anechoic speech excerpt. Because of this signal to noise ratio problem, the 6.3 kHz band is used instead of the problematic 8 kHz band.

D. Training and validation data sets

Artificial neural networks using supervised training need to learn from a large number of example-teacher pairs. A simple stochastic model for impulse response synthesis—multiplying white noise by an exponential decay function—was previous used by Schroeder to investigate MTF measurement methods¹⁴ and was said to be realistic in the late part of reverberation. However, such simple stochastic model does not give frequency dependency and is inaccurate in its description of early reflections. STI by definition is a frequency dependent parameter. An improved stochastic model

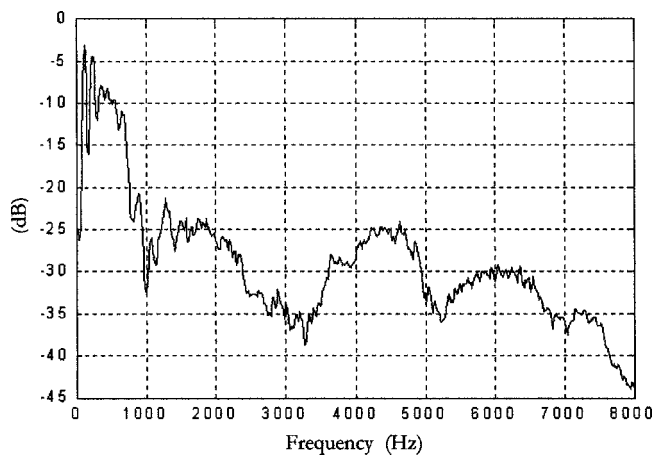


FIG. 7. An example of a speech spectrum.

was developed for this study which incorporates various possible early reflection patterns, frequency dependence and realistic reflection density. Figure 8 shows an example of an impulse response synthesized using the new stochastic model.¹¹

The stochastic impulse response synthesiser is coded such that it randomly generates all impulse response patterns while statistically it is controlled by governing diffuse field physical laws. Although governed by the diffuse field laws, the stochastic nature of the generator meant that distinctly nonlinear decays could be generated. When it is run for sufficiently large number of times, it hypothetically generates a superset of impulse responses found in reality. A small number of real impulse responses were also used, not enough to properly validate the systems ability for actual measurements, but some reassurance that the simulations are realistic.

The success of ANN methods is evaluated using validation tests. The validation tests use data not seen in training to test for generalization as is standard practice in ANN research. The data sets are split into two halves, the first half is used to train the ANNs, while the second half is used to validate the trained ANNs. Rigorous validation is achieved by ensuring that cases in validation sets have never been used in training. In this study, six untrained narrators were used to read excerpts from three different text materials. In each different text material two different excerpts were taken. The text extracts were contrasting samples, ranging

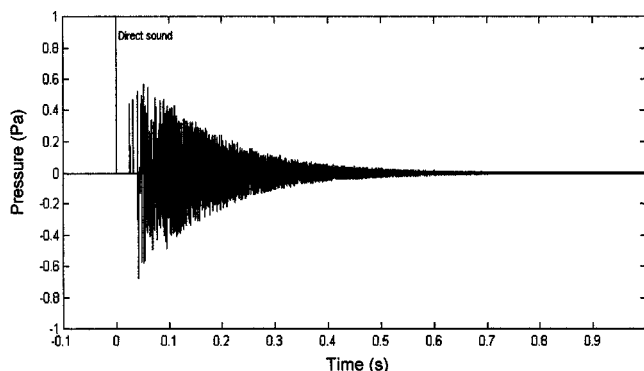


FIG. 8. An example of simulated impulse response.

from a descriptive passage from a classic 19th century novel to a conversation passage from a late 20th century popular novel. The teachers for training the ANNs are the true STI values of these speech examples. STI was calculated from impulse responses as follows. First, the MTFs are calculated using Eq. (1). Then the standard procedure for obtaining STI from the MTF is followed.^{3,6,7}

- (1) Calculation of mean apparent S/N ratio $\overline{(S/N)}_{app,F}$ in each octave band

$$\overline{(S/N)}_{app,F} = \frac{1}{14} \sum_{F=0.63}^{12.5} \max \left[\min \left[10 \log \left(\frac{MTF(F)}{1 - MTF(F)} \right), 15 \right], -15 \right]. \quad (10)$$

- (2) Calculation of overall mean apparent S/N by weighting the $\overline{(S/N)}_{app,F}$ of seven octave bands, and converting to an index ranging from 0 to 1:

$$\overline{(S/N)}_{app} = \frac{15 + \sum w_k \overline{(S/N)}_{app,F}}{30}. \quad (11)$$

The values of w_k are given in Refs. 3 and 6. In this paper, examples cover reverberation times from 0 to 5 s, signal to noise ratios from 0 dB to noise free. Added noise is white.

IV. APPLICATION OF ANNS TO STI EXTRACTION

A. Training on impulse responses

The capability of ANNs to extract STI values from impulse responses is first explored. The experiment intends to identify (i) if the ANN can generalize impulse responses and (ii) if the ANN can perform the nonlinear calculation needed to obtain STI values. The 14 000 simulated impulse responses are used. These are octave band filtered and used as input signals. The MTF values are obtained via Eq. (1). The 14 MTF values in the 1/3 octave band used in the standard STI method are fed into an artificial neural network. This network has the following neurons in the input, hidden and output layers: 14–10–8–1. Typically, 50 000 iterations (each iteration presents a block of 280 examples) are required for satisfactory results, but this may vary with different initial weights and learning rates. When training is completed, the network is validated. To demonstrate how the network generalizes to different impulse responses, Fig. 9 shows the standard deviation found over all validation tests. It is found that relatively large standard deviations are associated with higher level of noises because the noise interference never repeats. Even so, very low errors can be obtained in this case. This is not surprising, since neural networks can map almost any complicated function. This also shows that the network can generalize from impulse responses seen in the training phase, to ones not seen before in the validation phase. Since the precise relation between impulse response and MTF is known, the use of ANN here does not surpass traditional calculation in terms of accuracy. On the other hand, the ANN

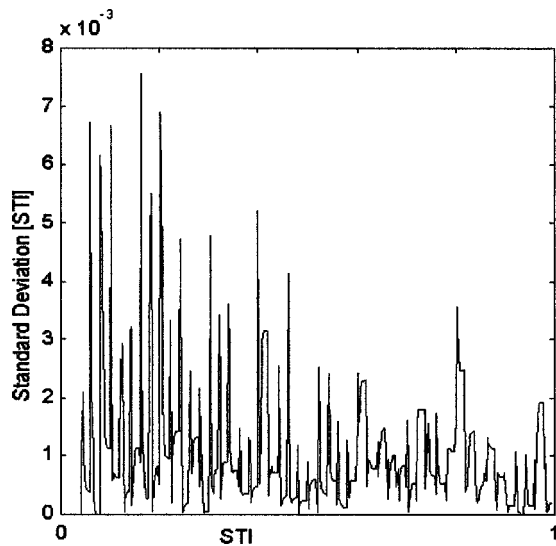


FIG. 9. Standard deviation found over validation tests.

method to map impulse responses to STIs would provide a useful alternative neural computing approach for implementation on very simple hardware.

B. Training on octave band STI

Octave band STI is defined as $\overline{((S/N)_{app,F} + 15)/30}$, where $(S/N)_{app,F}$ is defined in Eq. (10). The training phase and retrieval phases are as illustrated in Fig. 1. Again, envelope spectra values at 14 designated modulation frequency points are extracted and fed into the input layer of the neural networks. It is empirically found that a 14–20–8–1 network performs well for all octave bands. It is found that very high accuracy is possible. No signs of over-training were found before the maximum error in validation tests dropped to below 0.01 STI. (The reason why a more strict termination criterion was used here is because the extracted octave-band STI values will eventually be used to calculate full STI.) This method is found robust to different speakers, text types and mode of reading, as long as they are individually trained on those particular speech excerpts, i.e., each speaker or text requires a different network. This demonstrates that the ANN can be used to memorize speech spectra and compensate the error in measurement when natural running speech is used as excitation.

C. Full STI

As the full band STI is a linear combination of STI values in octave bands according to Eq. (12), this can be implemented with a fixed network structure as shown in Fig. 10. Seven neural networks representing corresponding octave bands are individually training as described above, they then form a bank of trained neural networks. The outputs of these are processed by the additional linear layer. As the weights for octave bands are known, no further training for this weighting layer is required.

The maximum prediction error found is 0.018 STI as shown in Fig. 11. Unfortunately, accurate error values for Steeneken and Houtgast's running speech method¹⁰ are not

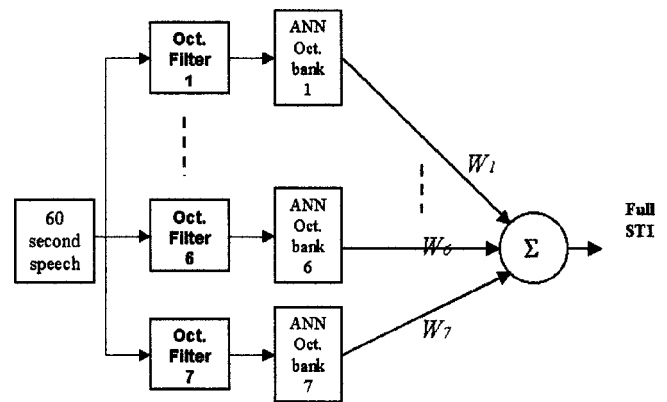


FIG. 10. How the ANNs trained to give octave band STI are combined to give full STI values.

available from their brief conference paper, but it is known that the correlation coefficient between their running speech method and the true STI was 0.971. The correlation coefficient between the ANN estimated and true STI is 0.9999, considerably higher.

D. Training on unfiltered speech

Artificial neural networks together with the preprocessor were trained on unfiltered speech to extract STI values. The motivation here is to form a more compact structure which would then be more efficient to implement in instrumentation. In this case, octave band filters are not inserted. Speech was fed into ANN via the spectra estimator directly. ANNs have been successfully used to enhance accuracy of spectrum analysis²² and this philosophy is followed. Due to the squaring operation in forming the envelope, high frequency components of the received speech appear also at low frequencies, these additional components being generated by the cross terms in the squaring operation. This gives the ANN access to additional information to base its parameter estimation on.

It is therefore sought to train the ANN to learn from unfiltered speech and automatically associate contributions from different octave bands to give reasonably accurate STI estimations. Envelope spectrum values for the 14 1/3-octave frequency bands used in the standard STI method (the energy being summed over the 1/3-octave band), are used as input vector for the ANN. Full STI values were used as teachers. Gradual convergence was shown in training process, however, ensemble errors were not reduced to a satisfactorily low value even after training for a long time. The test result showed a maximum prediction error of 0.07.

It has been empirically identified that the traditional STI envelope spectrum analysis^{1,9,12} is inadequate for extracting STI from unfiltered speech. To decrypt the intermodulated information, a higher resolution envelope spectrum estimator is needed. Consequently, the Hilbert transform detected envelope, low pass filtered at a cutoff frequency of 80 Hz, is used to allow more information to be used. High resolution power spectrum estimator gives envelope spectra at a 0.3125 Hz frequency step giving 40 linearly sampled envelope spectra from 0.3125 Hz to 12.18 Hz. The 40 data point are sub-

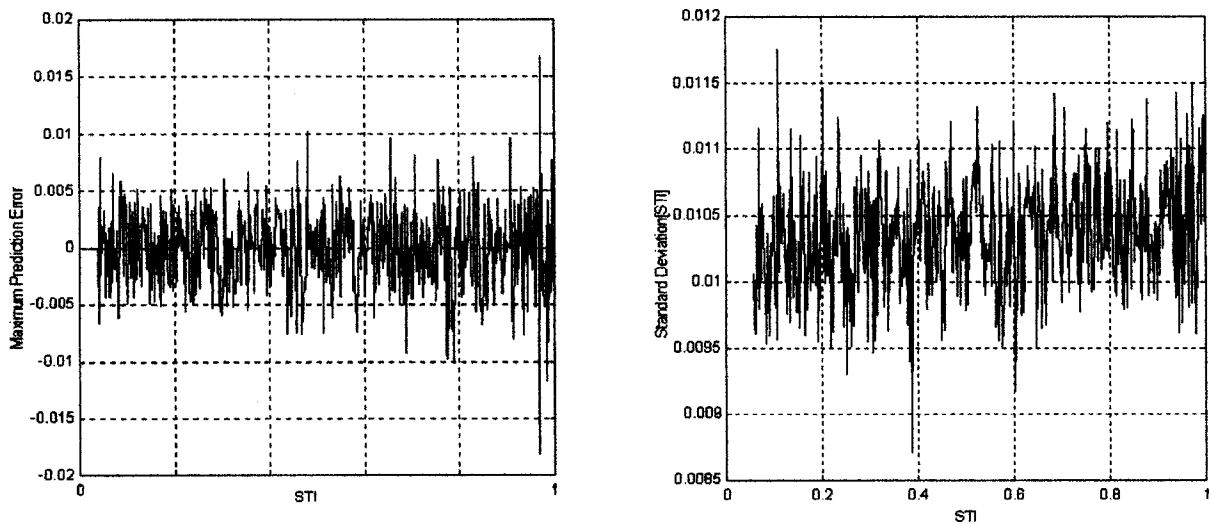


FIG. 11. Maximum prediction errors and standard deviations found in validation tests. STI obtained using the structure shown in Fig. 10.

sequently used to form the input vector for the ANN. A network size of 40–20–10–1 was found suitable. Validation results are shown in Fig. 12. The correlation coefficient between the ANN estimated and true STI is 0.999 97 and the maximum prediction error is 0.0197 STI.

E. Multiple speech excerpts: source independence

So far, a set of ANN models have developed to accurately extract STIs from received speech. The systems work with a specific speech excitation, as the statistical feature of that particular speech is learned and imbedded in the ANN. This means prerecorded speech signals have to be used. As one step further, the feasibility to learn from different speech excitations and generalize to arbitrary speech (source independent) extraction of STI is explored. This would enable the instrument to be much more flexible, with the potential to work with live speech.

The very low frequency envelope spectrum of anechoic running speech is known to be a generally stable.¹³ However, speech is very complicated stochastic process and the envelope spectra are not sufficiently stable to be regarded as constant for STI extraction. Figure 13 shows an over-plot of 18

envelope spectra of anechoic speech signals read by six untrained native English narrators. A maximum difference of approximately 7 dB is found.

The problem with arbitrary speech is that the attack and decay of the anechoic speech mixes with the reverberance in the room. To take a simple example, a word pronounced with a long decay (e.g., “bus”) in a dry room, can have the same envelope as a short decay word (e.g., “stop”) in a reverberant space. While using a long speech extract can help average out random variations, consistent differences in pronunciation will affect the envelope spectrum. These differences in the envelope spectrum caused by pronunciation can be indistinguishable from the changes due to reverberance in the room. For this reason, contradictions in the data set are seen, and the neural network fails to properly converge because it is asked to map similar envelope spectra to different STI values.

To deal with these contradictions, additional information must be fed to the ANN. One possibility is to feed additional information from frequency ranges not previously used (>25 Hz). As the envelope spectra are normalized, a speech having lower level spectrum in certain frequencies must result in

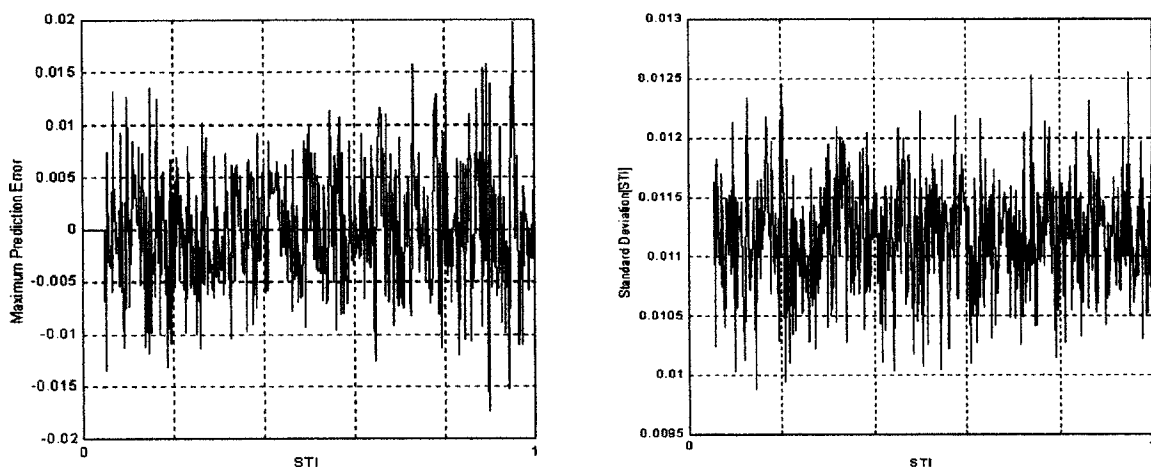


FIG. 12. Maximum prediction errors and standard deviations found in validation tests. STI obtained using one neural network from unfiltered speech.

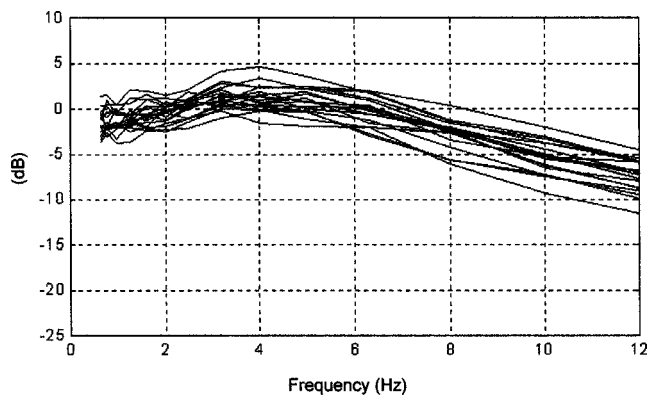


FIG. 13. Over plot of envelope spectra of 18 speech excerpts in 1 kHz band. Envelopes obtained using a digital implementation of a traditional envelope spectrum analyzer.

higher level spectrum at other frequencies as there is no leakage. In addition, it is assumed that in the vast quantity of data being filtered out by the preprocessor, there is information concerning how the speaker pronounces words, and this information is needed for the neural network to resolve the contradictions in the data set.

The envelope spectra are taken up to 80 Hz; these are estimated using Welch's average periodogram method.²³ Frequency contents are sampled at a 0.5 Hz step up to 80 Hz, providing 160 inputs for the ANN. The ANN has a 160–40–20–1 architecture. The 18 different anechoic speech examples and three different texts read by six narrators are used in the training.

Figure 14 illustrates the errors found with the validation tests. The maximum prediction error for STI found in is 0.13, and the correlation coefficient between actual and predicted STI is 0.9948. Better accuracy can be obtained by averaging over several different speech excerpts. When averaging the

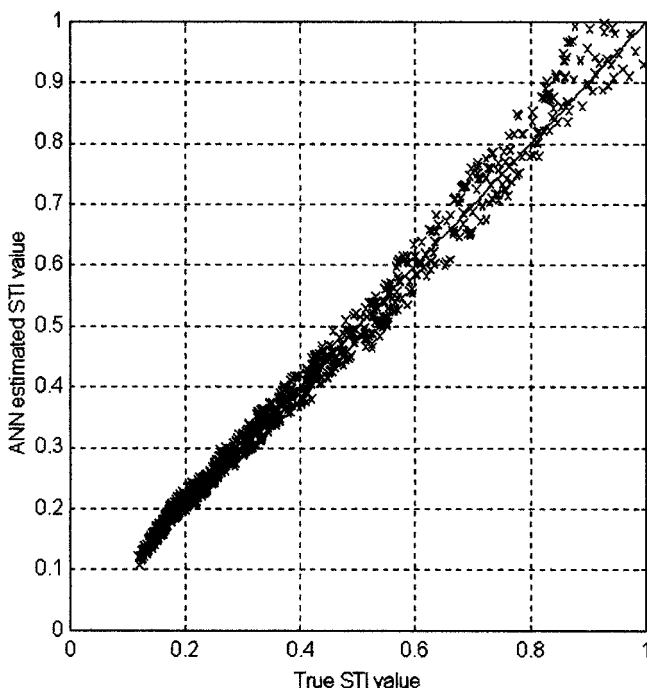


FIG. 14. Comparison of ANN predicted (worst cases) and actual STI over 18 different speech extracts.

estimated STI over three different speech excerpts read by different narrators, prediction errors can normally be reduced to less than 0.1. Consequently, accurate enough extraction from arbitrary speech with only output speech knowledge has not been achieved.

V. DISCUSSIONS AND CONCLUSIONS

A neural network method to improve the accuracy and repeatability of STI measurements with running speech is proposed and validated via simulations. This method can significantly improve the accuracy of STI measured with natural running speech, hence facilitating measurement in occupied conditions. The proposed neural network method works with both received broadband and octave band speech signals, providing an accuracy comparable to measurements made using artificial test signals, typically a standard deviation of less than 0.02, when a one-net-one-speech excerpt case is considered.

Source independent extraction of STI from speech, was explored. It seems that the proposed ANN method has a certain capability to learn from examples and adapt to different speakers and texts. The actual STI and ANN estimation show reasonable agreement when testing with speech excerpts not previously seen by the ANN. Further investigations are needed to fully develop such a technique to gain sufficient accuracy for a practical measurement system.

Only a few real impulse responses have been used in training and validations, to fully validate this method and evaluate its use, more on-site validations will be needed. Nevertheless, this method proposed and validated here mainly with simulations, provides a promising avenue towards accurately measuring STI from natural occurring sound sources. The work has only considered the case of natural sound reproduction into rooms, but there is no reason why this cannot work with public address systems. To achieve this, suitable examples of sound reproduced by a public address system would have to be included in the training set.

ACKNOWLEDGMENTS

This paper was funded by the Engineering and Physical Sciences Research Council, UK, under Grant No. GR/L89280. The authors would also like to thank colleagues at Salford University who helped make the anechoic speech recordings.

¹T. Houtgast and H. J. M. Steeneken, "Evaluation of speech transmission channels by using artificial signals," *Acustica* **25**, 355–367 (1971).

²T. Houtgast and H. J. M. Steeneken, "The modulation transfer function in room acoustics as a predictor of speech intelligibility," *Acustica* **28**, 66–73 (1973).

³H. J. M. Steeneken and T. Houtgast, "A physical method for measuring speech transmission quality," *J. Acoust. Soc. Am.* **67**, 318 (1980).

⁴T. Houtgast and H. Steeneken, "Predicting speech intelligibility in rooms from the modulation transfer function. part I. General room acoustics," *Acustica* **46**, 60–72 (1980).

⁵ISO TR 4870, "Technical report: Acoustics—The construction and calibration of speech intelligibility tests," 1991.

⁶"Technical Note, Device for measuring the speech transmission index," *J. Acoust. Soc. Am.* **71**, 1612 (1982).

⁷IEC 60268-16:1998 (also BS EN 60268-16 and BS 60268-16), "Sound

- system equipment, Part 16: Objective rating of speech intelligibility by speech transmission index," 1998.
- ⁸P. W. Barnett and P. H. Scarbrough, "From word scores to RASTI and back—An experimental study of the relationship between RASTI and word scores," *Proc. Inst. Acoust. UK* **22**, 73–82 (2000).
 - ⁹H. J. M. Steeneken, J. Verhave, S. McManus, and K. Jacob, "Development of an accurate, handheld, simple-to-use meter for the prediction of speech intelligibility," *Proc. Inst. Acoust., UK* **23**, 53–59 (2001).
 - ¹⁰H. J. M. Steeneken and T. Houtgast, "The temporal envelope spectrum and its significance in room acoustics," *Proceedings of the 11th ICA*, 7, Paris, 1983, pp. 85–88.
 - ¹¹T. J. Cox, F. Li, and P. Darlington, "Extraction of room reverberation time from speech using artificial neural networks," *J. Audio Eng. Soc.* **49**, 219–230 (2001).
 - ¹²F. F. Li and T. J. Cox, "Extraction of speech transmission index from speech signals using artificial neural networks," *Proceedings of the 110th AES convention*, Amsterdam, paper 5354, 2001.
 - ¹³T. Houtgast and H. J. M. Steeneken, "Envelope spectrum and intelligibility of speech in enclosure," *IEEE–AFCRL Speech Conference*, 1972.
 - ¹⁴M. R. Schroeder, "Modulation transfer functions: definition and measurement," *Acustica* **49**, 179–182 (1981).
 - ¹⁵S. Y. Kung, *Digital Neural Network* (Prentice-Hall, New York, 1993).
 - ¹⁶S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. (Prentice-Hall New York, 1999).
 - ¹⁷G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control, Signals Syst.* **2**, 303–314 (1989).
 - ¹⁸D. E. Rumelhart, G. Hinton, and J. R. Williams, "Learning internal representations by error propagations," *Parallel Distributed Processing* (MIT Press, Cambridge, MA, 1986), Vol. 1, Chap. 8.
 - ¹⁹M. Riedmiller, "Advanced supervised learning in multi-layer perceptrons—from back propagation to adaptive algorithms," *Int. J. Comput. Standards Interfaces* **16**, 265–275 (1994), special issue on neural networks.
 - ²⁰R. A. Jacobs, "Increased rate of convergence through learning rate adaptation," *Neural Networks* **1**, 295–307 (1988).
 - ²¹H. Kuttruff, *Room Acoustics*, 4th ed. (Spon Press, 2000), p. 249.
 - ²²F. Luo and R. Unbehauen, *Applied Neural Networks for Signal Processing* (Cambridge University Press, Cambridge, MA, 1996), pp. 74–120.
 - ²³M. H. Hayes, *Statistical Digital Signal Processing and Modeling* (Wiley, New York, 1996), pp. 415–420.

Correction of ultrasonic wave aberration with a time delay and amplitude filter

Svein-Erik Måsøy,^{a)} Tonni F. Johansen, and Bjørn Angelsen

Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Trondheim, Norway

(Received 10 June 2002; revised 2 January 2003; accepted 12 January 2003)

Two-dimensional simulations with propagation through two different heterogeneous human body wall models have been performed to analyze different correction filters for ultrasonic wave aberration due to forward wave propagation. The different models each produce most of the characteristic aberration effects such as phase aberration, relatively strong amplitude aberration, and waveform deformation. Simulations of wave propagation from a point source in the focus (60 mm) of a 20 mm transducer through the body wall models were performed. Center frequency of the pulse was 2.5 MHz. Corrections of the aberrations introduced by the two body wall models were evaluated with reference to the corrections obtained with the optimal filter: a generalized frequency-dependent phase and amplitude correction filter [Angelsen, *Ultrasonic Imaging* (Emantec, Norway, 2000), Vol. II]. Two correction filters were applied, a time delay filter, and a time delay and amplitude filter. Results showed that correction with a time delay filter produced substantial reduction of the aberration in both cases. A time delay and amplitude correction filter performed even better in both cases, and gave correction close to the ideal situation (no aberration). The results also indicated that the effect of the correction was very sensitive to the accuracy of the arrival time fluctuations estimate, i.e., the time delay correction filter. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1559174]

PACS numbers: 43.58.Ry, 43.20.Fn [LLT]

I. INTRODUCTION

Wave distortion in medical ultrasound imaging originates from the inhomogeneities in the tissue, especially in the human body wall. Distortion of an ultrasonic wave, also known as aberration, leads to poorer resolution in the obtained images.

In the literature it is possible to deduce three different methods for correction of ultrasonic wave aberration.

(1) A focusing and apodization method, where arrival time and amplitude fluctuations are corrected in the focusing and apodization of the transmitted signal. In this article this method will be denoted in the general case as a correction filter, or more specifically, as a time delay correction filter alone or in combination with an amplitude correction filter. Flax and O'Donnell¹ introduced a cross-correlation method for estimation of arrival time fluctuations in the received aberrated signal. This represents the classical method for estimation of arrival time fluctuations and has been extensively used in the literature on ultrasonic wave aberration.

(2) Fink² introduced the time reversal mirror, which incorporates information about the pulse form in addition to arrival time and amplitude fluctuations. This method requires a point reflector in the insonified medium, which rarely is the case.

(3) The backpropagation method introduced by Liu and Waag³ uses the time reversal principle to obtain an optimal situation for the estimation of aberration parameters. For the case of estimating a time delay correction filter, it is not yet

completely understood if backpropagation followed by arrival time estimation works better than arrival time estimation at the transmitting array for correction of the retransmitted wave.^{4,5}

The object of the work presented in this article was to find how well a time delay correction filter, or a time delay and amplitude correction filter, could correct for ultrasonic wave aberration. There has been performed much research on time delay and amplitude correction of ultrasonic wave aberration,^{1,6-9} but all of this work has used a very simple aberration model consisting of a single pure time delay screen placed just in front of, or some distance from the transmitting/receiving array. The work done by the ultrasound group at the University of Rochester¹⁰⁻¹⁵ has given valuable insight into how the inhomogeneities in the human body wall affect aberration of ultrasonic waves. This has been crucial in the understanding of this phenomenon, and possible corrections. Their work has also shown that a single pure time delay screen model is too simple and cannot adequately model ultrasonic wave aberration.

Berkhoff and Thijssen⁵ introduced a number of successive time delay screens to model ultrasonic wave aberration, but gave no characterization of the total accumulated effect of these successive screens. In the present article, the concept of successive time delay screens has been further developed as a model of the human body wall, although this has been criticized as not an appropriate method to model aberration introduced by propagation through the human body wall.¹⁵ The main reason to create such a body wall model was to simulate ultrasonic wave aberration with a relatively simple and nonexpensive method. Such a method can be used to

^{a)}Electronic mail: svein-erik.masoy@medisin.ntnu.no

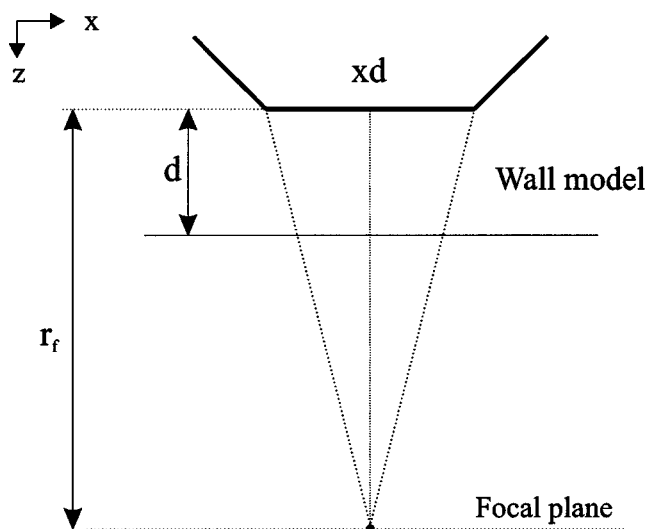


FIG. 1. Simulation setup.

evaluate various techniques of estimation and correction of wave aberration. This article shows that if care is used in designing such a wall model, it produces most of the reported aberration phenomena introduced by the human body wall.

A correction method using a generalized frequency-dependent phase and amplitude correction filter¹⁶ was used as a reference for the best possible aberration correction. Such a filter comprises spatial and frequency dependency of both amplitude and phase fluctuations at the receiving array. Approximations to the generalized frequency-dependent correction filter with a time delay filter alone, or in combination with an amplitude filter, were then evaluated from simulations of point sources placed in the focus of a focusing array. Simulations with point sources in the focus of an array provide the optimal situation for estimation of wave aberration. Six different methods were evaluated for the estimation of the time delay correction filter. Absorption effects were not included in the simulations.

II. THEORY AND METHOD

A. Simulation setup

The simulation setup is illustrated in Fig. 1, which depicts a simulation setup of a point source placed in the focus r_f of a focusing array. A pulse was propagated from the point source through a body wall model of thickness d to the receiving array. A correction filter was estimated from the received signal and applied on transmit of a focused pulse. This pulse was propagated back to the focal plane for evaluation of the applied correction filter.

B. Body wall modeling

The human body wall was modeled with different sets of several time delay screens, in order to simulate different cases of aberration. A time delay screen only introduces a shift in the phase of the ultrasonic signal with a pure time delay. The time delay screen was defined only in a plane (Fig. 2), and thus had zero thickness. The propagation

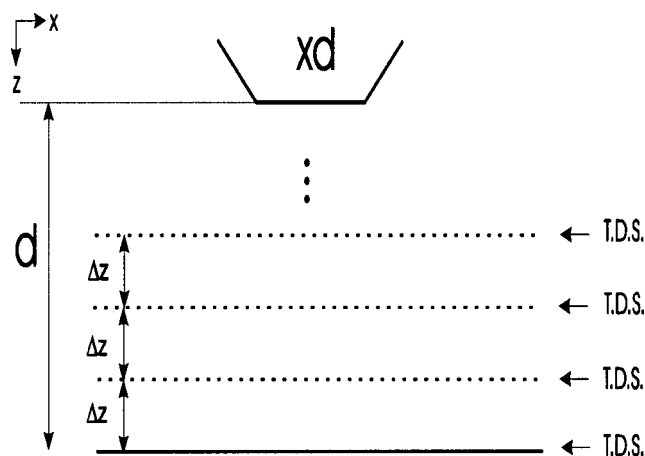


FIG. 2. Body wall model with time delay screens.

through the body wall model was performed by homogeneous propagation with an angular spectrum operator (described in Sec. II C) between each time delay screen.

In order to create different cases of aberration, eight sequences of random numbers with a standard normal distribution were generated. These sequences were then low-pass filtered to obtain reported correlation lengths for the time delays produced by the human abdominal wall.¹⁰

Four different configurations containing two, four, six, and eight sequences were selected in order to generate different aberration conditions. Various aberration situations could now easily be simulated by multiplying each configuration with an individual time delay amplitude (given in seconds), thus turning the configuration into a set of time delay screens. Note that in this article, all the time delay screens in a specific configuration were multiplied with the same time delay amplitude. The time delay screens in a configuration were uniformly spaced over a depth range d with a distance $\Delta z = d/(\text{number of screens})$ between them. The first time delay screen was placed a distance Δz from the array, and the last was placed at the distance d from the array (see Fig. 2). To summarize: a given model of a body wall consisted of one of the above-mentioned configurations, multiplied with a defined time delay amplitude given in seconds, and distributed uniformly in space over a depth d .

Two categories of human body wall models were created, representing cases of weak and strong aberration. These two categories were not rigorously defined, but differently created body wall models were placed into these categories after several selection criteria. These criteria were body wall statistics (correlation lengths and rms values of arrival time and energy level fluctuations), the level of waveform deformation (calculated by a waveform similarity factor), visual inspection of received arrival time and amplitude fluctuations, visual inspection of received waveforms, and visual inspection of aberrated beam profiles in the focal plane.

A first crude selection criterion was that the maximum peak to peak difference in the arrival time fluctuations should be less than half the period of the transmitted pulse for the weak aberration case. For the strong aberration case this value should be higher than half the period of the transmitted pulse. Based on this crude selection rule, four different body

wall models with two, four, six, and eight time delay screens, respectively, were generated in both the weak and strong aberration categories. Both the weak and strong aberration categories thus contained all four sequence configurations, multiplied with a different time delay amplitude.

One body wall model was then chosen from each of the two categories. The reason for this was to find the body wall model that produced the “worst” aberration in each category. In this procedure the additional selection criteria were applied subjectively. The most important measures of aberration from these criteria were side lobe level in aberrated beam profiles, width of main lobe in aberrated beam profiles, shape distortion and shape deformation of received waveforms, and finally comparison of body wall statistics with measured statistics for human abdominal wall.¹⁰

The two selected body wall models were denoted w6 and s6, w6 representing the weak aberration category and s6 the strong aberration category. The number six indicating that both body wall models contained six time delay screens. The s6 body wall model was meant to represent a case of severe ultrasonic wave aberration. A main point for the s6 wall was that in addition to all the other selection criteria, the received waveform should be severely shape deformed and the body wall statistics should be very “tough.” This means that the rms values of the arrival time and energy level fluctuations should be well above the mean of those reported in Ref. 10, and the correlation lengths in the low region of the reported values.

To summarize: this elaborated process was carried out in order to secure that the two final selected body wall models, w6 and s6, really generated aberration comparable to what is observed in ultrasonic imaging. These were also chosen because they represented the worst aberration in their respective categories, and note that simulation results with correction filters from only these two selected body wall models are presented in this article.

C. Linear angular spectrum propagation operator

The propagation operator used in the simulation program was obtained by the solution of the linear wave equation given as

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0, \quad p = p(x, z, t), \quad (1)$$

where p is the pressure. The spatial coordinate x defines the array direction and z the propagation direction (see Fig. 1). The Fourier transform with respect to time t and the spatial coordinate x gives the equation of the form

$$\frac{\partial^2 P_{tx}}{\partial z^2} + k_z^2 P_{tx} = 0, \quad k_z = \sqrt{\left(\frac{\omega}{c}\right)^2 - k_x^2}, \quad (2)$$

where the capital P represents the Fourier transform of the pressure, and the indices indicate the coordinate that is transformed. The wave number in the propagation direction is defined as k_z , and k_x represents the wave number in the x direction. The solution of this equation, with z as the positive propagation direction, can be expressed as

$$P_{tx}(k_x, z, \omega) = P_{tx}(k_x, z_0, \omega) e^{-jk_z(z-z_0)}, \quad (3)$$

where $P_{tx}(k_x, z_0, \omega)$ is the pressure given at $z = z_0$, which represents the array surface in the simulations. The propagation operator is given as $e^{-jk_z(z-z_0)}$, where z is the propagation distance. Note that the propagation operator propagates a two-dimensional plane defined by the spatial coordinate x and the time coordinate t , in the propagation direction z . It should also be noted that k_x marks the limit for propagation in the propagation direction. If $\omega/c > k_x$, the operator is complex and represents propagation in the direction of z . If $\omega/c < k_x$ the operator becomes real valued and represents a damping of the wave, but no propagation. This phenomenon is known as evanescent waves.

D. Correction filters and the generalized frequency-dependent screen

The solution of the wave equation with a point impulse source for an inhomogeneous medium can be modeled in the frequency domain as¹⁶

$$G_f(\mathbf{r}, \mathbf{r}_s; \omega) = S(\mathbf{r}, \mathbf{r}_s; \omega) G_h(\mathbf{r} - \mathbf{r}_s; \omega), \quad (4)$$

where G_f is the free-space Green's function in the inhomogeneous medium as a function of the field coordinate \mathbf{r} , the source location \mathbf{r}_s , and the angular frequency ω . The capital letters indicate the temporal Fourier transform of the functions. The free-space Green's function for the homogeneous medium G_h is given as

$$G_h(\mathbf{r} - \mathbf{r}_s; \omega) = \frac{e^{-ik|\mathbf{r} - \mathbf{r}_s|}}{4\pi|\mathbf{r} - \mathbf{r}_s|}, \quad (5)$$

where k is the wave number. $S(\mathbf{r}, \mathbf{r}_s; \omega)$ is hence the frequency response of a filter that produces the distortion of the ultrasonic wave as it propagates through the inhomogeneous medium. Here this filter is referred to as the generalized frequency-dependent screen, and is composed of a frequency-dependent amplitude and phase screen defined as¹⁶

$$S(\mathbf{r}, \mathbf{r}_s; \omega) = A_s(\mathbf{r}, \mathbf{r}_s; \omega) e^{i\Theta_s(\mathbf{r}, \mathbf{r}_s; \omega)}, \quad (6)$$

where A_s and Θ_s are the frequency-dependent amplitude and phase screen, respectively. The indices s indicate that the phase and amplitude screens are dependent of the source position.

The generalized frequency-dependent screen can be used for correction of the wave distortion from a body wall as a time reversal mirror.² Assume that a signal from a point impulse source at \mathbf{r}_s is measured over an array with limited aperture. Because of the reciprocity of the wave equation, re-transmitting the time reversed signal received at each point on the array will reproduce the diffraction limited impulse field at \mathbf{r}_s . Time reversal of a real signal gives a complex conjugation of the signal's temporal Fourier transform. This means that filtering a transmitted signal from an array on each channel with a filter $S^*(\mathbf{r}_a, \mathbf{r}_s; \omega)$, \mathbf{r}_a denotes the array coordinate, is identical to time reversal, and will reproduce the diffraction limited impulse field at \mathbf{r}_s . The complex

conjugate of the generalized frequency-dependent screen thus represents the ideal correction filter for the wave distortion.¹⁶

In the present paper the complex conjugate of the generalized frequency-dependent screen was used as an aberration correction filter along with two approximations of this defined as

$$\begin{aligned} S_1(\mathbf{r}_a, \mathbf{r}_s) &= e^{-i\omega\tau(\mathbf{r}_a, \mathbf{r}_s)}, \\ S_2(\mathbf{r}_a, \mathbf{r}_s) &= A(\mathbf{r}_a, \mathbf{r}_s) e^{-i\omega\tau(\mathbf{r}_a, \mathbf{r}_s)}, \end{aligned} \quad (7)$$

where S_1 is a pure time delay correction filter, and S_2 is a time delay and amplitude correction filter.

III. SIMULATION PARAMETERS AND DATA PROCESSING

The simulation area was 10.24 cm in the x direction, with a grid size of 0.1 mm. The sampling frequency was 62.4 MHz providing a time window of 16.4 μ s. Center frequency of the pulse was 2.5 MHz with a -6 dB bandwidth of 1.6 MHz. The pulse was filtered on transmit with a 4 MHz band-pass filter centered around the center frequency of the pulse. An array size of 20 mm with a pitch of 0.1 mm was chosen. The focal depth of the array was set to 60 mm (distance \mathbf{r}_f in Fig. 1). Water at 37 °C was used as the propagation medium. On receive, all signals were corrected for geometric focusing for the homogeneous medium before processing of results.

The generalized frequency-dependent screen was estimated by filtering the received signal U with a Wiener-type filter H_W in the frequency domain given by

$$S(\mathbf{r}_a, \mathbf{r}_f; \omega) = H_W(\mathbf{r}_a, \mathbf{r}_f; \omega) U(\mathbf{r}_a, \mathbf{r}_f; \omega), \quad (8)$$

where \mathbf{r}_a represents the array coordinate and \mathbf{r}_f the focus of the array (see Fig. 1). The transfer function for the filter was

$$H_W(\mathbf{r}_a, \mathbf{r}_f; \omega) = \frac{U_t^*(\mathbf{r}_a, \mathbf{r}_f; \omega)}{|U_t(\mathbf{r}_a, \mathbf{r}_f; \omega)|^2 + \max(|U_t(\mathbf{r}_a, \mathbf{r}_f; \omega)|^2)/\text{SN}}, \quad (9)$$

where U_t is the transmitted frequency-dependent signal. The factor SN represents a signal-to-noise-ratio parameter in the filter.

The time delay correction filter approximation of the generalized frequency-dependent screen, defined in Sec. II D, is identical to the arrival time fluctuations in the received signal, after correction of geometric focusing for the homogeneous medium. Six different methods were used for the estimation of the arrival time fluctuations. Three of the techniques calculated the angle of the cross-correlation function of the complex envelope at zero lag in the received signal.¹⁷ The difference between them being in the choice of a reference for the calculation of the cross correlation. The different methods are labeled (a), (b), (c), (d), (e), (f), and this notation will also be used in the presentation of results later in the article.

(a) *Phase front tracking*: A phase front tracking algorithm was developed as an arrival time fluctuation estimator. In essence the phase front tracking algorithm determined the peak of the first period of the received waveforms. The track-

ing was then performed on each channel, and the time delay between two channels was defined as the difference in time between the two peaks of the respective channels.

(b) *Neighbor cross-correlation*: The cross-correlation technique (as described earlier) which calculated the cross-correlation of the complex envelopes, at zero lag, between neighboring channels in the receiving array. The total arrival time fluctuations were calculated as the cumulative sum of the time delays between each neighbor channel, starting from the first channel in the array.

(c) *Beamsum cross-correlation*: The cross-correlation technique which used a reference signal and calculated the cross-correlation of all the channels at zero lag in the receiving array, with respect to the reference signal. The reference signal was the beamsum of the received signal. The beamsum was defined as the received signal summed over all channels in the array direction, resulting in a time-dependent signal only.

(d) *Reference channel cross-correlation*: Same as (c), but the reference signal was now chosen as the signal at channel 1 in the receiving array.

(e) *Frequency mean*: Arrival time fluctuations across the receiving array were determined by taking the Fourier transform in time of the received signal on each channel. The time delay on each channel of the array was calculated as the arithmetic mean of the phase, of the now frequency-dependent signal, over a band of frequencies ranging from 2 to 3 MHz and divided by the center frequency of the transmitted pulse (2.5 MHz). This band was chosen empirically as it produced satisfactory estimates in the weak aberration case.

(f) *Ray delays*: The last estimator was used as a comparative measure of the performance of the latter described estimators. Since the time delay of each time delay screen was known, arrival time fluctuations were calculated by summing the time delay along a straight line from the focus to a channel on the receiving array.

Amplitude variations across the receiving array were determined in the same fashion as the time delay estimator (e). The received waveform was Fourier transformed in time. The amplitude on each channel of the array was calculated as the arithmetic mean of the frequency-dependent amplitudes over the same band of frequencies.

For the characterization of the different body wall models, leading to the selection of the two generating the most severe aberration, rms values of arrival time and energy level fluctuations were determined along with correlation lengths of these. A linear fit was subtracted from the arrival time fluctuations before the calculation of the correlation lengths and rms values, in order to remove direction steering of the beam by the body wall. The energy level fluctuations were simply defined as the square of the amplitude variations across the receiving array. Correlation lengths were defined as the full width between the half maximum points of the autocorrelation function.

In addition to these measures a waveform similarity factor was determined in order to quantify the extent of waveform deformation in the received signal. Wave form deformation severely complicates time delay estimation due to

low correlation in the received signal. Two similar versions of such a waveform similarity factor^{3,18} have been proposed independently. In this article the approach of Liu and Waag³ was chosen. The similarity of N received waveforms $u_1(t), u_2(t), \dots, u_N(t)$ is defined to be maximum if

$$u_i(t) = a_i u(t - \tau_i), \quad i = 1, \dots, N, \quad (10)$$

where N is the number of elements on the receiving array. From the definition of Eq. (10), $u_i(t + \tau_i) = a_i u(t)$ must also be true, so that

$$\sum_{i=1}^N u_i(t + \tau_i) = \sum_{i=1}^N a_i u(t). \quad (11)$$

Squaring and integrating Eq. (11) over time gives

$$\int_{-\infty}^{+\infty} \left(\sum_{i=1}^N u_i(t + \tau_i) \right)^2 dt = \left(\sum_{i=1}^N a_i \right)^2 \int_{-\infty}^{+\infty} u^2(t) dt. \quad (12)$$

Equation (10) gives

$$a_i^2 = \int_{-\infty}^{+\infty} u_i^2(t) dt / \int_{-\infty}^{+\infty} u^2(t) dt. \quad (13)$$

Substituting Eq. (13) into Eq. (12) gives

$$\int_{-\infty}^{+\infty} \left(\sum_{i=1}^N u_i(t + \tau_i) \right)^2 dt = \left(\sum_{i=1}^N \sqrt{\int_{-\infty}^{+\infty} u_i^2(t) dt} \right)^2. \quad (14)$$

This leads to the definition of the waveform similarity factor r as

$$r = \sqrt{\int_{-\infty}^{+\infty} \left(\sum_{i=1}^N u_i(t + \tau_i) \right)^2 dt} / \sum_{i=1}^N \sqrt{\int_{-\infty}^{+\infty} u_i^2(t) dt}. \quad (15)$$

Mallart and Fink,¹⁸ who independently proposed a focusing criterion as $C = r^2$, showed that for a point source the value of C lies between 0 and 1. Thus this also holds for r , but note that r can only be 1 if the τ_i 's are selected properly.

Beam profiles in the focal plane of the array were calculated as the rms value in time of each spatial position. These profiles were used as visual characterization of the aberration caused by the two established body wall models. They were also used for the visual evaluation of the effect of the different aberration correction filters.

Additional evaluation of the aberration correction was performed by analysis of the focus quality of the transmitted aberration corrected beam. Liu and Waag¹⁹ have developed a method for evaluation of the focus quality of a transmitted beam. This method was also used here. A brief explanation of this method, as described in Ref. 19, is given here. As described in Sec. II C, the signal that is propagated has a spatial (array direction) and temporal dimension. The method calculates the envelope of the analytic signal in the focal plane, thus obtaining a two-dimensional matrix of positive numbers which can be used for various focal quality measures. A first measure is obtained by projection of the maximum value of the analytic envelope amplitudes in both array and time direction. This provides one-dimensional amplitude profiles in both these directions. Effective widths of these

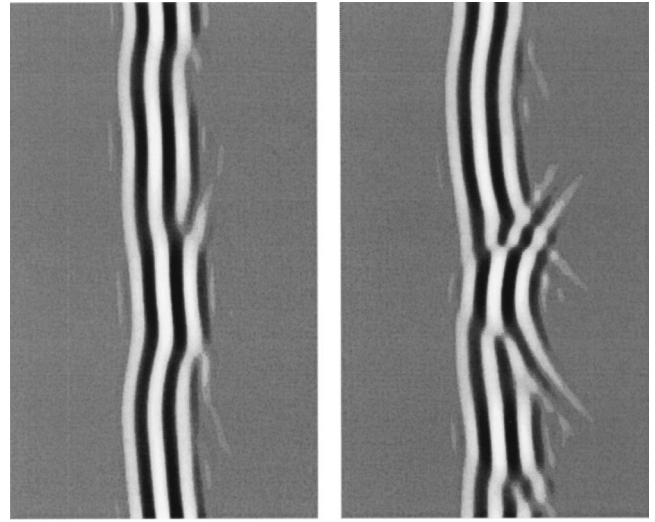


FIG. 3. Received waveforms after correction of geometric focusing for the homogeneous medium. Each panel shows a temporal logarithmic gray scale picture of the waveforms at the receiving array. The dynamic range of the logarithmic scale is 40 dB. The horizontal axis represents time and spans 4.8 μ s and the vertical axis represents the receiving array and spans 20 mm. The left-hand panel shows the waveform received after propagation through the w6 body wall model and the right-hand panel shows the waveform from the s6 body wall model.

profiles were found at -10 dB below the peak value. The effective width of such a profile was defined as the square root of the product of the radius in each direction of the profile, at the -10 dB level from the peak. A measure denoted the effective diameter was then calculated as the square root of the product of the effective widths in each direction. The effective diameter is hence the diameter of a circle with the same area as an ellipse defined by the effective widths in space and time. This provided a measure of the size of the focus at -10 dB from the peak value. The effective widths are also used in the evaluation of the -10 dB peripheral energy ratio. An ellipse was specified with axes equal to half of the effective widths in space and time, defining a central and peripheral region of the signal. The -10 dB energy ratio was then calculated as the energy outside the ellipse to the energy inside it. These measures, in combination with visual inspection of the beam profiles of the focal plane, gave substantial information of the focus quality of the corrected beam.

IV. RESULTS

Figure 3 shows the received waveforms after correction of geometric focusing for the homogeneous medium. The left-hand panel shows the weak aberration case and the right-hand panel the strong aberration case. Shape distortion of the waveform is clearly visible, and in the strong aberration case deformation of the waveform also occurs in several places. This was, as earlier described, one of the main selection criteria for the strong aberration body wall model as a case of severe aberration. Since the waveforms are pictured with a logarithmic scale and a dynamic range of 40 dB, amplitude variations do not show.

TABLE I. Body wall models statistics and waveform similarity factors (WSF).

Wall model	Arrival time fluctuations		Energy level fluctuations		WSF (*1e-3)
	rms value (ns)	Corr len (mm)	rms value (dB)	Corr len (mm)	
w2	36.0	3.6	2.9	1.6	971
w4	37.5	3.6	2.4	1.6	984
w6	49.4	4.4	3.1	1.8	975
w8	47.7	5.0	2.5	2.2	992
s2	39.6	3.0	3.6	2.0	957
s4	54.5	19.4	4.0	2.0	924
s6	53.7	5.8	4.1	1.4	914
s8	69.8	9.2	3.8	1.8	949

Table I shows the statistical properties and the waveform similarity factors for the four body wall models in each class (weak and strong), where the w6 and s6 walls are shown in boldface. The values for the arrival time fluctuations are based on the phase front estimation method (a). The rms values of the arrival time fluctuations for the w6 and s6 walls are well above the mean value of those reported by Hinkelman *et al.*¹⁰ In particular, the energy level fluctuations for the s6 wall model are higher than any of the reported values. Correlation lengths for the arrival time and energy fluctuations are well below the mean of the reported values for both w6 and s6. The waveform similarity factors have been calculated after correction for geometric focusing and aberration correction with the phase front tracking algorithm. Compared to the received waveform, the phase front tracking method shows a very accurate representation of the time delay variation of the wave front. For the w6 wall the waveform similarity is quite high, indicating low level of wave deformation which also can be seen in Fig. 3(a). The waveform similarity for the s6 wall on the other hand is quite low quantifying the amount of waveform deformation clearly visible in Fig. 3(b).

Figure 4 shows the estimated arrival time fluctuations

TABLE II. Normalized L_2 norm. The L_2 norm between the phase front tracking method (a) and all the other methods for both the weak and strong aberration cases, where the L_2 norm is normalized with respect to (a).

w6 case		s6 case	
Method	L_2	Method	L_2
(a)	0.0	(a)	0.0
(b)	0.34	(b)	0.98
(c)	0.40	(c)	0.45
(d)	0.39	(d)	0.28
(e)	0.42	(e)	1.43
(f)	0.17	(f)	0.33

for the weak aberration case (w6 body wall model). Figure 4(a) shows the arrival time fluctuations as determined by the phase front tracking algorithm. This has been used as a reference in the other plots, indicated by a dotted line, for comparison of the estimated arrival time fluctuations to the actual time delay variation of the phase front. The figure shows that the different estimation methods produce very similar results for the weak aberration case. Table II shows the normalized L_2 -norm between the arrival time fluctuation estimation methods, where the phase front tracking method (a) serves as the reference. Table II shows that there are some deviations between the frequency mean method (e), the different cross-correlation techniques (b), (c), (d), and the time delay variation of the phase front, but they are relatively small. The ray delays method (f) and the arrival time fluctuations of the phase front are almost identical in this case. It must be emphasized that method (f) is not an estimation method, but is merely used as a qualitative measure of the arrival time fluctuations.

In Fig. 5 the arrival time fluctuation estimates from the strong aberration case is presented. In this case several of the arrival time fluctuation estimation methods fail to estimate the time delay variation of the phase front of the received waveform. The neighbor cross-correlation method (b) and the frequency mean method (e), show large deviations in the

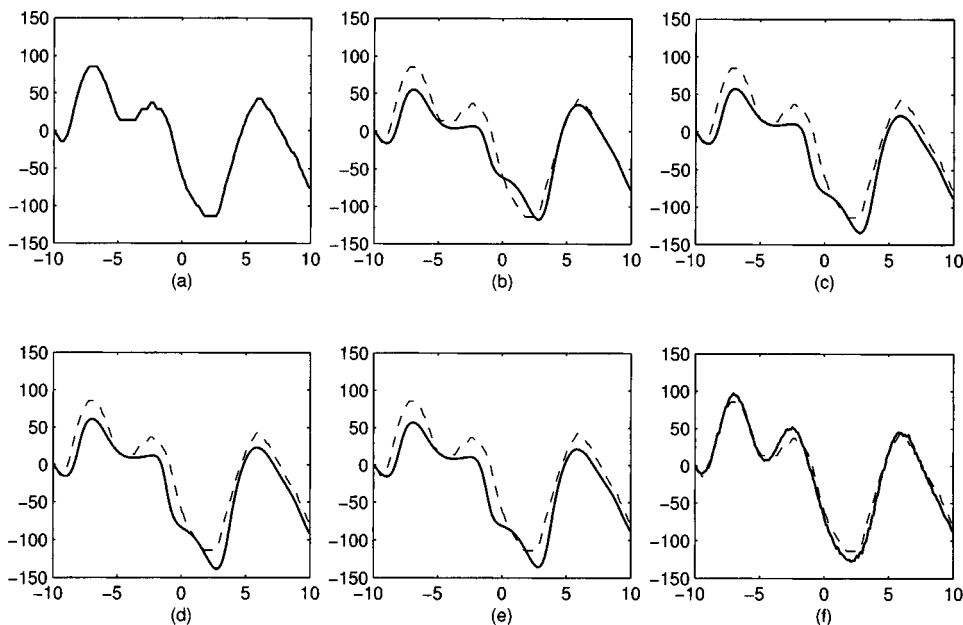


FIG. 4. Arrival time fluctuation estimates from the w6 wall model. The notation (a), (b), etc., in the different panels follows the notation introduced for the different arrival time estimators in Sec. III. The horizontal axis represents the array direction (mm), and the vertical axis the time delays (ns). For the horizontal axis, zero represents the center of the array and -10 mm the first channel in the array, which was used as reference for the reference channel cross-correlation method (d), and starting point of the cumulative summation in method (b). In each panel the solid line represents the estimate, and the dotted line the arrival time fluctuations of the phase front method (a), which is only plotted as a reference.

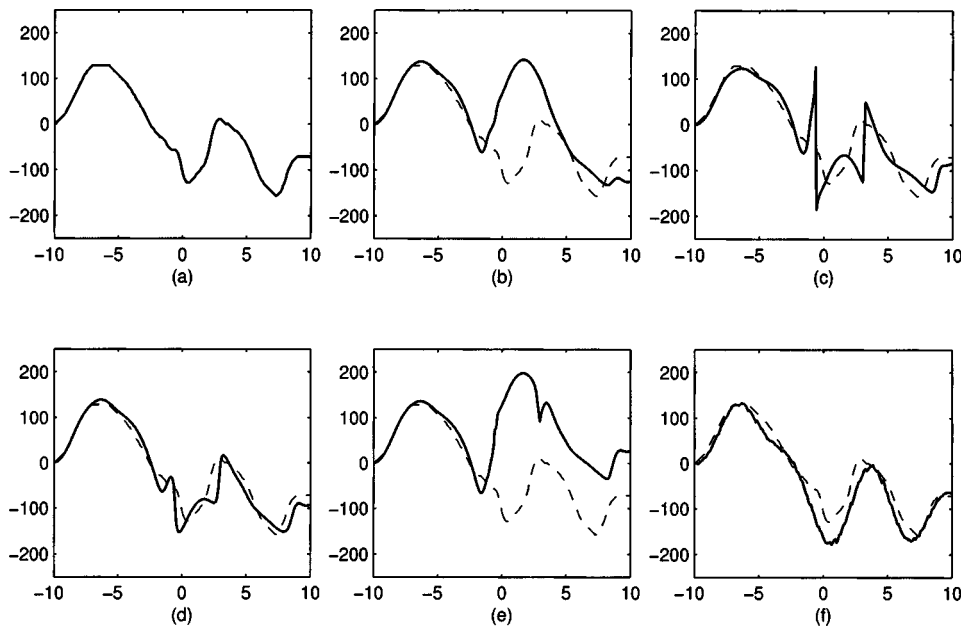


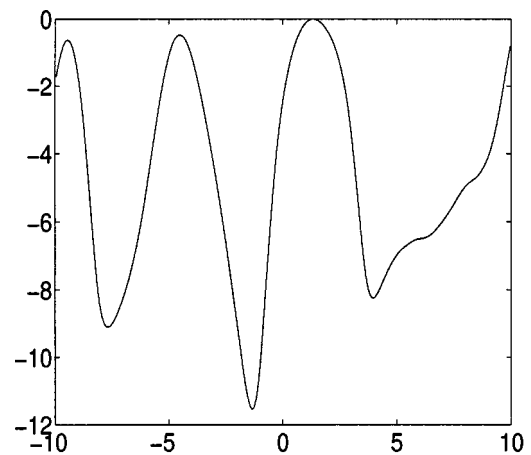
FIG. 5. Arrival time fluctuation estimates from s6 wall. The same notation as in Fig. 4 is used.

estimates and the L_2 norm is large. The beamsum cross-correlation method (c) shows “jumps” in the estimates. By closer inspection these deviations and “jumps” all occur in the shape deformed region of the received waveform. The reference channel cross-correlation method (d), estimates the time delay variations very well, but also here the estimates show smaller “jumps” in the shape deformed region of the received waveform. There are only smaller variations in the L_2 norms for methods (c), (d), and (f).

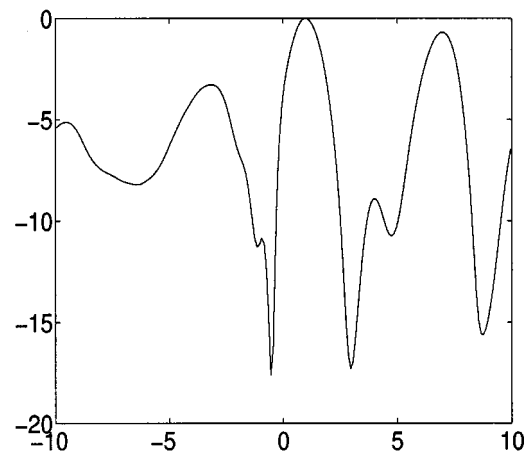
Amplitude variations across the receiving array for the weak and strong aberration cases are presented in Fig. 6. In the strong aberration case amplitude variations of approximately 18 dB are observed.

Figure 7 shows beam profiles corrected with the generalized frequency-dependent correction filter for both the weak and strong aberration cases. This is the best correction obtainable with a transmit correction filter of any kind, and the correction is very close to the ideal case. Table III presents the results of the focus quality calculations as described in Sec. III. These values give a more detailed description of what is observed in the beam profiles as time [which is the same as depth in the z direction (see Fig. 1)] is also included. It must be stressed that the focus quality variables in Table III cannot be directly linked to the presented beam profiles, as they are calculated in a different way. After correction with the generalized frequency-dependent correction filter, the peripheral energy ratios are almost identical to the ideal case even if the effective widths are slightly higher.

Beam profiles in the focal plane after correction with a time delay correction filter for the weak aberration case are shown in Fig. 8. The aberrated profiles show side lobe level of -5 dB from the peak value, and the profile is quite wide. Since the arrival time fluctuation estimates for the weak aberration case are very similar, only small differences in the different corrected profiles are visible. By visual inspection of the beam profiles, a general reduction of about 10 dB in the strongest side lobes is obtained with the time delay correction filter. Further away from the central axis the correc-



(a) w6 case



(b) s6 case

FIG. 6. Amplitude variations across the receiving array for the weak and strong aberration cases. The horizontal axis represents the array direction (mm) and the vertical axis the amplitude (dB).

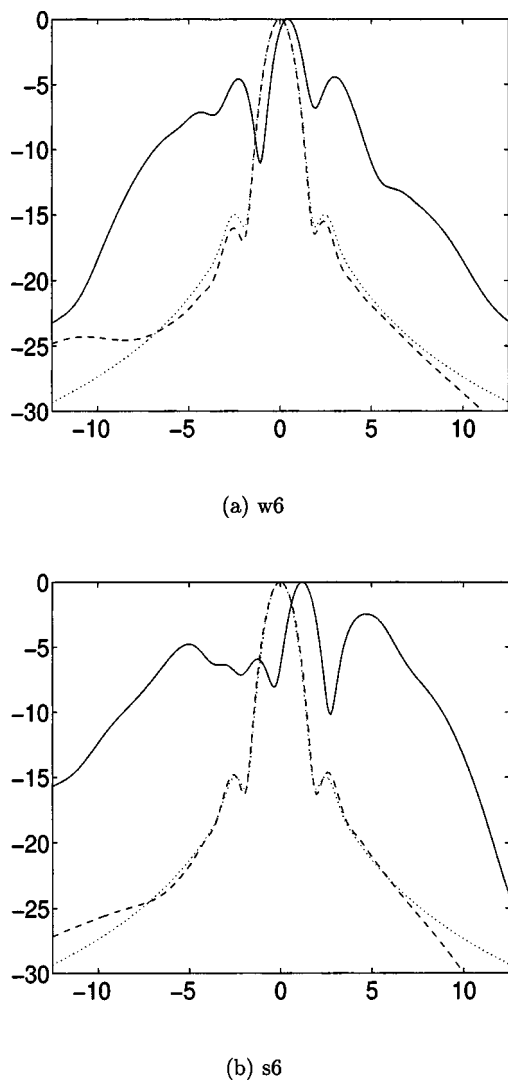


FIG. 7. Beam profiles corrected with generalized frequency-dependent correction filter. Left-hand panel shows the weak aberration case and the right-hand panel the strong aberration case. The horizontal axis represents the focal plane (mm) and the vertical axis the energy (dB). The solid line represents the aberrated profile, the dashed line the corrected profile, and the dotted line the profile with no aberration, which serves as a reference. The zero point on the horizontal axis denotes the center axis of the transmitting array. All profiles are normalized to their maximum value.

tion is less substantial. The effective widths at -10 dB in space and time are identical to the ideal case, but the peripheral energy ratio is relatively high, because the side lobe level is still high in comparison to the ideal case. Note that the ray delays method (f) gives the poorest correction.

Beam profiles corrected with a time delay and amplitude correction filter for the weak aberration case are presented in Fig. 9. Now the strongest side lobes are reduced to the side lobe level of the nonaberrated profile, and a strong reduction in the side lobe level further away from the central axis can be noticed. Also here the effective widths in Table III at -10 dB are very close to the ideal case. The -10 dB peripheral energy ratios with time delay and amplitude correction are very close to the ideal case, and almost half the value of time delay correction only. The peripheral energy ratios for the phase front tracking method (a) and the ray delays method (f) are somewhat higher than for the other methods.

Figure 10 shows beam profiles with time delay correction for the strong aberration case. Now the aberrated profile is severely distorted and very wide. As shown in Fig. 5, the estimation neighbor cross-correlation method (b) and the frequency mean method (e), fail to predict the arrival time fluctuations of the wave front, and they produce little or no correction at all. The other methods perform very well, especially the beamsum and reference channel cross-correlation methods (c) and (d), but it is quite obvious that this is a more difficult case. From Table III the -10 dB peripheral energy ratios are somewhat higher for this case than for the weak aberration case, even if the effective widths have not increased much. It should be noted that there is great uncertainty about the focus quality parameters of methods (b) and (e) as there is no clearly defined main lobe. This poses problems for the calculation of the effective widths. The same argument holds for the aberrated profiles. It should also be pointed out that even if there is only a small difference in the L_2 norm between method (d) and the ray delays method (f), the difference in the -10 dB peripheral energy ratio is large.

Beam profiles with time delay and amplitude correction are shown in Fig. 11. Now the correction is quite substantial with all methods except (b) and (e). The focus quality parameter in Table III shows that the beamsum cross-correlation and reference channel methods (c) and (d) perform very well also for the strong aberration case. With time delay and amplitude correction the peripheral energy ratios with these methods are very close to those for the weak aberration case (and thus the ideal case), indicating their robustness. Notice that even if method (c) has “jumps” in the arrival time fluctuations estimates, it still performs quite well. Notice also that the peripheral energy ratio for method (f) is almost double the value of method (d), even if the differences in the estimated arrival time fluctuations are relatively small (see Fig. 5). This indicates the sensitivity of the arrival time estimate and illustrates the importance of a “good” estimate.

V. DISCUSSION

Several studies have modeled the human body wall with single or multiple time delay screens.^{3,5-9} Some of these have used measured or estimated human body wall statistics as a basis for a realization of a simple random distribution, used as a time delay aberration screen. Usually the statistics are rms values and correlation lengths of arrival time and amplitude fluctuations. In Ref. 15, such models are criticized on the basis that even if they produce wave distortion with statistics similar to the human body wall, they are not truly able to reproduce human body wall aberration. This is simply due to the fact that a body wall does not represent a true random medium, and that it is the tissue structure of the wall that introduces the aberration. An example of this is septa aligned close to the propagation direction in the subcutaneous fat layer. As demonstrated by Mast *et al.*, this structure causes scattering of energy and generates arrival time fluctuations that increase cumulatively with propagation through the fat layer.¹⁵ The use of time delay screen models is thus criticized as not accounting for aberrations created by the

TABLE III. Table of focus quality with respect to applied correction method. The peripheral energy ratio and all effective widths have been calculated at -10 dB level from the peak. The two column headings refer to the weak and strong aberration cases, respectively. The column named “Estimator” indicates which time delay fluctuation estimation method that is used. The “Correction” column specifies the type of correction where τ denotes time delay correction, and τ &a time delay and amplitude correction. The three lowest rows have no value in the “Correction” column. The first of these rows shows the correction with the generalized frequency-dependent correction filter, the second shows the aberrated case, and the third shows the ideal case with no aberration.

w6 case						s6 case					
Estimator	Correction	Peripheral energy ratio	Effective diameter (mm)	Effective width in array direction (mm)	Effective width in time direction (mm)	Estimator	Correction	Peripheral energy ratio	Effective diameter (mm)	Effective width in array direction (mm)	Effective width in time direction (mm)
(a)	τ	0.486	1.84	2.70	1.26	(a)	τ	0.797	1.84	2.80	1.21
(b)	τ	0.446	1.84	2.70	1.26	(b)	τ	2.370	3.16	4.70	2.13
(c)	τ	0.426	1.84	2.70	1.26	(c)	τ	0.575	1.91	2.90	1.26
(d)	τ	0.426	1.84	2.70	1.26	(d)	τ	0.579	1.88	2.80	1.26
(e)	τ	0.426	1.84	2.70	1.26	(e)	τ	0.850	3.51	5.90	2.13
(f)	τ	0.518	1.84	2.70	1.26	(f)	τ	0.932	1.91	2.90	1.26
(a)	τ &a	0.301	1.84	2.70	1.26	(a)	τ &a	0.393	1.91	2.90	1.26
(b)	τ &a	0.264	1.88	2.80	1.26	(b)	τ &a	1.140	3.31	6.00	1.82
(c)	τ &a	0.242	1.88	2.80	1.26	(c)	τ &a	0.288	1.91	2.90	1.26
(d)	τ &a	0.242	1.82	2.80	1.26	(d)	τ &a	0.273	1.91	2.90	1.26
(e)	τ &a	0.242	1.88	2.80	1.26	(e)	τ &a	1.170	2.50	3.70	1.69
(f)	τ &a	0.331	1.84	2.70	1.26	(f)	τ &a	0.544	1.94	3.00	1.26
Filter		0.220	1.87	2.70	1.30	Filter		0.210	1.84	2.70	1.26
abb		1.220	2.58	4.50	1.47	abb		1.430	3.68	6.50	2.08
Ideal		0.227	1.84	2.70	1.26	Ideal		0.227	1.84	2.70	1.26

subcutaneous fat layer, and that they are more appropriate for describing aberration caused by muscle layers.¹⁵ Mast *et al.* show that muscle layers mainly introduce phase distortion and weak amplitude distortion, and that amplitude distortion basically originates from the inhomogeneities in the subcutaneous fat layer.¹⁵

In this article the body wall statistics has only been one of several criteria in the making of the body wall models. A selection of eight normally distributed random sequences was combined, filtered, and tuned to obtain eight different body wall models. Then a careful selection of two specific wall models among the eight was performed. The selection criteria were visual inspections of gray scale pictures of re-

ceived waveforms, calculation of body wall statistics and waveform similarity factors, visual evaluation of arrival time fluctuations and amplitude variations, and evaluation of aberrated beam profiles in the focal plane. The body wall statistics were compared with similar statistics as measured correlation lengths for the human abdominal wall.¹⁰ This evaluation was performed in order to select two body wall models that produced weak and strong aberration, respectively. As seen from Table I the w6 body wall model has the toughest body wall statistics in the weak category even if the waveform similarity is somewhat higher than the w2 wall. When comparing all the above-mentioned criteria the w6 model still came out as the worst one in its category. In the

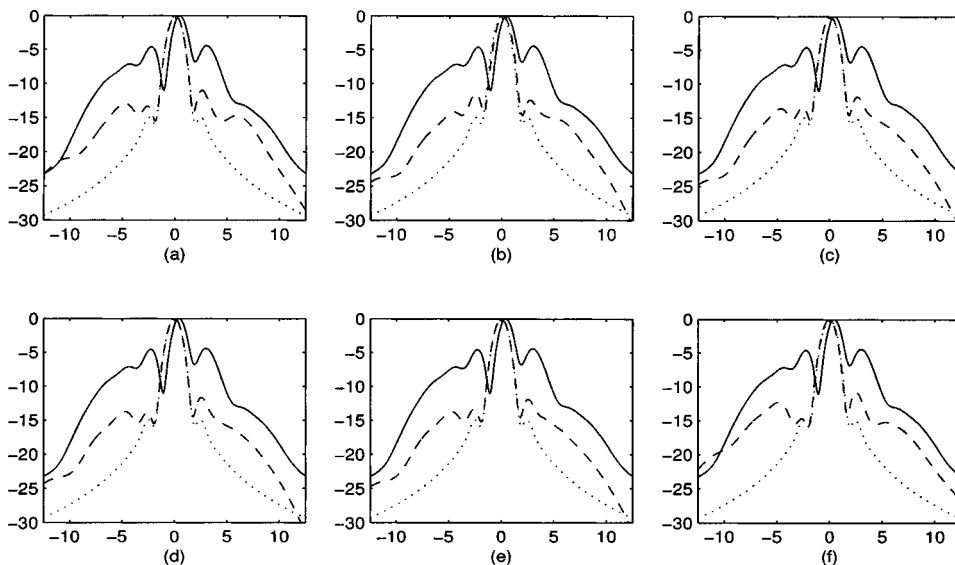


FIG. 8. Beam profiles in the focal plane corrected with a time delay correction filter for the weak aberration wall model. Same notation as in Fig. 7.

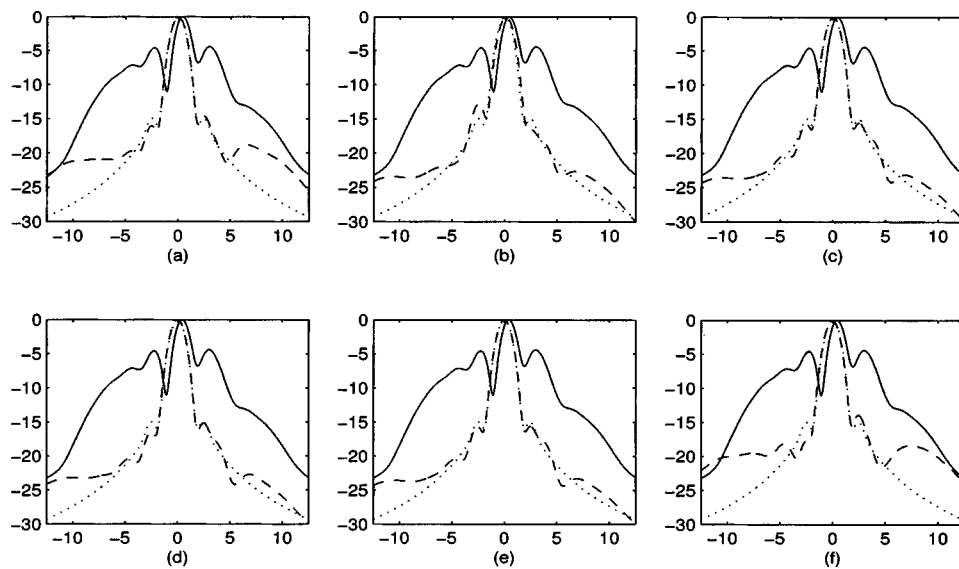


FIG. 9. Beam profiles in the focal plane corrected with a time delay and amplitude correction filter for the weak aberration wall model. Same notation as in Fig. 7.

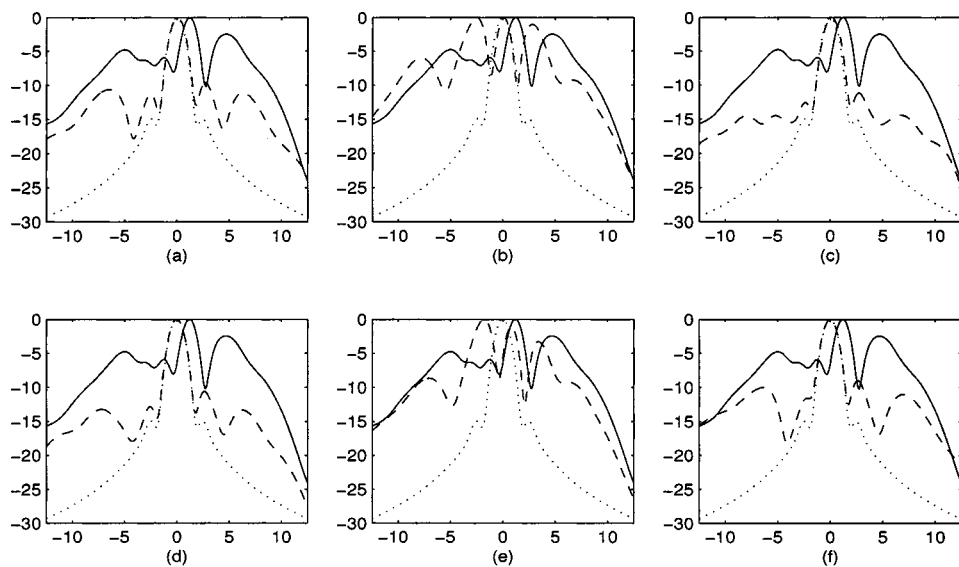


FIG. 10. Beam profiles with time delay correction for the strong aberration wall model. Same notation as Fig. 7.

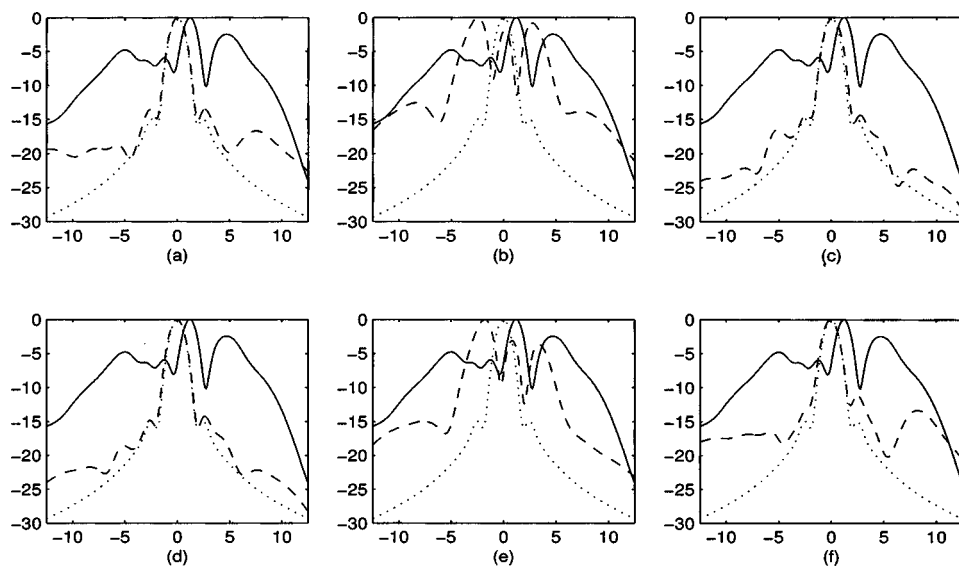


FIG. 11. Beam profiles with time delay and amplitude correction for the strong aberration wall model. Same notation as Fig. 7.

strong category the s8 wall has very high rms value of the arrival time fluctuations, but a relatively high correlation length and waveform similarity factor. This makes arrival time estimation easier, and this in combination with the other criteria is why the s6 wall was preferred as a case of severe aberration.

A close inspection of the arrival time fluctuations for the weak and strong aberration cases (Figs. 4 and 5) show variations that can be interpreted as cumulative time delays resulting from propagation through larger or smaller structures. A comparison of the time delay screens for these two cases with the arrival time fluctuations obtained with time delay estimator (f), i.e., summing the time delay along a straight line from the focus to a channel on the receiving array, showed that the specific variations did not stem from a single strong time delay screen alone. Large variations were the result of high correlation between all six time delay screens in the direction of summation, smaller variations were results from lower correlation. These observations indicate that the w6 and s6 model walls do model correlated structures of different sizes. A structure such as septa aligned close to the propagation direction in the subcutaneous fat layer would result in a large drop in amplitude level over a small region on the receiving array. The amplitude variation of the s6 model wall [Fig. 6(b)] shows this type of amplitude variation. An amplitude drop of about 18 dB is observed, which is close to what is observed in Ref. 18.

The waveform similarity factor was only used to quantify the amount of waveform deformation in the received signals. As explained in Sec. IV, the waveform similarity factor was calculated after correction for geometric focusing and aberration correction with the phase front tracking algorithm. Since the factor is dependent on the selected time delays τ_i [see Eq. (15), Sec. III], it is dependent on the arrival time estimation method. In this article simulations were performed with point sources without noise, and the phase front tracking algorithm will in such a situation always give a reliable estimate of the time delay variation of the wave front, since it is clearly defined. The method proved to give an accurate estimate of the time delay variation of the received wave front, giving a very good focusing of the received signal. This means that the variations in the waveform similarity factor in Table I are only due to the difference in waveform deformation of the received aberrated signal, and thus gives an accurate quantification of the waveform deformation. Note that in a more “real” case, with interference from multiple scatterers and acoustic and electronic noise, the phase front tracking method will probably be useless since there no longer exists an evident wave front.

A weakness in the simulations was that absorption effects were not included. Mast *et al.* have shown that both absorption and scattering effects makes a significant contribution to the total insertion loss in the human body wall.¹⁵ They suggest that estimation of attenuation by the human abdominal wall for correction of amplitude aberration should consider both scattering and absorption effects. They indicate that scattering accounts for about 36% of total insertion loss in the human abdominal wall at 3.75 MHz.¹⁵ Attenuation of

the amplitudes by absorption would probably deteriorate the amplitude correction.

In both the weak and strong aberration cases, correction with a generalized frequency-dependent correction filter produced very close to ideal correction. For the weak aberration case the different arrival time fluctuation estimators all performed well. In the strong aberration case both neighbor cross-correlation method (b), and the frequency mean method (e), failed severely. This is obviously due to the waveform deformation in this case, which also caused problems for the beamsum and reference channel cross-correlation methods [(c) and (d)]. In Fig. 5, it is clearly visible that the reference channel cross-correlation method encounters problems in the same region as the cross-correlation beamsum method, but they are less severe. This indicates that the choice of reference is of utmost importance in cross-correlation techniques, where strong waveform distortion and deformation occurs. On the other hand these methods perform rather well despite the problems. This can be explained by comparing the “jumps” in the estimates with methods (c) and (d), and the amplitude variation for the strong aberration case [Fig. 6(b)]. The large drops in the amplitude variation correspond with the regions where the waveform is deformed [Fig. 3(b)]. The “jumps” in the arrival time fluctuation estimates occur in the exact same places as the amplitude drops, due to the waveform deformation. Therefore the “jumps” in the arrival time fluctuation estimates do not influence the correction, because the amplitudes are negligible in these regions on transmit. It is also worth commenting that the arrival time fluctuations obtained with the phase front tracking algorithm do not produce the best aberration correction, although this could be expected intuitively since this method gives the actual time delay variation of the wave front. This is just a statement of the difficulty of properly defining the time delay between two signals in ultrasonic imaging.

The peripheral energy ratios after correction with the generalized frequency correction filter are actually somewhat better than the ideal nonaberration situation for both body wall models. This is probably due to the Wiener-type filtering process that is performed to obtain the generalized frequency correction filter. When the filter is applied on transmit, it slightly changes the transmitted pulse, and in this case this results in lower peripheral energy ratios than the ideal situation.

The differences in the L_2 norms should also be commented on further. For the weak aberration case, the ray delays method (f) has more than half the L_2 norm compared to beamsum cross-correlation method (c), but the peripheral energy ratio is higher. For the strong aberration case there are only smaller differences in the L_2 norms between methods (c), (d), and (f). Still the peripheral energy ratio of the ray delays method (f) is almost double the value of both methods (c) and (d). This indicates that in some cases, the effect of the correction can be very sensitive to how the arrival time fluctuations are estimated. Relatively small differences in the arrival fluctuation estimates can give large variations in the focus spot size of the corrected beam.

VI. CONCLUSIONS

Different correction filters for ultrasonic wave aberration have been investigated. The results in this paper indicate that correction of aberration on transmit with a time delay and amplitude correction filter produce substantial correction on the focus quality of the transmitted beam. Correction with a time delay filter alone performed poorer than correction with a time delay and amplitude filter, but improved focus quality considerably. Two cases were studied, a weak and strong aberration case, where the strong aberration case represented a case of severe wave aberration. For both cases correction with a time delay and amplitude filter gave results close to the ideal nonaberration case, indicating that this correction method is very robust. The main criterion for this method to work was that the arrival time fluctuations were estimated correctly. Six different arrival time estimation methods were investigated. The best result was obtained with a method that calculated the cross-correlation at zero lag of the complex envelope of the received signal, with a reference signal. Two reference signals were tested: the beamsum and a signal from a channel on the transducer. The two different references produced approximately the same aberration corrections, even if there were some differences in the estimated arrival time fluctuations.

ACKNOWLEDGMENT

The work presented in this article was supported by the Medicine and Health program of the Research Council of Norway.

- ¹S. W. Flax and M. O'Donnell, "Phase-aberration correction using signals from point reflectors and diffuse scatterers: Basic principles," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **35**, 758–767 (1988).
- ²M. Fink, C. Prada, F. Wu, and D. Cassereau, "Self focusing in inhomogeneous media with time reversal acoustic mirrors," *IEEE Ultrasonics Symposium*, 1989, pp. 681–686.
- ³D.-L. Liu and R. C. Waag, "Correction of ultrasonic wavefront distortion using backpropagation and a reference waveform method for time-shift compensation," *J. Acoust. Soc. Am.* **96**, 649–660 (1994).
- ⁴J. C. Lacefield and R. C. Waag, "Evaluation of backpropagation methods

- for transmit focus compensation," *IEEE Ultrasonics Symposium* 2001, Vol. 2, pp. 1495–1498.
- ⁵A. P. Berkhoff and J. M. Thijssen, "Correction of concentrated and distributed aberrations in medical ultrasound imaging," 1996 *IEEE Ultrasonics Symposium Proceedings*, Vol. 2, pp. 1405–1410.
- ⁶S. W. Flax and M. O'Donnell, "Phase-aberration correction using signals from point reflectors and diffuse scatterers: Measurements," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **35**, 768–774 (1988).
- ⁷L. A. Ødegaard, "Using signals scattered from diffuse inhomogeneities to correct for phase aberration caused by a phase-screen far from the transducer," 1995 *IEEE Ultrasonics Symposium Proceedings*, Vol. 2, pp. 1443–1447.
- ⁸S. Krishnan, P. C. Li, and M. O'Donnell, "Adaptive compensation of phase and magnitude aberrations," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 44–55 (1996).
- ⁹W. E. Walker and G. E. Trahey, "Aberrator integration error in adaptive imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 780–791 (1997).
- ¹⁰L. M. Hinkelman, D.-L. Liu, R. C. Waag, Q. Zhu, and B. D. Steinberg, "Measurements of ultrasonic pulse arrival time and energy level variations produced by propagation through abdominal wall," *J. Acoust. Soc. Am.* **95**, 530–541 (1994).
- ¹¹L. M. Hinkelman, D.-L. Liu, R. C. Waag, Q. Zhu, and B. D. Steinberg, "Measurement and correction of ultrasonic pulse distortion produced by the human breast," *J. Acoust. Soc. Am.* **97**, 1958–1969 (1995).
- ¹²L. M. Hinkelman, T. L. Szabo, and R. C. Waag, "Measurements of ultrasonic pulse distortion produced by human chest wall," *J. Acoust. Soc. Am.* **101**, 2365–2373 (1997).
- ¹³T. D. Mast, L. M. Hinkelman, M. J. Orr, V. W. Sparrow, and R. C. Waag, "Simulation of ultrasonic pulse propagation through the abdominal wall," *J. Acoust. Soc. Am.* **102**, 1177–1190 (1997).
- ¹⁴L. M. Hinkelman, T. D. Mast, L. A. Metlay, and R. C. Waag, "The effect of abdominal wall morphology on ultrasonic pulse distortion. I. Measurements," *J. Acoust. Soc. Am.* **104**, 3635–3650 (1998).
- ¹⁵T. D. Mast, L. M. Hinkelman, M. J. Orr, and R. C. Waag, "The effect of abdominal wall morphology on ultrasonic pulse distortion. II. Simulations," *J. Acoust. Soc. Am.* **104**, 3651–3664 (1998).
- ¹⁶B. A. J. Angelsen, *Ultrasound Imaging* (Emantec, Norway, 2000), Vol. II, <http://www.ultrasoundbook.com>.
- ¹⁷M. O'Donnell and W. E. Engeler, "Real-time phase aberration correction system for medical ultrasound imaging," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 12, 1990.
- ¹⁸R. Mallart and M. Fink, "Adaptive focusing in scattering media through sound speed inhomogeneities: The Van Cittert Zernicke approach and focusing criterion," *J. Acoust. Soc. Am.* **96**, 3721–3732 (1994).
- ¹⁹D.-L. Liu and R. C. Waag, "Time-shift compensation of ultrasonic pulse focus degradation using least-mean-square error estimates of arrival time," *J. Acoust. Soc. Am.* **95**, 542–555 (1994).

Modeling individual differences in ferret external ear transfer functions

Jan W. H. Schnupp,^{a)} John Booth, and Andrew J. King

University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, United Kingdom

(Received 19 July 2002; accepted for publication 16 December 2002)

Individual variations in head and outer ear size, as well as growth of these structures during development, can markedly alter the values of the binaural and monaural cues which form the basis for auditory localization. This study investigated individual differences in the directional component of the head-related transfer function of both adult and juvenile ferrets. In line with previous studies in humans and cats, intersubject spectral differences were found to be reduced by scaling one of the directional transfer functions on a log-frequency axis. The optimal scale factor correlated most highly with pinna cavity height. Optimal frequency scaling reduced interear spectral difference equally well for adult–juvenile comparisons as for comparisons between pairs of adult ears. This illustrates that the developmental changes in localization cue values should be at least partly predictable on the basis of the expected growth rate of the outer ear structures. Predictions of interaural time differences (ITDs) were also derived from the physical dimensions of the head. ITDs were found to be poorly fitted by the spherical head model, while much better predictions could be derived from a model based on von Mises spherical basis functions. Together, these findings show how more accurate estimates of spatial cue values can be made from knowledge of the dimensions of the head and outer ears, and may facilitate the generation of virtual acoustic space stimuli in the absence of acoustical measurements from individual subjects. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1547460]

PACS numbers: 43.64.Ha, 43.66.Qp, 43.66.Pn [LHC]

I. INTRODUCTION

Sound traveling from a free-field source to the eardrum is filtered by interactions with the head and the external ear. This transformation is known as the head-related transfer function (HRTF) (Butler, 1987; Wightman and Kistler, 1989). The HRTF is commonly thought of as the sum of two components, a “canal component,” which includes ear canal resonances and other acoustic features that are independent of the angle of incidence of the sound wave, and a “directional transfer function” (DTF), which contains spectral features (peaks and notches) that vary as a function of sound source direction and provide important spectral cues to sound source location (Batteau, 1967; Blauert, 1969; Middlebrooks and Green, 1990; Shaw, 1974). Artificially filtering signals delivered over headphones with the DTF makes it possible to generate “virtual free-field sounds,” which can be externalized and localized with good accuracy (Wightman and Kistler, 1989a, 1989b). Due to variation in the size and shape of the head and outer ears, DTFs differ between individuals (Mehrgardt and Mellert, 1977; Pralong and Carlile, 1994; Wightman and Kistler, 1989). Consequently, listeners localizing sounds in virtual acoustic space (VAS) are likely to make more and larger localization errors when the VAS stimuli are derived from the ears of a different subject, rather than their own DTF (Wenzel *et al.*, 1993). Recently, Middlebrooks (1999a) reported that individual differences in DTF amplitude spectra can be reduced when one of the DTFs is scaled along a log-frequency axis. The optimal scale factor

(OSF) was highly correlated with the pinna height ratio of the two individuals. Furthermore, when human listeners were required to localize sounds in VAS based on nonindividual (“foreign”) DTFs, localization performance improved when these foreign-ear DTFs were optimally frequency scaled (Middlebrooks, 1999b). Frequency scaling has also been reported to reduce individual DTF differences in cats (Xu and Middlebrooks, 2000), although not quite to the same extent as in humans. In this study we have investigated systematic variations in individual DTFs of the ferret (*Mustella putorius*) and possible correlations with the physical characteristics of individual animals. Ferrets have become an increasingly important animal model in physiological and psychophysical research.

One question of interest is how the auditory system is able to adapt during ontogenetic development to changes in acoustic localization cues due to the growth of the head and outer ear. Middlebrooks’ study (1999a) showed that pinna size differences in adults are to some extent predictive of differences in individual DTFs. If this observation can be extended to comparisons between adult and infant ears, then at least some of the developmental changes in the DTF may also be predictable. We therefore included DTFs from juvenile as well as adult ferrets in this study, and tested directly whether size differences are as predictive of acoustic differences in adult–juvenile comparisons as they are in adult–adult comparisons.

In addition to DTFs, interaural time difference cues (ITDs) form another key component of VAS stimuli. We also considered how ferret ITDs may be predicted accurately from measurements of the physical dimensions of the head.

^{a)}Electronic mail: jan.schnupp@physiol.ox.ac.uk

II. METHOD

A. Measurement of physical dimensions and directional transfer functions

Acoustical recordings were taken from 25 ferrets (18 male, 7 female), aged between 27 and 897 days, with a mean of 145 days. To quantify the rate of growth of the head and outer ear in juveniles, we also measured the heads and outer ears of a further seven infant animals (4 male, 3 female, aged 7–20 days, mean 16 days). All animals used were purpose bred for scientific experimentation, and all procedures were carried out in compliance with A.S.A. guidelines and U.K. home office regulations.

Digital photographs were taken of both sides and the dorsal aspect of the ferrets' head. A centimeter rule was held against the animal, in frame, in each photograph. Physical size measurements were then derived from the photographs.

DTFs were recorded during the preparation of electrophysiological studies reported elsewhere, and further details on the procedures can be found in Mrsic-Flogel *et al.* (2001). Briefly, the animals were anesthetized with Saffan (2 ml/kg i.p., supplementary doses i.v. as required). The caudal aspect of each auditory meatus was exposed through an incision made behind the outer ears. A small opening was made in the wall of the meatus as close to the skull as practicable (ca. 5 to 7 mm from the distal end of the meatus). Probe tube microphones consisting of polythene tubes (length 20 mm, o.d. 1.52 mm, i.d. 0.86 mm) mounted on Sennheiser microphone capsules (KE-4-211-2) were inserted into the meatus through this surgical opening. The animal was positioned inside an anechoic chamber with the head held at the center of a robotic hoop (radius 65 cm) which moved a loudspeaker (Kef T27) automatically through 433 positions around the animal's head. We thus sampled almost the entire range of azimuths and elevations at ca 10° intervals, except for a 40° wide region behind the animal or positions of less than –60° elevation, which the hoop could not reach. (Here, negative azimuths denote positions to the animal's left, and negative elevations positions below the interaural plane.) At each speaker position 512-point Golay code pairs (sampled at 80 kHz, antialias filtered at 30 kHz) were presented through the loudspeaker and re-recorded through the probe tube microphones. DTFs were extracted from the recorded signals using the techniques described by Middlebrooks and colleagues (Middlebrooks, 1999a; Middlebrooks and Green, 1990; Xu and Middlebrooks, 2000; Zhou *et al.*, 1992). ITDs were extracted from the recorded impulse response functions using cross correlation after bandpass filtering (0.5–3.5 kHz). ITDs are largely, but not entirely, independent of sound frequency. Bandpass filtering the impulse responses prior to cross correlation “weights” our observed ITD function towards frequencies below the phase-locking limit. However, ITD values obtained with or without bandpass filtering differed typically by less than 20 μ s.

B. Quantification of individual DTF differences

Adopting the methodology used by Middlebrooks (1999a), we passed the measured DTFs through a filter bank of 104 triangular filters, with filter slopes of 105 dB per

octave, and center frequencies spaced in equal intervals of 1/35 octave from 3.75 to 30 kHz. Processing the data in this way roughly equalizes the number of samples per perceptual critical band and facilitates the scaling of DTFs in log frequency space. Bandpass filtering also blunts the sharp peaks and notches that may occur in the DTFs at the very high frequency end. These very high frequency notches are highly individual, as they may be influenced by very small structural features of the outer ear, but they are probably too narrow to be preserved after the sound is encoded by the cochlea. By approximating the effect of cochlear filters (Glasberg and Moore, 1990), the triangular filtering method provides an expedient method for selecting those parts of the spectral pattern of the DTFs that are likely to be perceptually relevant.

Animal-to-animal differences in DTFs were quantified by an *interear spectral difference* (IESD) measure. This was calculated in the following manner: dB amplitudes in each of the 104 filtered DTF components were subtracted, one ear from the other, to create a difference spectrum for each sound source location. The variance of the 104 difference spectrum components was then calculated. The IESD is defined as the mean of these variance values over all 433 sampled source directions. Like Middlebrooks and colleagues (Middlebrooks, 1999a; Xu and Middlebrooks, 2000), we investigated whether the observed IESDs could be reduced by scaling one of the filtered DTFs in logarithmic frequency space. This was done by recalculating the IESD repeatedly after shifting one of the filtered DTF spectra up or down by up to 30 steps of 1/35 octave. Sixty-one IESD values were calculated in all, one for each of 61 frequency scaling factors from –6/7th to +6/7th of an octave in 1/35-octave steps. The *optimal scale factor* (OSF) was defined as the frequency scaling (shift) factor associated with the lowest IESD. OSFs were computed to describe the frequency disparities of DTF sets for “same side” (i.e., two right or two left ears), as well as for “opposite side” ears (i.e., a left and a right ear).

III. RESULTS

We computed OSFs in the manner detailed above for all 1225 available ferret ear combinations (comparing each ear with every other ear, including the opposite ear of the same animal). Figure 1(A) illustrates the effect of frequency scaling on the DTFs of two ferret ears for one specific sound source direction. Note how downward scaling of the DTF of one ferret (F0018) aligns prominent high-frequency notches, thus reducing the difference between the DTFs. For this particular sound direction, scaling reduced the IESD from 8.7 to 4.8 dB². Prominent notches, like those seen near 25 kHz in Fig. 1(A), are often found at frontal sound source directions. More lateral directions, or those at higher elevations, may lack such notches but exhibit prominent peaks instead, as illustrated in Fig. 1(B). Here, frequency scaling results in an alignment of those peaks and reduced the IESD from 4.2 to 2.9 dB². Note that the magnitude of the frequency shift required to align the peaks in Fig. 1(B) (ca. 1/6th of an octave) is the same as that which aligned the notches shown in Fig. 1(A). Figure 1(C) shows the distribution of OSF values ob-

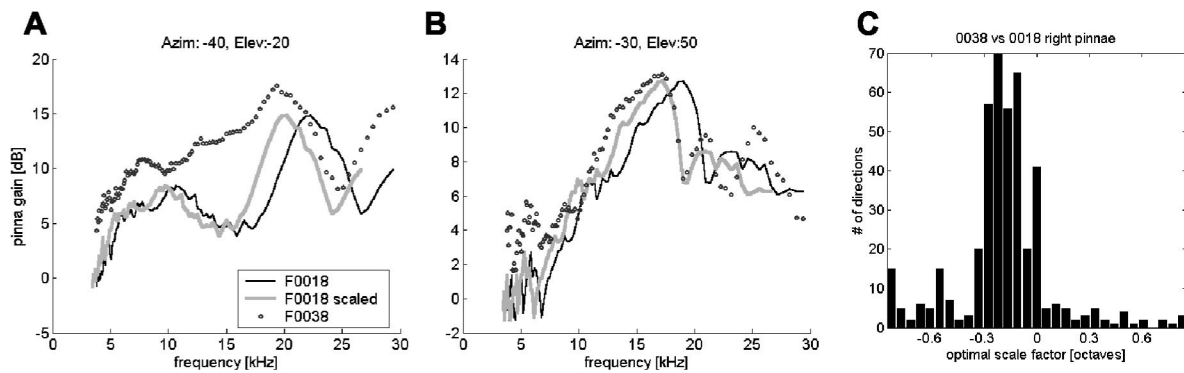


FIG. 1. Frequency scaling of DTFs for one particular ferret ear pair, the right ears of F0038, a large male, and F0018, an average size female. (A) Illustration of frequency scaling to reduce individual differences in DTFs. The DTFs of both ears at -40° azimuth, -20° elevation are shown, as well as the DTF of ferret F0038 after downward frequency scaling. The prominent high-frequency notches visible in the DTFs are brought into register by scaling. (B) Another example of the effect of frequency scaling, taken from the same animals and using the same scale factor, but for a different position at high elevations where notches in the spectrum are less apparent. Here, the effect of scaling was to align the prominent spectral peaks near 15 kHz. (C) Histogram revealing the distribution of best frequency scale factors for all the sound-source directions sampled, in the comparison of the right ears of F0018 and F0038. The large peak centered on about -0.25 octaves is indicative of the overall optimal scale factor.

tained at each sound direction. The fact that many sound source directions yield very similar OSF values indicates that frequency scaling can indeed reduce a general systematic frequency disparity between the DTFs of different ears. In this particular example, the distribution is centered around a shift of 0.18 octaves, the OSF calculated for the mean of all source locations; 47.4% of directions lie in the 8 bars between -0.21 and 0.014 octaves. Interestingly, the distribution shown in Fig. 1(B) shows a second small peak between approximately -0.52 and -0.8 octaves. Middlebrooks (1999a) reported similar secondary maxima in his study on human DTF differences. Here, as in Middlebrooks' study, the major contributor to these are sound directions behind the head or at low elevations, where signal-to-noise ratios in the DTF measurements are poor.

Figure 2 shows the distribution of absolute OSFs for the pairwise comparisons of all ears in our database. The median and maximal absolute OSFs were 0.1429 and 0.6571 octaves, respectively. The amount of IESD reduction achieved

by frequency scaling is quite variable, as illustrated in Fig. 3(A). IESDs were reduced by 11.76% or more for half of the ear-pair combinations and by more than 40% for 9.63% of ear pairs. For almost 20% of ear pairs, optimal

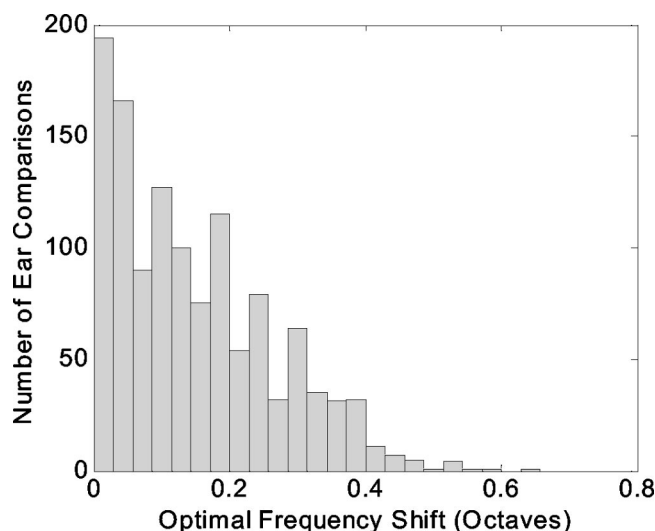


FIG. 2. Distribution of absolute optimal scale factors obtained for all ear-pair combinations ($n=1225$).

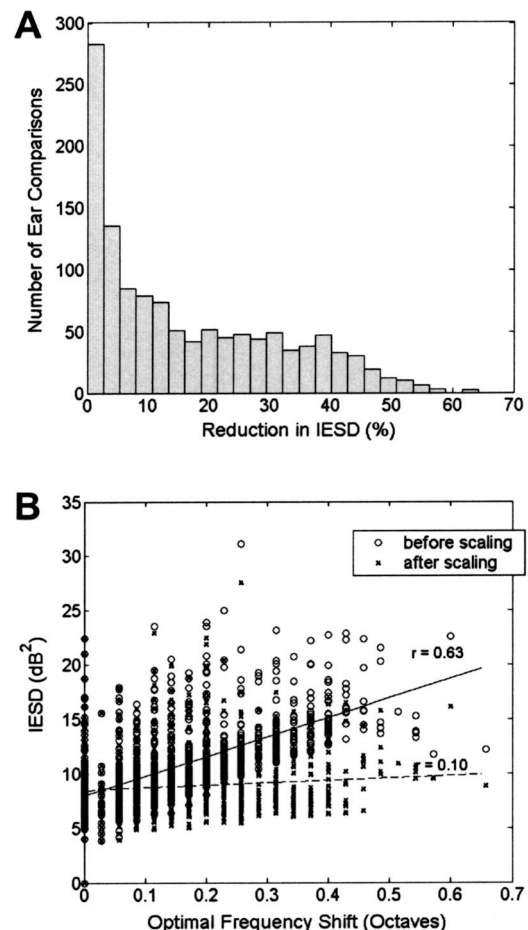


FIG. 3. Reduction in spectral difference achieved by frequency scaling. (A) Distribution of percentage IESD reductions achieved by optimal frequency scaling over all ear-pair combinations ($n=1225$). (B) IESD plotted as a function of OSF before and after scaling. The lines represent regression lines, and are labeled with the Pearson correlation coefficient.

scaling produced no appreciable IESD reduction. However, when assessing the effectiveness of optimal frequency scaling, one must bear in mind that the initial IESD can be very variable, and it may be unrealistic to expect large IESD reductions when the initial IESD is small. Like Middlebrooks (1999a), we therefore also plotted the IESD before and after frequency scaling against OSF, which is shown in Fig. 3(B). Prior to frequency scaling, IESD values (circles) correlated highly with OSF, whereas after optimal scaling this relationship disappears. This suggests that frequency scaling removes a component contributing to the IESD.

A. Relating optimal scale factors to physical dimensions

The spectral features of the DTFs are generally thought to be the consequence of resonance and interference patterns generated when the incident sound wave interacts with the external ear. Middlebrooks' study (1999a) showed that it may be possible to account, and compensate, for DTF differences in terms of differences in physical size. A larger outer ear presumably produces resonance and interference at longer wavelengths than a smaller but otherwise similarly shaped outer ear. If this is the case, then optimal frequency scaling should align spectral features in the DTFs and substantially reduce the IESD. Furthermore, the OSF required to align spectral features should correlate highly with the ratio of the sizes of the outer ear. Middlebrooks (1999a) showed that these expectations are indeed borne out for the human outer ear. In this study, we found that the same principle applies to the outer ear of the ferret. Ear and head measurements were obtained from digital photographs as described above. These measurements comprised pinna-cavity height [intertragal notch to the rim of the helix; Fig. 4(A), A to B], intertragal notch to antihelix (A to C), notch to crus of the helix (A to D), tragus to antitragus (E to F), pinna-cavity width (E to G), ear tip separation [Fig. 4(b) a to b] and ear base separation (c to d). Body weight was also included in the analysis, in the light of the findings of Xu *et al.* (2000), who observed a moderate correlation between body weight ratios and OSF in the cat.

Interear ratios, or, in the case of animal weight and head measurements, intersubject ratios were calculated for each of the physical measurements. Correlation coefficients for each of these ratios with OSF across the entire data set are shown in Table I. Pairwise correlation coefficients among the various physical size ratios are also shown. The ratio of pinna cavity heights [AB distance in Fig. 4(A)] correlated most strongly with OSF ($r = -0.838$), and all measured physical ratios correlated with OSF with a coefficient of $r = -0.511$. The fact that all the physical measurements correlated with OSF is perhaps unsurprising, since a high degree of covariance among these measurements is to be expected. In the following we consider pinna-cavity height as the single best predictor of optimal scale factor.

A linear regression of the pinna-cavity height ratio (ABratio) on OSF gave the formula

$$\text{predicted OSF (octaves)} = -0.6547 \cdot \text{ABratio} + 0.6734.$$

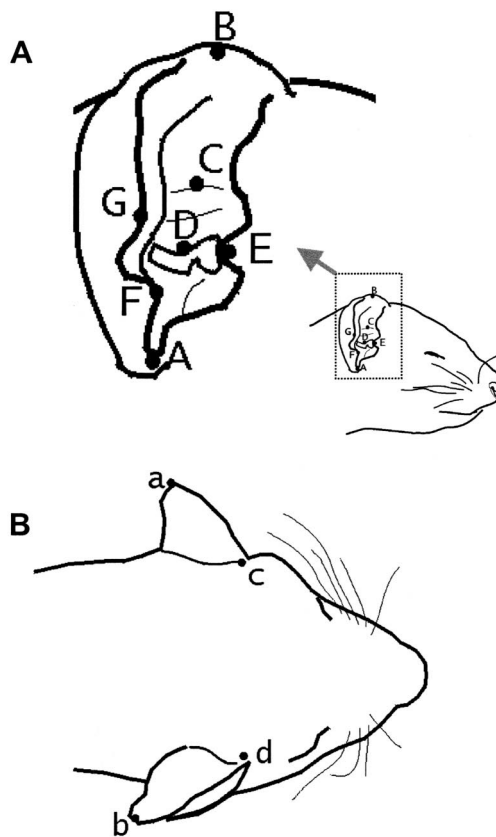


FIG. 4. Morphology of the ferret head and outer ear. (A) Points on the ferret ear used as measurement landmarks for size comparisons (A, intertragal notch; B, rim of helix; C, antihelix; D, crus of helix; E, tragus; F, antitragus; G, edge of pinna cavity). (B) Points used to measure ear separation. (a, b, ear tips; c, d, ear bases).

B. Frequency scaling of DTFs between juvenile and adult ferrets

The animals in this study covered a wide age range, and included a number of juveniles. A pertinent question to ask is whether scaling of DTFs in frequency between juveniles and adults is justifiable. If the growth of the outer ear during postnatal development incorporates significant changes in pinna shape, then these shape changes could produce qualitative differences between juvenile and adult DTFs which might be inadequately captured by simple frequency scaling. To address this question, we first had to define "mature" and "juvenile" age ranges in terms of outer ear and head growth. Figure 5 plots pinna-cavity height [Fig. 5(A)] and ear separation [Figs. 5(B),(C)] versus age. Exponential asymptotes of the form $y = y_o + a(1 - e^{-bx})$ were fitted to the plots, as an aid to estimating the ages by which pinna height and ear separation reach the adult range. For both pinna height and ear separation, the ferrets in this dataset aged >60 days appear to fall within the asymptotic ("mature") range. The juvenile (<60 -day-old) animals were subdivided into three groups: postnatal day (P) 50 (i.e., approximately 50 days old), P35 and P20. The P20 animals, being developmentally extremely immature, were not considered further.

To assess whether frequency scaling between juvenile and adult ferrets is justifiable, we first performed an exhaustive interear comparison for the four mature ($>P60$) animals

TABLE I. Pearson correlation coefficients. The upper numbers in each row are the correlation coefficient, while the small numbers beneath give the number of pairwise comparisons used to estimate the correlation coefficient. All correlations were significant at $p < 0.01$. (AB, AC, AD, EF, EG, ab, cd relate to the ear and head dimensions shown in Fig. 4 and are explained in the text. OSF-optimal scale factor.)

	AC	AD	EF	EG	ab	cd	Weight	OSF
AB	0.921 780	0.884 780	0.789 780	0.488 780	0.810 561	0.743 561	0.764 630	-0.838* 780
AC	1	0.863 780	0.779 780	0.406 780	0.755 561	0.693 561	0.723 630	-0.802 780
AD		1	0.702 780	0.445 780	0.671 561	0.594 561	0.660 630	-0.810 780
EF			1	0.448 780	0.725 561	0.649 561	0.550 630	-0.679 780
EG				1	0.287 561	0.196 561	0.296 630	-0.511 780
ab					1	0.966 561	0.753 561	-0.697 561
cd						1	0.726 561	-0.628 561
Weight							1	-0.611 1035

in our dataset, and determined the mean residual IESD after optimal frequency scaling. For the purpose of this analysis, same animal left ear–right ear comparisons were excluded. This was then compared to mean residual IESDs after optimal scaling for a large random selection of P50 versus adult ear comparisons, as well as for P35 versus adult ear comparisons. The mean residual IESDs obtained in this manner were very similar for the three groups: 10.21 dB² for the adult–adult comparison, 9.065 dB² for the adult–P50, and 9.889 dB² for the adult–P35. Optimum frequency scaling between adult and juvenile ears therefore produced residual IESDs no larger than those seen for pairs of adult ears. It therefore appears that frequency scaling of DTFs between juvenile and adult ears is equally justifiable as scaling between two adult ears, at least for ferrets aged 35 days or older, and for the range of sound frequencies considered. Given that the onset of hearing in ferrets is close to P35 (Moore and Hine, 1992), extending this analysis to even lower ages would be of limited interest.

C. Developmental changes in DTF gain

IESDs measure only differences in the shape of the DTF spectra. Differences in overall mean DTF gain do not enter

into the IESD calculation, as they cancel when IESD is calculated as the variance of the difference spectrum. However, the growth of the head and outer ears in development is bound to cause systematic changes in DTF gain. For sound sources near the pinna acoustic axis the DTF gain will typically increase with pinna size, since a larger pinna and concha produces a larger “sound funnel.” Sound source directions far from the acoustic axis will experience a decrease in overall gain due to increased “head shadows” produced by the larger head and neck. Figure 6(A) illustrates this pattern of developmental changes in overall DTF gain. This figure shows a map of differences in average DTF gains for a sample five adult and five (P35–P38) juvenile animals. An increase in gain for directions near the acoustic axis (about 50° azimuth, 10° elevation) of up to 8 or 9 dB in the adult animals is readily apparent, as is a decrease for directions below and behind the animal. This pattern of developmental changes in overall DTF gain is fairly stereotyped, with relatively little individual variation. This is illustrated in Fig. 6(B), which maps the distribution of standard deviations of gain differences for 25 adult–juvenile DTF gain comparisons. [Each of the five adult DTFs contributing to Fig. 6(A) was compared in turn with each of the five juveniles.]

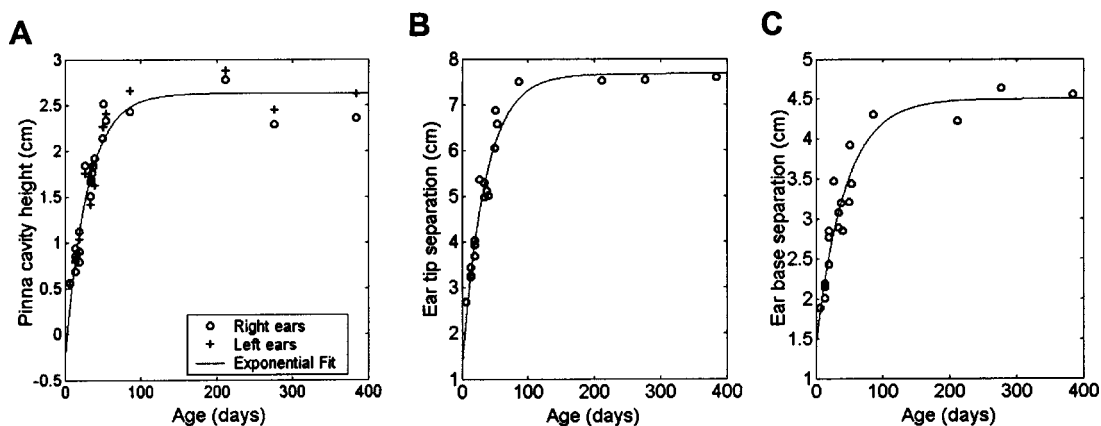


FIG. 5. Growth of the ferret outer ear. (A) Pinna-cavity height plotted against age ($n = 41$). (B) Ear tip separation plotted against age ($n = 18$). (C) Ear base separation plotted against age ($n = 18$). The data in all three panels were fitted with a curve of the form $y = y_0 + a(1 - e^{-bx})$.

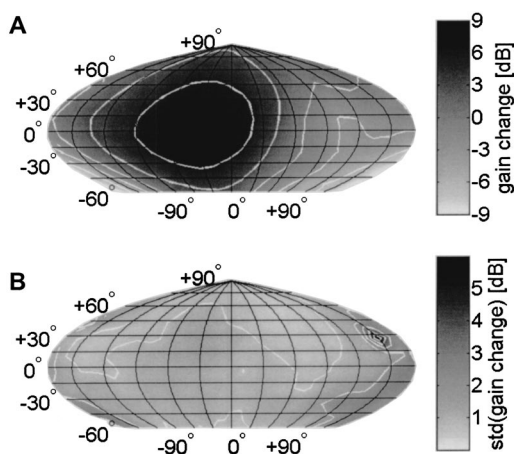


FIG. 6. (A) Change in overall DTF gain as an animal matures from onset of hearing to adulthood. The map shows the difference between the mean gain for five adult animals and the mean for five juveniles aged between P35 and P37. (B) Variability in gain changes. The map shows the standard deviation of 25 adult-juvenile DTF gain comparisons (each of the five adults from A compared with each of the juveniles in turn).

D. Analysis of ferret interaural time differences (ITDs)

ITDs provide an important localization cue, and the ability to model ITDs accurately without acoustical recordings from the ear canal could provide a useful research tool. We

extracted ITDs for each ferret from the HRTF measurements, as described above and compared them to predicted ITDs based on the “spherical head model” (SHM) (Woodworth and Schlosberg, 1962). The SHM predicts ITD values according to the formula $ITD = r/c[\theta + \sin(\theta)]$, where r is the radius of a spherical head, c is the speed of sound (here 340 m/s) and θ is the “lateral angle,” i.e., the angle of the sound source direction relative to the median plane. Figure 7(A) shows observed ITDs as a function of the lateral angle, as well as ITDs predicted by the “best-fit” SHM. The shape of a ferret’s head is rather poorly approximated by a sphere, and the value r obtained in the model fit is perhaps best thought of as an estimated “effective acoustic radius” of the head. Figure 7(B) compares this effective radius for the animals in our database against a corresponding physical measurement, namely half the ear base separation, i.e., half of $c-d$ distance ($cdDist$) as seen in Fig. 4(B). These values show a strong positive correlation (correlation coefficient 0.781, $N = 16$). A linear regression on these data, constrained to give a y intercept at zero, yielded the equation

$$r = 1.8355 \cdot (\frac{1}{2} \cdot cdDist).$$

Given this correlation, one could use the SHM to generate a set of approximate ITDs for a ferret using the ear-base separation multiplied by $1.8355 \cdot \frac{1}{2}$ or 0.91775 to obtain

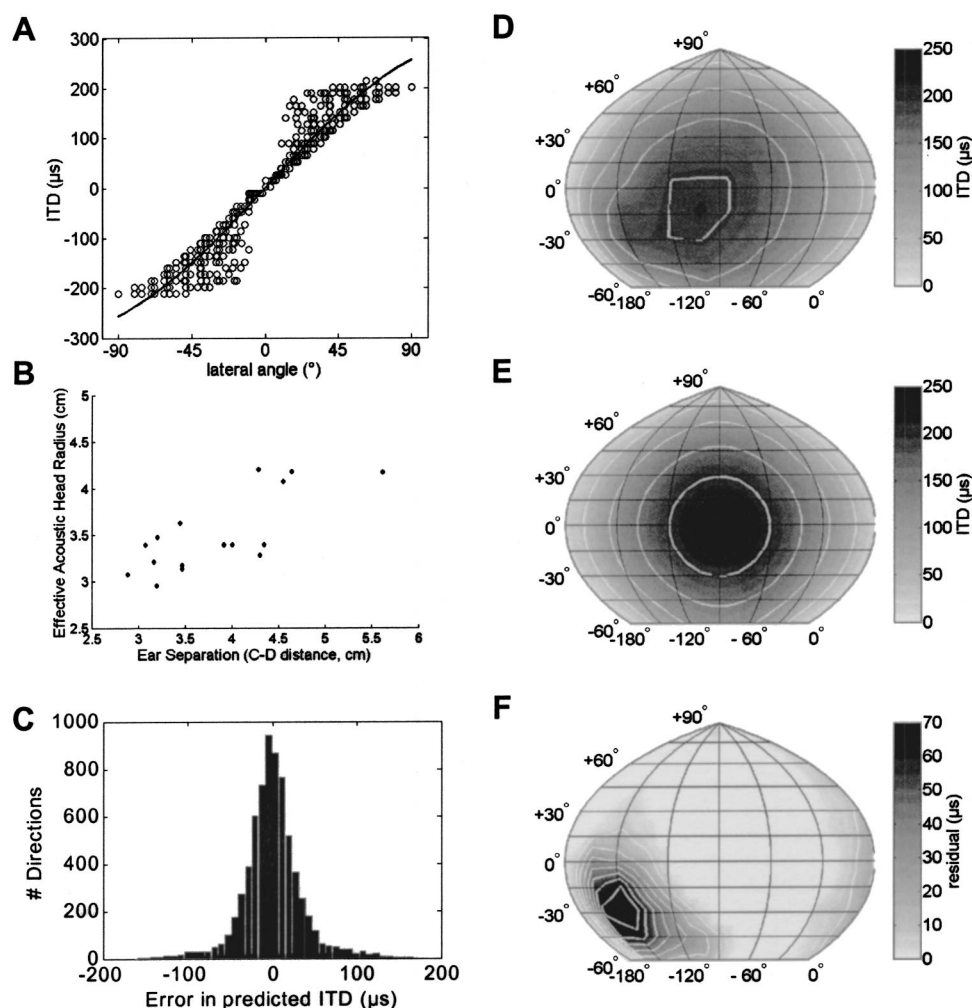


FIG. 7. Modeling ferret ITDs with the “spherical head model” (SHM). (A) The interaural delays measured for ferret F9966, plotted as a function of lateral angle. The thick black line shows ITD values predicted by the SHM. (B) Comparison between measured half-ear base separations [$c-d$ distance in Fig. 4(B)] and the “effective acoustic head radius” estimated by fitting the spherical head model to observed ITD data ($n = 16$). (C) Histogram showing the distribution of differences between observed ITDs and values predicted by the SHM, pooled over all animals and sound directions. (D) Representative ITD map for one animal (F9966). The ITD map is plotted using a quartic-authalic equal area map projection (MATLAB mapping toolbox, MathWorks, Natick, MA). The animal is facing toward the right of the figure. The white lines show iso-ITD contours at 50, 100, 150, and 200 μs . (E) Map of ITDs predicted for animal F9966 by the SHM. (F) Difference between observed map shown in (D) and prediction from SHM shown in (E).

the effective radius. However, Fig. 7(A) clearly shows that, for certain speaker positions, the recorded delays deviate quite markedly from the SHM. Particularly striking is the large spread of observed ITDs at lateral angles around $\pm 30^\circ$, which the SHM cannot capture. Similar characteristic deviations are apparent in the ITDs of all the ferrets in our database. Figure 7(C) summarizes the errors in the predictions made by the SHM, pooled for 16 ferrets from our database, in histogram form. To evaluate the performance of the model one requires an estimate of how large an error in the predicted ITD might be deemed negligible. Ferret just-noticeable ITD limens are unknown, but for humans they may be as little as $10 \mu\text{s}$ (Klump and Eady, 1956), while for the cat ITD discrimination thresholds are estimated to be greater than $20 \mu\text{s}$ but less than $50 \mu\text{s}$ (Masterton *et al.*, 1966). In comparison, for about 43% of the measured ITD values the absolute error in the predictions based on the SHM exceeds $20 \mu\text{s}$, and for 22% the absolute error exceeds $50 \mu\text{s}$. Thus, the predictions of the SHM would appear to be inadequate for a sizable proportion of the sound sources considered. Figures 7(D)–(F) illustrate the likely origin of some of the largest deviations from the model. Figure 7(D) shows a map of observed ITD directions for one animal (F9966). Only ITDs for directions to the right of the animal are shown. ITD functions are highly symmetric, so that the function for the left hemifield looks very similar. Figure 7(E) plots the ITDs predicted by the SHM, and Fig. 7(F) plots the difference between the observed ITD function and the spherical head prediction. In Fig. 7(F) it is readily apparent that the greatest discrepancies between the observed ITDs and the SHM occur for directions to the rear of the animal and below the midline. These are probably due to acoustic delays introduced by the animal's neck and body.

Since the SHM does not provide an adequate approximation of ferret ITD functions, it is natural to look for a better alternative. We found that von Mises basis functions (VMBFs) can provide a suitable alternative numerical model. The use of VMBFs was pioneered by Jenison and colleagues (1998) to develop analytic receptive field models of neurons in cat primary auditory cortex, but the technique can be used in principle to approximate any smooth function on spherical coordinates. A VMBF describes a “hill on a sphere.” It is uniquely defined by three parameters: two spherical coordinates α and β , which give the azimuth and elevation, respectively, of the hill's peak, while a “concentration parameter” κ determines its size and steepness. The VMBF for a spherical position with azimuth θ and elevation ϕ is then given by the equation

$$\text{VMBF}_{\alpha,\beta,\kappa}(\theta,\phi) = \exp(\kappa \cdot [\sin \phi \cdot \sin \beta \cdot \cos(\theta - \alpha) + \cos \phi \cdot \cos \beta]).$$

Spherical ITD functions can be approximated by the weighted sum of a sufficiently large number of suitably chosen VMBFs plus a constant offset o .

$$\text{ITDest}(\theta,\phi) = o + \sum_{i=1}^n w_i \cdot \text{VMBF}_i(\theta,\phi).$$

Increasing the number of VMBFs in the model generally allows a more accurate approximation of details inherent in the function to be modeled, but practical considerations make it desirable to keep the number of VMBFs in the model small.

To arrive at a model which could predict the ITDs for a given ferret on the basis of physical dimensions only, we used VMBFs to model a normalized average ferret ITD function (normITD). Individual ITD functions are then estimated by scaling normITD by a factor derived from the individual ear separation [$c-d$ distance “cdDist,” Fig. 4(B)] using linear regression. The model equation is therefore

$$\text{individualITD}(\theta,\phi) = \text{normITD}(\theta,\phi) \cdot (\zeta \cdot \text{cdDist} + \eta).$$

To estimate normITD we averaged the individual ITD functions after normalizing each function to the same mean absolute ITD value of $123.2 \mu\text{s}$. The averaged normalized ITD function obtained in this way, shown in Fig. 8(A), was then approximated using eight VMBFs. The parameters α_i , β_i , κ_i , w_i , and o of the VMBF model for normITD were estimated using a gradient descent method as described in Jenison *et al.* (1998) and are given in Table II. Figure 8(B) illustrates the VMBF approximation of normITD. Figure 8(C) charts the differences between the averaged ITD function and the VMBF model. [When comparing Fig. 8(C) with Fig. 7(F), please note that the gray scale in Fig. 8(C) is over four times finer.]

Finally, we estimated the parameters ζ and η relating individual ear separation to ITD scale factor by linear regression. For cdDist values in cm, this analysis resulted in estimates of $\zeta=0.1217$ and $\eta=0.5328$. To assess the fit of this model we once more calculated the residuals, i.e., the differences between observed and predicted ITD values for our entire database. The distribution of the residuals is shown in histogram form in Fig. 8(D). To facilitate a comparison with the SHM, the distribution of residuals for the SHM [distribution from Fig. 7(C)] is also shown (superimposed gray line). It is clear that the VMBF model performs considerably better than the SHM. The proportion of observed ITDs approximated with an error of greater than $20 \mu\text{s}$ has fallen from 43% with the SHM to 20% with the VMBF model, and the proportion of ITDs estimated with an error in excess of $50 \mu\text{s}$ has fallen from 22% to only 1.2%. The VMBF model therefore appears to generate quite adequate predictions of an animal's ITD from head size measurements for the large majority of sound directions.

IV. DISCUSSION

A. Frequency scaling of DTFs and potential applications

The primary aim of this investigation was to examine variability in individual DTFs of the acoustic periphery of ferrets, to assess whether optimal frequency scaling can account for this variability and how highly OSFs correlated with the physical dimensions of the animal. Our results indicate that in these respects the DTFs of ferrets behave much like those of humans (Middlebrooks, 1999a), although the IESD reductions achievable in ferrets appear to be slightly

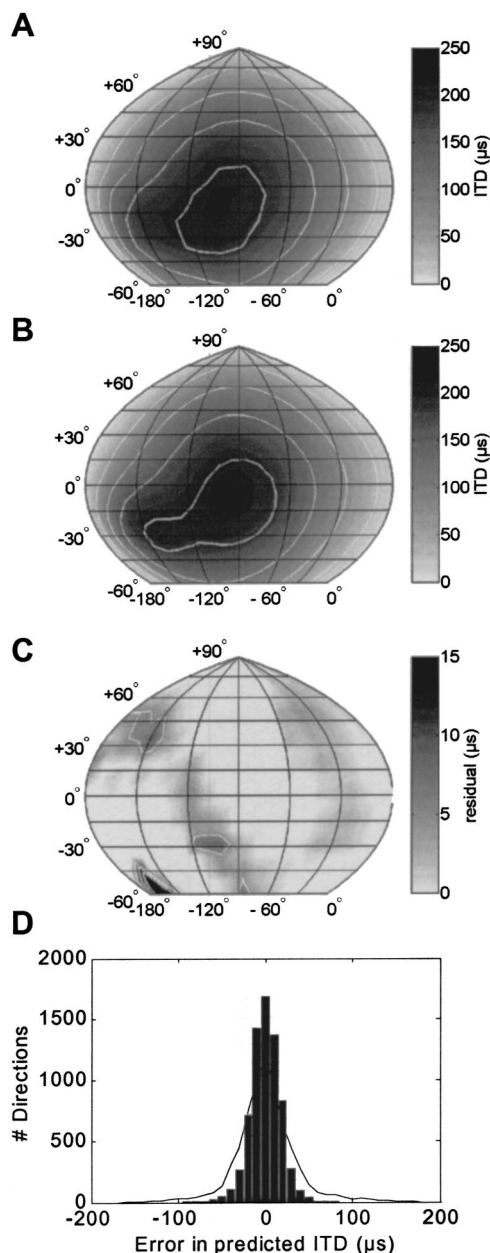


FIG. 8. Modeling ferret ITDs with von Mises spherical basis functions (VMBFs). (A) Map of averaged, normalized ferret ITD function. (B) Approximation of the averaged, normalized ITD function obtained using eight VMBFs. (C) Difference between average ITD function and von Mises fit. (D) Histogram of differences between observed ITDs and predictions generated by scaling the VMBF model. To facilitate comparisons with the spherical head model, a superimposed gray line shows the spherical head model error distribution from Fig. 7(C).

smaller than those reported for humans. For example, Middlebrooks (1999a) reported a reduction of $\geq 15.5\%$ for 50% of human ear comparisons, whereas the median IESD reduction for ferret ear comparisons here was only 11.76%, similar to a value reported for cats by Xu *et al.* (2000). As Fig. 1 illustrates, optimal frequency scaling tends to align prominent spectral features in the DTF between different animals. Nevertheless, in many cases sizable interear spectral differences remain after scaling.

A strong correlation ($r = -0.838$) was found between OSFs and pinna-cavity height, allowing OSFs to be estimated from pinna-cavity height by linear regression. For humans, Middlebrooks formulated a regression equation to predict OSFs from pinna-cavity height and head width, as these two variables had similarly high degrees of correlation with OSF, but relatively low mutual correlation. Overall, predicted human OSFs computed with the Middlebrooks equation correlated with actual OSF values at $r = 0.82$. This correlation coefficient is of similar magnitude to the value we obtained for the ferret using only pinna height as the explanatory variable in a simple linear regression ($r = -0.838$).

These findings indicate that it may be possible to generate a reasonably realistic virtual acoustic environment without the need for detailed acoustic measurements in each individual. In recent years, VAS stimulation has become an extremely valuable research tool for studying both the psychophysics (Bronkhorst and Houtgast, 1999; Pralong and Carlile, 1996; Shinn-Cunningham, 2001; Wightman and Kistler, 1989b) and the physiology (Brugge *et al.*, 1996; Delgutte *et al.*, 1999; Keller *et al.*, 1998; Schnupp *et al.*, 2001; Zella *et al.*, 2001) of spatial hearing. Virtual sound sources can be of sufficient fidelity as to be indistinguishable from genuine free-field sources (Hartmann and Wittenberg, 1996; Kulkarni and Colburn, 1998; Langendijk and Bronkhorst, 2000). But, to achieve a high fidelity it is usually necessary to compute the VAS stimuli from HRTFs measured for each individual. The use of nonindividual HRTFs leads to impaired localization accuracy (Middlebrooks, 1999b; Wenzel *et al.*, 1993) and can alter the spatial tuning properties of neurons in primary auditory cortex (Mrsic-Flogel *et al.*, 2001).

These studies illustrate that individual acoustical measurements are desirable for the generation of faithful VAS stimuli, but measuring HRTFs in each subject is a technically demanding and time-consuming process. The possibility of predicting acoustical properties of the auditory periphery on

TABLE II. Parameters of the von Mises basis function model of the averaged, normalized ferret ITD function. Each column gives the parameter values for one VMBF.

Param	VMBF 1	VMBF 2	VMBF 3	VMBF 4	VMBF 5	VMBF 6	VMBF 7	VMBF 8
α_i (deg)	154.60	-180.00	-55.79	-176.51	154.83	-71.24	-140.52	85.8311
β_i (deg)	-27.37	58.68	-30.82	-27.77	-27.41	-24.81	53.50	-17.00
κ_i	6.29	1.00	1.87	5.55	6.26	1.33	1.45	1.25
w_i (μs)	36.56	-94.50	-30.98	1.34	-38.08	100.00	59.63	-63.20
o (μs)	29.61							

the basis of a small number of easily obtained size measurements would therefore constitute an important advance. The present investigation, like the earlier studies of Middlebrooks (1999a, 1999b) and colleagues (Xu and Middlebrooks, 2000), were carried out with that goal in mind. However, while optimal frequency scaling can reduce the IESD, large intersubject differences often remain, and it is not obvious whether the achieved IESD reductions are sufficient to improve the fidelity of VAS stimuli. Middlebrooks (1999b) addressed this question in a psychophysical study. He found that, when subjects localized nonindividual VAS stimuli, the error rate and magnitude correlated with the IESD between the subject's own DTF and the foreign-ear DTF used to generate the VAS stimuli. He also reported a substantial improvement in localization accuracy when foreign-ear DTFs were optimally frequency scaled. Interestingly, the percent reduction in localization errors reported by Middlebrooks often exceeded the percent reduction in IESD achieved by frequency scaling. This type of psychophysical validation would be technically very difficult to replicate in nonhuman species. A validation by physiological means may be possible in the future, e.g., by recording from acoustically responsive neurons in the superior colliculus, whose spatial receptive fields are arranged to form a topographic map of auditory space (King and Hutchings, 1987; Middlebrooks and Knudsen, 1984; Palmer and King, 1982).

B. Modeling of ferret ITDs

Construction of VAS also requires the presentation of appropriate ITD cues and, where individual acoustical measurements are unavailable, these too would have to be estimated from physical dimensions. For the auditory periphery of the human, ITDs are usually quite well approximated by the "spherical head model" (Middlebrooks, 1999a; Woodworth and Schlosberg, 1962). For the ferret, however, we found that the SHM works poorly, deviating from measured values by $\geq 50 \mu\text{s}$ for almost one quarter of the sound directions tested. The SHM performed particularly poorly for posterior sound directions toward the back and below the interaural axis. We anticipate that ITD functions for many other animal species commonly used in auditory research may be similarly poorly approximated by the SHM. As a more accurate alternative we introduced von Mises spherical basis functions, which allowed us to capture the lack of radial symmetry in the ferret ITD function relative to the interaural axis. Linear scaling of a von Mises approximation to a standard ferret ITD function by a factor calculated from the ear separation enabled us to estimate ITDs with an error of less than $20 \mu\text{s}$ in 80% of cases and less than $50 \mu\text{s}$ in 98.8% of cases. Little data are available on ITD resolution in carnivores, but published estimates for the cat (Masterton *et al.*, 1966) suggest that this degree of accuracy is likely to be adequate.

C. Frequency scaling of HRTFs between juveniles and adults

The present study also indicated that DTF frequency scaling can to some extent map DTFs of juvenile ferrets onto

those of adult animals. Optimal scaling of DTFs across age groups, including infant–adult comparisons, reduced the remaining mean IESD to values no greater than those obtained with adult–adult scaling. This has potential implications for our understanding of the development of the auditory system. Behavioral (Hofman *et al.*, 1998; King *et al.*, 2000, 2001; Knudsen *et al.*, 1984; Shinn-Cunningham, 2001) and physiological (Gold and Knudsen, 2000; King *et al.*, 1988) studies have demonstrated that both the developing and mature auditory system is capable of adapting to experimentally altered localization cue values. During development, this plasticity is thought to enable the recalibration of the neural circuits involved in spatial hearing to continuous changes in the DTFs which accompany growth of the head and outer ears. If these changes are at least partly predictable, then the association between spectral cues and particular sound directions could be partly preprogrammed, and would not have to be entirely learned through experience.

ACKNOWLEDGMENTS

Supported by Defeating Deafness, the Dunhill Medical Trust, the Association of Physicians, and the Wellcome Trust.

- Batteau, D. W. (1967). "The role of the pinna in human localization," *Proc. R. Soc. London, Ser. B* **168**, 158–180.
- Blauert, J. (1969). "Sound localization in the median plane," *Acustica* **22**, 205–213.
- Bronkhorst, A. W., and Houtgast, T. (1999). "Auditory distance perception in rooms," *Nature (London)* **397**, 517–520.
- Brugge, J. F., Reale, R. A., and Hind, J. E. (1996). "The structure of spatial receptive fields of neurons in primary auditory cortex of the cat," *J. Neurosci.* **16**, 4420–4437.
- Butler, R. A. (1987). "An analysis of the monaural displacement of sound in space," *Percept. Psychophys.* **41**, 1–7.
- Delgutte, B., Joris, P. X., Litovsky, R. Y., and Yin, T. C. (1999). "Receptive fields and binaural interactions for virtual-space stimuli in the cat inferior colliculus," *J. Neurophysiol.* **81**, 2833–2851.
- Glasberg, B. R., and Moore, B. C. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138.
- Gold, J. I., and Knudsen, E. I. (2000). "Abnormal auditory experience induces frequency-specific adjustments in unit tuning for binaural localization cues in the optic tectum of juvenile owls," *J. Neurosci.* **20**, 862–877.
- Hartmann, W. M., and Wittenberg, A. (1996). "On the externalization of sound images," *J. Acoust. Soc. Am.* **99**, 3678–3688.
- Hofman, P. M., Van Riswick, J. G., and Van Opstal, A. J. (1998). "Relearning sound localization with new ears," *Nat. Neurosci.* **1**, 417–421.
- Jenison, R. L., Reale, R. A., Hind, J. E., and Brugge, J. F. (1998). "Modeling of auditory spatial receptive fields with spherical approximation functions," *J. Neurophysiol.* **80**, 2645–2656.
- Keller, C. H., Hartung, K., and Takahashi, T. T. (1998). "Head-related transfer functions of the barn owl: Measurement and neural responses," *Hear. Res.* **118**, 13–34.
- King, A. J., and Hutchings, M. E. (1987). "Spatial response properties of acoustically responsive neurons in the superior colliculus of the ferret: a map of auditory space," *J. Neurophysiol.* **57**, 596–624.
- King, A. J., Hutchings, M. E., Moore, D. R., and Blakemore, C. (1988). "Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus," *Nature (London)* **332**, 73–76.
- King, A. J., Parsons, C. H., and Moore, D. R. (2000). "Plasticity in the neural coding of auditory space in the mammalian brain," *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11821–11828.
- King, A. J., Schnupp, J. W. H., and Doubell, T. P. (2001). "The shape of ears to come: Dynamic coding of auditory space," *Trends Cog. Sci.* **5**, 261–270.
- Klumpp, R., and Eady, H. (1956). "Some measurements of interaural time difference thresholds," *J. Acoust. Soc. Am.* **28**, 859–864.

- Knudsen, E. I., Esterly, S. D., and Knudsen, P. F. (1984). "Monaural occlusion alters sound localization during a sensitive period in the barn owl," *J. Neurosci.* **4**, 1001–1011.
- Kulkarni, A., and Colburn, H. S. (1998). "Role of spectral detail in sound-source localization," *Nature (London)* **396**, 747–749.
- Langendijk, E. H., and Bronkhorst, A. W. (2000). "Fidelity of three-dimensional-sound reproduction using a virtual auditory display," *J. Acoust. Soc. Am.* **107**, 528–537.
- Masterton, B., Jane, J. A., and Diamond, I. T. (1966). "Role of brainstem auditory structures in sound localization. I. Trapezoid body, superior olive and lateral lemniscus," *J. Neurophysiol.* **30**, 341–359.
- Mehrgardt, S., and Mellert, V. (1977). "Transformation characteristics of the external human ear," *J. Acoust. Soc. Am.* **61**, 1567–1576.
- Middlebrooks, J. C. (1999a). "Individual differences in external-ear transfer functions reduced by scaling in frequency," *J. Acoust. Soc. Am.* **106**, 1480–1492.
- Middlebrooks, J. C. (1999b). "Virtual localization improved by scaling non-individualized external-ear transfer functions in frequency," *J. Acoust. Soc. Am.* **106**, 1493–1510.
- Middlebrooks, J. C., and Green, D. M. (1990). "Directional dependence of interaural envelope delays," *J. Acoust. Soc. Am.* **87**, 2149–2162.
- Middlebrooks, J. C., and Knudsen, E. I. (1984). "A neural code for auditory space in the cat's superior colliculus," *J. Neurosci.* **4**, 2621–2634.
- Moore, D. R., and Hine, J. E. (1992). "Rapid development of the auditory brainstem response threshold in individual ferrets," *Brain Res. Dev. Brain Res.* **66**, 229–235.
- Mrsic-Flogel, T. D., King, A. J., Jenison, R. L., and Schnupp, J. W. H. (2001). "Listening through different ears alters spatial response fields in ferret primary auditory cortex," *J. Neurophysiol.* **86**, 1043–1046.
- Palmer, A. R., and King, A. J. (1982). "The representation of auditory space in the mammalian superior colliculus," *Nature (London)* **299**, 248–249.
- Pralong, D., and Carlile, S. (1994). "Measuring the human head-related transfer functions: A novel method for the construction and calibration of a miniature in-ear recording system," *J. Acoust. Soc. Am.* **95**, 3435–3444.
- Pralong, D., and Carlile, S. (1996). "The role of individualized headphone calibration for the generation of high-fidelity virtual auditory space," *J. Acoust. Soc. Am.* **100**, 3785–3793.
- Schnupp, J. W. H., Mrsic-Flogel, T. D., and King, A. J. (2001). "Linear processing of spatial cues in primary auditory cortex," *Nature (London)* **414**, 200–204.
- Shaw, E. A. (1974). "Transformation of sound-pressure level from the free field to the eardrum in the horizontal plane," *J. Acoust. Soc. Am.* **56**, 1848–1861.
- Shinn-Cunningham, B. (2001). "Models of plasticity in spatial auditory processing," *Audiol. Neuro-Otol.* **6**, 187–191.
- 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.
- Wightman, F. L., and Kistler, D. J. (1989a). "Headphone simulation of free-field listening. I. Stimulus synthesis," *J. Acoust. Soc. Am.* **85**, 858–867.
- Wightman, F. L., and Kistler, D. J. (1989b). "Headphone simulation of free-field listening. II. Psychophysical validation," *J. Acoust. Soc. Am.* **85**, 868–878.
- Woodworth, R. S. (1938). *Experimental Psychology* (Holt, Rinehart and Winston, New York), pp. 349–361.
- Xu, L., and Middlebrooks, J. C. (2000). "Individual differences in external-ear transfer functions of cats," *J. Acoust. Soc. Am.* **107**, 1451–1459.
- Zella, J. C., Brugge, J. F., and Schnupp, J. W. H. (2001). "Passive eye displacement alters auditory spatial receptive fields of cat superior colliculus neurons," *Nat. Neurosci.* **4**, 1167–1169.
- 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.

Transient emission suppression tuning curve attributes in relation to psychoacoustic threshold^{a)}

Erika M. Zettner^{b)} and Richard C. Folsom

Department of Speech and Hearing Sciences, University of Washington, JG-15, Seattle, Washington 98195

(Received 13 July 2002; revised 15 January 2003; accepted 21 January 2003)

Ipsilateral suppression characteristics of transiently evoked otoacoustic emissions (TEOAEs) are described in relation to psychoacoustic threshold at 4000 Hz and the presence or absence of spontaneous otoacoustic emissions in 41 adults with normal hearing. TEOAE amplitudes were measured in response to 4000-Hz tonebursts presented in linear blocks at 40 and 50 dB SPL while puretone suppressors were introduced at a variety of frequencies and levels ipsilateral to and simultaneously with the tonebursts. Suppressors close to the toneburst frequency were most effective in decreasing the amplitude of the TEOAEs, while those more remote in frequency required significantly greater intensity for a similar amount of suppression. Consequently, characteristic tuning curve shapes were obtained. Tuning-curve tip levels were closely associated with the level of the toneburst and tip frequencies occurred at or above the toneburst frequency. Tuning-curve widths (Q_{10}), however, varied significantly across subjects with similar psychoacoustic thresholds in quiet determined by a two-alternative forced-choice method. The results suggest that a portion of that variability may be explained by the presence or absence of spontaneous otoacoustic emissions in an individual ear. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1560191]

PACS numbers: 43.64.Jb, 43.64.Kc [BLM]

I. INTRODUCTION

One of the main correlates of sensorineural hearing loss and, more specifically, outer hair cell dysfunction is a decrease in frequency resolution (Bonding, 1979; Liberman and Dodds, 1984; Moore, 1986). A loss of frequency resolution is thought to contribute to speech perception difficulties in those with hearing loss (Festen and Plomp, 1983; Stelmachowicz *et al.*, 1985). We have used a noninvasive measure of peripheral frequency resolution, the suppression of otoacoustic emissions (OAEs), to provide greater insight into the mechanisms for the auditory processes responsible for frequency resolution and sensitivity. Ipsilateral OAE suppression is a consequence of normal cochlear function and occurs at the preneural, mechanical stage of auditory transduction within the cochlea (Cooper and Rhode, 1992; Patuzzi *et al.*, 1984; Robles *et al.*, 1991; Ruggero *et al.*, 1992; Sellick and Russell, 1979). In addition, while suppression-tuning curves (STCs) are not identical, they have been compared to psychoacoustic measures of frequency resolution (Abdala *et al.*, 1996; Zwicker and Wesel, 1990).

Ipsilateral suppression of all types of OAEs in humans has been demonstrated [reviewed by Harris and Glatcke (1992)] with the majority of investigations reporting the suppression of distortion product OAEs (DPOAE). These re-

ports measured emission suppression tuning curves under experimental conditions that can be assumed to have changed psychoacoustic threshold, but which did not all report threshold. These studies used the temporary ototoxic effects of salicylates, furosemide and noise (Howard *et al.*, 2002; Long *et al.*, 1991; Martin *et al.*, 1998; Zettner *et al.*, 1996). Unexpectedly sharper tuning of DPOAE suppression was reported during salicylate toxicity (Zettner *et al.*, 1996). In addition, Martin *et al.* (1998) used furosemide in rabbits to temporarily induce hearing loss and reported a trend for sharper DPOAE STCs during periods when cochlear function was disrupted. Similarly, Howard *et al.* (2002) showed that DPOAE STCs were sharper during temporary, noise-induced, DPOAE-level reductions in rabbit ears. Under the effects of salicylates, spontaneous OAE (SOAE) STCs were shifted down in level, that is, lower suppression levels were needed to suppress the SOAE by the criterion amount but tuning was unchanged (Long *et al.*, 1991). Relevant work by Abdala (2001) showed sharper DPOAE tuning in premature infants as compared to term infants and adults. DPOAE bandpass filter measures ($f2/f1$ ratio sweeps) have also not succeeded in demonstrating changes in tuning following aspirin ingestion (Brown *et al.*, 1993). Since broader tuning would be expected with degraded or immature cochlear function, these studies argue for an alternate interpretation of STCs as estimations of cochlear tuning.

Both the production and the suppression of each type of OAE are thought to reflect the same underlying mechanism. However, qualitative differences exist between STCs of each OAE type. STCs of DPOAEs, for example, often exhibit “double-tipped” curves and other shape irregularities not characteristic of other measures of frequency resolution (Kummer *et al.*, 1995; Taschenberger and Manley, 1998).

^{a)}This article is based on a dissertation submitted by the first author to the Graduate School of the University of Washington in partial fulfillment of the requirements for the Doctor of Philosophy degree. Portions of this work were presented in “Transient emission suppression tuning curves as a function of psychoacoustic threshold,” poster presentation at the Association for Research in Otolaryngology Midwinter Research Meeting, St. Petersburg Beach, Florida, February 1999.

^{b)}Correspondence to: School of Hearing, Speech, and Language Sciences, Ohio University, Athens, OH 45701. Electronic mail: zettner@ohio.edu

There is increasing evidence that tuning curves obtained by suppressing DPOAEs represent the suppression of distortions generated at more than one site on the cochlear partition. This may explain at least some of the irregularities seen in DPOAE tuning curve shapes and highlights problems in estimating frequency resolution of a particular ear using DPOAE suppression (Siegel *et al.*, 2000; Stover *et al.*, 1996; Talmadge *et al.*, 1999). Suppression of transiently evoked OAEs (TEOAEs) does not exhibit this complexity in tuning curve shape and may argue for a localized site for TEOAE generation (Folsom *et al.*, 1995; Kemp and Chum, 1980; Tavartkildadze *et al.*, 1994; Wilson, 1980).

While a precise correspondence is not expected between OAE and psychoacoustic threshold or tuning measures, there are a number of associations that have been reported. First, SOAEs are often associated with frequencies of threshold minima (i.e., increased threshold sensitivity) (Burns *et al.*, 1984; Long and Tubis, 1988). Second, sharper psychoacoustic tuning curves have been reported not only at SOAE frequencies (Bright, 1985) but also a generalized sharpening effect of SOAEs on psychoacoustic tuning may occur at some frequencies (Micheyl and Collet, 1994). Third, Micheyl and Collet also showed that ears with lower level overall TEOAE responses demonstrated sharper tuning at 2000 Hz. And fourth, presence of SOAEs has been shown to increase amplitude and latency of TEOAEs as well as dominate the spectral characteristics of TEOAEs (Kulawiec and Orlando, 1995; Prieve and Falter, 1995). Given the evidence for an association between OAE, tuning, and threshold, there is sufficient rationale to investigate possible associations between TEOAEs STCs, psychoacoustic threshold, and SOAE.

The broad goal of this work was to provide a clearer understanding of normal inner ear processes that contribute to low psychoacoustic threshold and sharp frequency resolution. The assumptions underlying this goal were (1) that both threshold and frequency resolution are largely outer hair cell mediated characteristics resulting from frequency specific amplification to cochlear partition motion, (2) that OAEs are by-products of the amplification provided by the outer hair cells, and (3) that TEOAEs are generated by mechanisms within a confined region on the organ of Corti directly related to the frequency of the stimulus. The specific goal for this work was to investigate the relation between TEOAE STCs and psychoacoustic thresholds at 4000 Hz. TEOAE STCs in 41 adults with normal hearing were analyzed for tuning-curve width, tip level, and sideband slopes in relation to psychoacoustic threshold and TEOAE amplitude at the same frequency in the same group of ears. It was hypothesized that if TEOAE STCs accurately reflect properties of cochlear tuning, then broader curves were expected in ears with lower OAEs and poorer behavioral thresholds. Furthermore, since SOAEs probably contribute to certain aspects of TEOAEs, the potential for a generalized effect of synchronized time-domain averaged SOAEs (SSOAEs) on suppression tuning curve characteristics was also investigated (Kulawiec and Orlando, 1995; Prieve and Falter, 1995; Probst *et al.*, 1986).

II. METHODS

A. Participants

Forty-five participants between the ages of 20 and 48 years (mean 31.07 years) were recruited without regard for gender or presence of SSOAE. All were screened by interview and video otoscopy for negative audiologic and otologic history, clear ear canals, and normal appearing tympanic membranes. The test ear was randomly selected. The criterion response amplitude of >7 dB signal/noise using a 50 dB SPL toneburst (described below) was necessary to conduct suppression experiments in order to provide sufficient response level for amplitude manipulation. In two subjects this criterion was not met with the first randomly tested ear, but was met in the opposite ear. Four subjects did not have adequate response amplitude in either ear and were not included in further testing. Of the 41 subjects remaining, 5 were male and 36 were female. Twenty-nine left ears were tested and 45% of these had SSOAEs. Twelve right ears were tested and 50% of these had SSOAEs. Presence of SSOAEs was determined by visual inspection of responses in the frequency domain display of emissions equipment described below. Tympanometry was used to confirm that middle-ear pressure was between -25 and $+25$ daPa and compliance was between 0.4 and 1.5 ml at the time of testing. In addition, standard audiometric thresholds were obtained for each of the 41 subjects to confirm normal hearing thresholds (i.e., ≤ 25 dB HL). Test sessions were conducted with participants seated in a recliner in a sound-treated booth and lasted approximately 2 h.

B. Instrumentation and stimuli

A test frequency of 4000 Hz was chosen because it was hypothesized that greater variance in psychoacoustic threshold would be seen across subjects at the higher frequencies, thus increasing the probability that a relation between STC characteristics across threshold would be revealed. Further, this frequency was chosen to avoid noise below 1000 Hz often encountered when measuring TEOAEs and to avoid the microphone response roll-off beyond approximately 5500 Hz.

1. Psychoacoustic threshold

Psychoacoustic thresholds for a 4000-Hz tone were determined using a microcomputer interfaced with signal generation, modification instrumentation, and a participant response box. A 500-ms 4000-Hz tone was generated digitally by a Data Translation (DT2821) D/A board. The 4000-Hz tone was then high-pass filtered at 3400 Hz and low-pass filtered at 5000 Hz. It was cosine ramped with a rise-fall time of 16 ms and duration of 468 ms. This signal was digitally attenuated (Wilsons, PATT), passed through a dual filter in series (Kemo VBF8), amplified, and impedance matched. The signal was then delivered to the ear with an Etymotic ER-1 insert-phone fitted with a foam tip. Psychoacoustic threshold for this stimulus was obtained for each subject in a two-alternative forced-choice (2AFC) paradigm using a one-up, two-down adaptive procedure (Levitt, 1971) and halving the step size upon each reversal as delineated by parameter

estimation sequential testing (PEST) rules (Taylor and Creelman, 1967). Threshold was defined as the mean of ten reversals from the 2AFC procedure.

2. Otoacoustic emission measures

Three otoacoustic emission measures were recorded using the ILO92 system and TEOAE software (Version 6, Otodynamics-Ltd.). These included (1) SSOAEs, (2) toneburst evoked TEOAEs (described below), and (3) a TEOAE STC generated by suppressing the TEOAE response with puretones presented ipsilaterally and simultaneously with the toneburst. This toneburst was an eight-cycle 50 dB SPL toneburst centered at 4000 Hz that was generated using ILO92 software stimulus generation options (frequency = 4000 Hz, length in cycles = 8, amplitude = 100, multiplex = 1). The toneburst was presented in linear-stimulus blocks of four identical tonebursts in terms of phase and amplitude. The low-level linear toneburst series minimized the potential for stimulus artifact. In addition, confirmation of complete toneburst cancellation was determined by complete suppression of OAEs in each ear with at least one suppressor tone. Calibration of the toneburst was carried out in each ear prior to the acquisition of each averaged response of 260 stimulus blocks by using the “check-fit” procedure of the ILO92 testing sequence.

The two-receiver probe assembly supplied by Otodynamics for distortion-product emission recording was used in all conditions. Micropore tape applied over the probe assembly and pinna held the probe assembly in place for the duration of the session. The first receiver of the probe assembly delivered the toneburst. The second receiver delivered puretone suppressors used in tuning curve measurements. A cancellation procedure was necessary to eliminate the suppressor tone from the ear canal response before response averaging took place. Cancellation of the suppressor tone was achieved by intercepting the output of the microphone, electrically introducing a second puretone (cancellation tone), and adjusting its level and phase to be 180° out of phase with the suppressor tone as verified visually on an oscilloscope. The cancellation tone was generated by splitting the suppressor tone and routing it to a custom-built phase shifter. This intercepted and “suppressor-cancelled” response was then averaged. Adjustments in the cancellation were always necessary after suppressor frequency and amplitude changes. Fine adjustments in phase and intensity were occasionally necessary to maintain adequate cancellation during averaging.

A synthesizer/function generator (Hewlett-Packard, Model 3325A) generated the suppressor tones and a programmable attenuator (Tucker Davis, PA4) served to attenuate the tones. Suppressor tones were calibrated in each ear canal by using the spectrum analyzer option of the ILO92 DPOAE software to read the suppressor SPL at the plane of the probe. Frequency and intensity values were controlled manually. Suppressor frequencies were presented individually at $\frac{1}{4}$ -oct intervals from one octave below 4000 Hz to $\frac{1}{2}$ oct above and at $\frac{1}{8}$ -oct steps around the tip of the tuning curve. Five to seven of the frequencies listed in Table I were selected for each curve depending on the ability to define a

TABLE I. List of suppressor frequencies used to construct the tuning curves. Distance in octaves relative to the probe frequency are listed.

Octave interval	Frequency (Rounded to the nearest Hz)
–1	2000
–3/4	2378
–5/8	2594
–1/2	2828
–3/8	3084
–1/4	3363
–1/8	3668
Probe	4000
1/8	4362
1/4	4757
3/8	5187
1/2	5657

tuning curve tip in an individual ear. At a minimum, the goal was to adequately define the tip region, and then obtain points on both the low- and high-frequency sides that occurred at least 10 dB above the tip of the tuning curve. The tip of the tuning curve most often occurred within a $\frac{1}{4}$ oct above the toneburst frequency so that suppressors up to a $\frac{1}{2}$ oct above 4000 Hz adequately defined the high-frequency side of the tuning curves. Each suppressor frequency was presented in 5 or 10 dB SPL increments up to a maximum of 90 dB SPL. The stopping rules for increasing the level of the suppressor tone were (1) when the emission level was below the noise floor, (2) when the suppressor tone reached 90 dB SPL, or (3) if harmonic distortion from the suppressor occurred (visible in the averaged emission response window). Distortion occurred for low-frequency suppressors and was revealed as either noise (uncorrelated energy) or emission (correlated energy). Such responses were not included in further data analysis. The suppressor tone level in the ear canal was verified prior to and following each test session using the Otodynamics calibration-tones function.

C. Data analyses

Suppression tuning curves were plotted off-line. Response amplitudes were defined as the dB SPL of a ± 500 -Hz band around the toneburst frequency (i.e., 3500–4500 Hz) occurring from 2.5 to 20 ms regardless of whether a SSOAE occurred in this frequency region. Thus, energy recorded outside the 4000-Hz response region was not considered in the analysis but SSOAE energy within the response region could have contributed to the response. A SSOAE occurred within this region in four subjects (014, 038, 042, and 043) and all but one of these had response levels within one standard deviation of the mean. The OAE response level relative to an unsuppressed control response within the 1000-Hz-wide band was used to plot the rate (or growth) of suppression at each suppressor frequency [Fig. 1(a)]. Suppressor levels producing 1–10 dB of suppression for each frequency were plotted as a function of suppressor frequency, thus forming iso-suppression curves or tuning curves. A typical data set including 1 to 10 dB iso-suppression curves is shown in Fig. 1(b). Tuning curve tips were identified as the suppressor frequency requiring the lowest sound pressure level to suppress

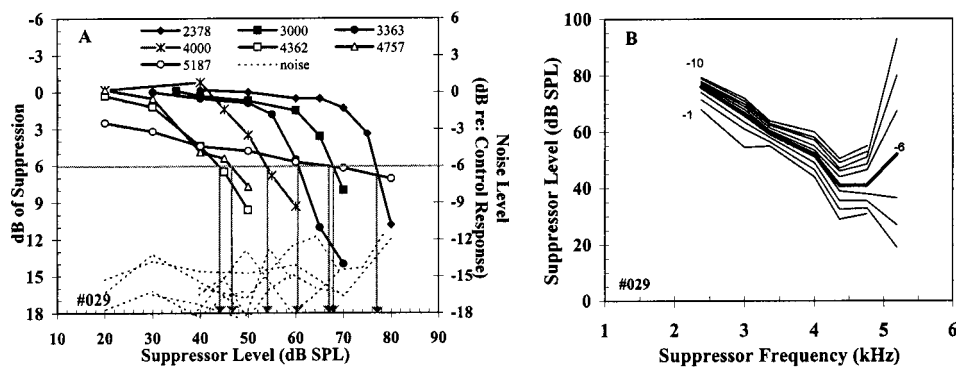


FIG. 1. (a) Sample set of seven suppression-rate curves from a representative subject (029). The parameter is suppressor frequency. Points are plotted relative to an unsuppressed control response. Suppressor tones below 4000 Hz are shown with filled symbols, suppressor tones above 4000 Hz with open symbols, and the suppressor tone at 4000 Hz with asterisks. Dashed lines indicate noise levels relative to the response level. A horizontal line intersects each curve at the criterion level reduction of 6 dB. Vertical arrows at each of these points of intersection indicate the suppressor level causing the criterion level reduction. These suppressor levels were used to construct the tuning curve [shown with a thick line in (b)]. (b) Typical iso-suppression curves (subject 029) from 1 to 10 dB of suppression of a transiently evoked otoacoustic emission using a 4000-Hz toneburst at 50 dB SPL and puretone suppressors. Each line represents an equal amount of suppression effect across suppressor frequency. The criterion of 6 dB (thick line) was used for measuring tuning curve width (Q_{10}) and side slopes. Suppression of 1 and 10 dB was not obtained at the highest frequency (5187 Hz) in this subject. Shown on a linear scale for easier visualization of individual curves. SSOAEs were present at 1306 Hz at -18.2 dB SPL and at 1880 Hz at -25.7 dB SPL.

the emission by the criterion of 6 dB. This criterion represented a halving of the TEOAE pressure and ensured that a significant amount of suppression was achieved. Further, this allows more direct comparison with neural tuning curves which have been frequently reported using a 6-dB criterion. Q_{10} values were calculated by dividing the bandwidth of the 6-dB isosuppression curve 10 dB up from the tip of the curve by the toneburst frequency (4000 Hz).

III. RESULTS

A. Psychoacoustic thresholds

Psychoacoustic thresholds ranged from -3.65 to 29.9 dB SPL (mean = 6.43 dB SPL ± 6.5). These correspond to thresholds better than 25 dB HL and within the range of normal hearing defined as better than 25 dB HL (or 34.5 dB SPL). These data are listed in Table II. A one-tailed t -test showed that the average psychoacoustic threshold for ears without SSOAEs (mean = 8.12 dB SPL) was not statistically different than ears with SSOAEs (mean = 4.83 dB SPL) ($t = 1.66$; $df = 39$; $p = 0.052$). In addition, there was no clear relationship between threshold and SSOAE frequency and

the two subjects whose SSOAEs were within 150 Hz of the test frequency (4000 Hz) demonstrated psychoacoustic thresholds within one standard deviation from the mean.

B. Otoacoustic emission responses

Representative transient OAE responses to the toneburst from three participants are shown in Fig. 2 and demonstrate the range of response levels and bandwidths observed in this study. In all cases, unsuppressed emission spectra were localized to the 4000-Hz frequency region. This is consistent with previously published data of toneburst evoked otoacoustic emissions (Norton and Neeley, 1987; Xu *et al.*, 1994). The average OAE response level was -1.76 dB SPL (± 4.9 dB) and ranged from -9.3 to 9.7 dB SPL (Table II). Twenty-one of the 41 participants (51%) had at least one SSOAE in the ears tested, and the remaining subjects had no measurable SSOAEs. Ears with SSOAEs demonstrated an average overall OAE response level of 1.05 dB SPL (± 4.1 dB), while ears without SSOAEs had an average OAE re-

TABLE II. Quantitative data (means and standard deviations) from the study are presented for all participants and then separately for subjects with synchronized spontaneous otoacoustic emissions (SSOAE) and without SSOAEs. Data are presented for the age of participants, psychoacoustic threshold (dB SPL) at 4000 Hz using a two-alternative forced-choice paradigm, otoacoustic emission response level (dB SPL) to a 50-dB SPL 4000-Hz toneburst, as well as transiently evoked otoacoustic emission suppression tuning curve characteristics [tuning curve sharpness (Q_{10}), tip level (dB SPL), and slope of the low- and high-frequency sides (dB/oct)].

		Age	Psychoacoustic threshold	OAE level	Q_{10}	Tip level	Low side slope	High side slope
All participants								
$N = 41$	mean	31.07	6.43	-1.76	4.44	49.22	-50.11	169.37
	sd	7.5	6.5	4.9	1.3	7.8	16.0	64.9
Without SSOAE								
$n = 20$	mean	30.05	8.12	-4.70	4.69	51.30	-51.66	163.58
	sd	7.7	7.8	4.0	1.1	6.8	13.1	69.1
With SSOAE								
$n = 21$	mean	32.15	4.83	1.05	4.19	47.21	-48.64	174.89
	sd	7.2	4.6	4.1	1.5	8.6	17.5	63.6

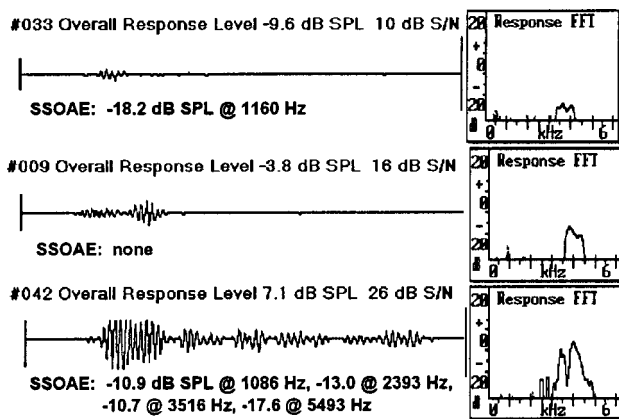


FIG. 2. Representative transient-evoked otoacoustic emission responses from three subjects to a 4000-Hz linear toneburst series presented at 50 dB SPL. Time waveforms (0- to 20-ms window) on the left side of the figure are shown with the corresponding FFT to the right of the waveform. Overall response levels and S/N ratios are indicated above each of the time waveforms. Each response obtained contained 260 averages. Responses shown were measured using an ILO88/92 system (Otodynamics Ltd). Frequencies and levels of synchronized spontaneous otoacoustic emissions (not shown in response) for these ears are given below the time waveforms.

sponse level of -4.70 dB SPL (± 4.0 dB). This difference was statistically significant ($t=4.56$; $df=39$; $p<0.01$, one-tailed t -test).

C. Suppression tuning curves

Forty-one STCs were analyzed quantitatively for Q_{10} , tip frequency, tip level, and slopes of the low- and high-frequency tuning curve sides. STCs shown in Fig. 3 were arbitrarily sorted by tip level (30–39.9, 40–49.9, 50–59.9, and 60–69.9 dB SPL). Twenty-nine (71%) of the curves had tip levels between 40 and 59 dB SPL. Tuning curves were always single-tipped which occurred at or up to $\frac{1}{4}$ oct above the toneburst frequency (4000–4757 Hz). The mean tip level was 49.22 dB SPL (± 7.8) with no significant difference between ears exhibiting SSOAEs and those without (Table II) ($t=1.67$; $df=39$; $p=0.098$). The average Q_{10} was 4.44 (± 1.3). The difference in Q_{10} between ears without SSOAEs was not statistically different from ears with SSOAEs ($t=1.24$; $df=39$; $p=0.11$). All curves exhibited steeper high-frequency than low-frequency slopes (mean = 169.37 dB/oct and -50.11 dB/oct, respectively). These data are shown in relation to Q_{10} in Fig. 4. Steeper slopes are indicated by more

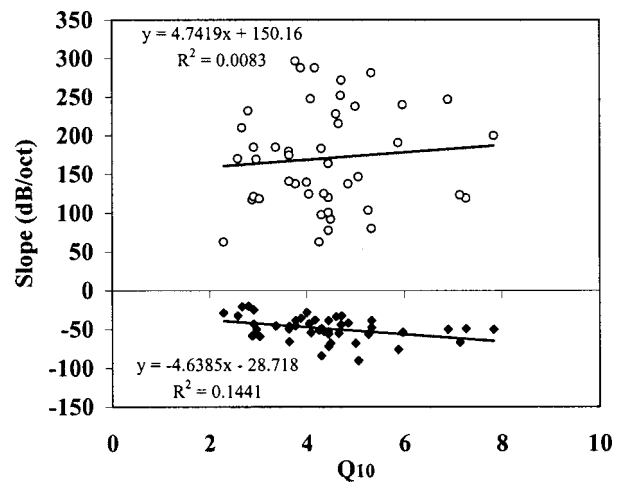


FIG. 4. Slopes (dB/oct) of both high-frequency sides (open circles) and low-frequency sides (filled diamonds) of all tuning curves shown in Fig. 3 are plotted as a function of Q_{10} . The linear regression of Q_{10} on the low frequency slope was significant ($p<0.01$), but not on the high frequency slope.

positive values on the high frequency side (open circles) and by more negative values on the low-frequency side (filled diamonds).

D. Suppression rate of growth

Representative iso-frequency or suppression-rate growth curves are shown in Fig. 1 from a typical participant (#029). Suppressors lower in frequency than the toneburst produced little or no effect at low levels but showed a substantial suppressive effect beyond a critical level, which varied depending on its frequency. Suppressors higher in frequency than the toneburst gradually increased in effectiveness (greater than 2 dB of suppression) beginning at very low suppressor levels (<20 dB SPL) and into higher levels (80 dB SPL). The slope of each frequency curve was estimated using a linear fit (R^2 was 0.90 or better) and represented the rate of growth of suppression. Only points greater than 2 dB of suppression were included to avoid inclusion of data points showing normal fluctuations of response amplitude. Figure 5 shows the estimated slope of the input/output curves as a function of frequency pooled across all subjects since there was no difference in the rate of suppression for ears with versus without SSOAEs at any frequency ($p>0.05$ for each frequency). The rate and variability of suppression across tuning curves was greatest at lower frequencies and gradu-

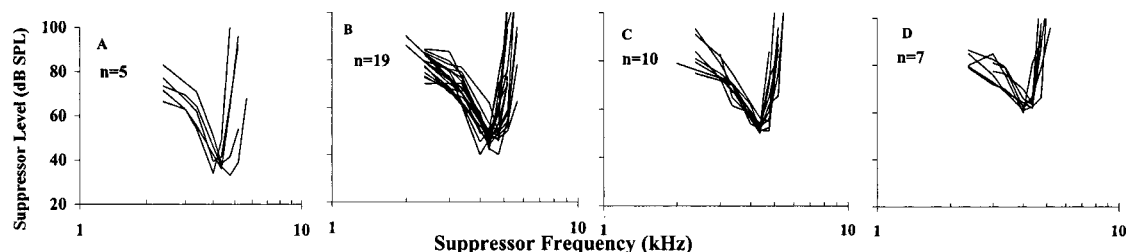


FIG. 3. Otoacoustic emission suppression tuning curves for all subjects ($n=41$) arbitrarily sorted by tip level. Panel (a) shows tuning curves (-6 dB iso-suppression curves) with tip levels between 30 and 39.5 dB SPL; (b) tip levels 40 to 49.5 dB SPL; (c) tip levels 50 to 59.5 dB SPL; and (d) tip levels 60 to 69.5 dB SPL. The toneburst was a 4000-Hz 50 dB SPL toneburst for all tuning curves.

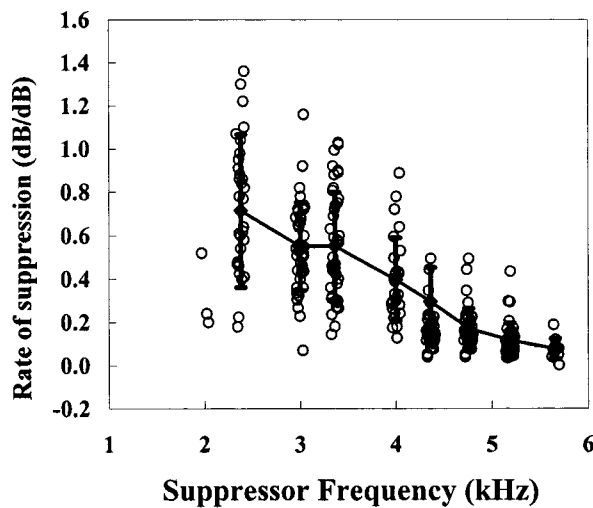


FIG. 5. Estimated rate (growth) of suppression (dB/dB) for each suppressor frequency used to construct the tuning curves. This was obtained by measuring the slope of a regression line fit to the iso-suppression curves (refer to Fig. 1) from 2 dB of suppression. The open circles indicate individual data and the thick vertical lines show standard deviations.

ally decreased in the higher frequencies. For example, at 2378 Hz the mean rate of suppression was 0.71 dB/dB (± 0.35) and at 5187 Hz the mean rate of suppression was 0.01 dB/dB (± 0.08).

E. Effect of toneburst level

The TEOAE suppression paradigm was tested using a 40 dB SPL toneburst in addition to the 50 dB SPL stimulus in 8 of the original 41 subjects in order to demonstrate the feasibility of using lower stimulus levels and to compare STCs from this study to other studies using DPOAE at multiple stimulus levels. This subgroup exhibiting >7 dB S/N OAE responses to a 40 dB SPL toneburst were retested using this reduced stimulus level. Figure 6 shows the tuning curves obtained at this lower level which are matched with a tuning curve using the higher level toneburst from the same ear. Asterisks indicate the level of each SSOAE identified. Overall, use of a 40 dB SPL toneburst did not result in significantly sharper tuning curves (mean $Q_{10} = 5.47 \pm 2.27$) than for a 50 dB SPL toneburst (mean $Q_{10} = 4.69 \pm 1.78$) ($t = 0.764$; $df = 14$; $p = 0.228$, one-tailed test). Sharper tuning was measured in three subjects, broader tuning in one subject and essentially no difference was measured in four subjects. However, tip level was significantly lower using the 40-dB SPL toneburst (mean 35.25 dB SPL) than the 50-dB SPL toneburst (mean 46.44 dB SPL) ($t = 3.03$; $df = 14$; $p < 0.01$). This represents a mean tip-level difference of 11.19 dB (± 5.6) and corresponded closely with the 10-dB difference in toneburst levels. In five cases, tip frequency was unchanged with use of the 40 dB SPL toneburst, in two cases it was lower, and one it was higher than when a 50 dB SPL toneburst was used.

F. Threshold versus OAE response level

Figure 7(a) shows the relation between emission response levels and psychoacoustic threshold for ears with ver-

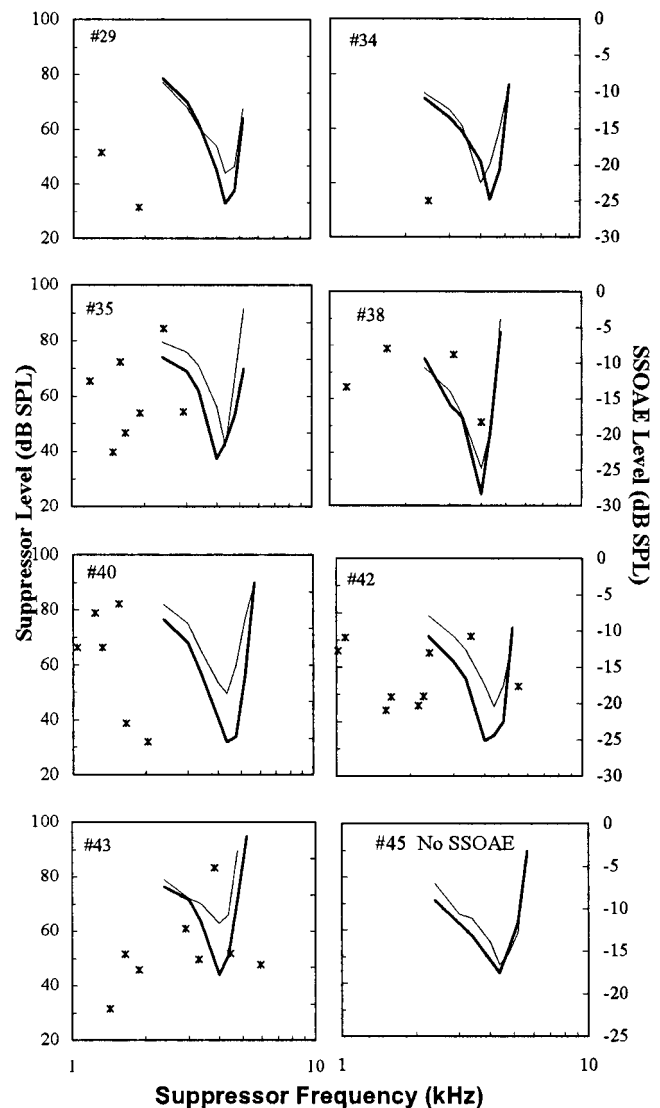


FIG. 6. TEOAE suppression tuning curves from eight subjects for whom both high- and low-toneburst levels were used. Thin lines indicate suppression tuning curves obtained using a 50 dB SPL 4000-Hz toneburst whereas thick lines indicate a 40 dB SPL toneburst was used. Asterisks indicate frequencies and corresponding amplitudes of synchronized spontaneous otoacoustic emissions. Secondary axis indicates SSOAE level in dB SPL.

sus without SSOAEs. Overall, higher toneburst evoked OAE response levels were obtained when SSOAEs were present (filled squares versus open diamonds) (average = 1.05 dB SPL versus -4.70 dB SPL). In both groups the highest OAE response levels were obtained when psychoacoustic threshold was low and levels decreased with increased thresholds. This relation was significant in both groups ($r = -0.640$, $df = 20$, $p < 0.01$ with SSOAE; $r = -0.490$, $df = 19$, $p < 0.05$ without SSOAE).

G. Q_{10} versus threshold

Figure 7(b) shows the relation of tuning curve sharpness (Q_{10}) as a function of psychoacoustic threshold in ears with and without SSOAEs. A wide range of Q_{10} values was measured across subjects, ranging from 2.29 to 7.84. In addition, there was substantial overlap in Q_{10} values across the two groups, particularly at lower thresholds. Only a few ears

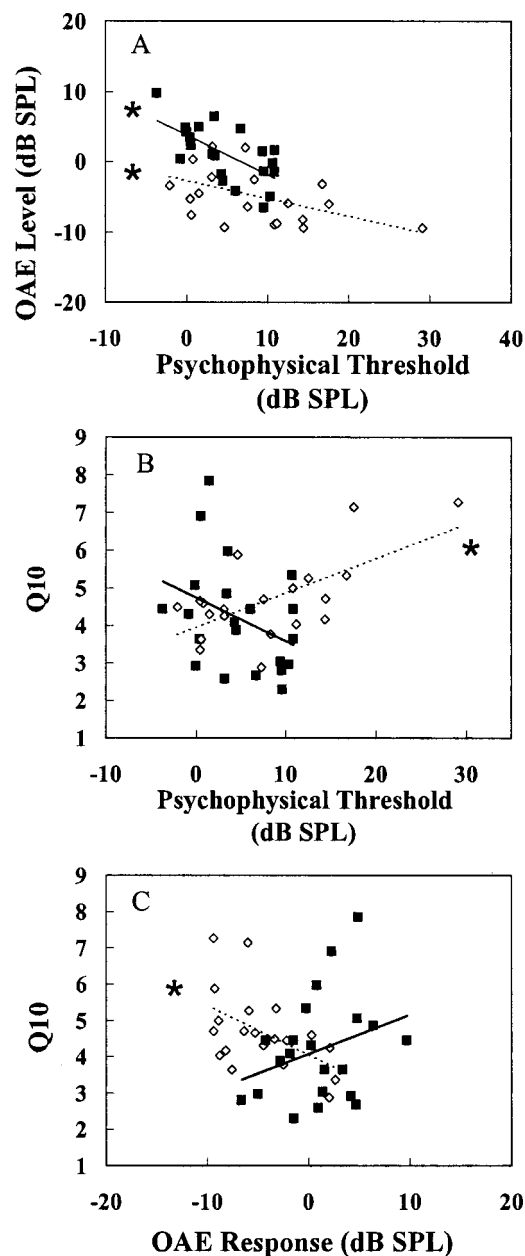


FIG. 7. (a) Transient OAE response levels to a 50 dB SPL toneburst as a function of psychoacoustic threshold for a 4000-Hz tone in ears with synchronized spontaneous otoacoustic emissions (SSOAE) (filled squares) and ears without SSOAEs (open diamonds). (b) Tuning-curve sharpness (Q_{10}) as a function of psychoacoustic threshold for a 4000-Hz tone obtained using a two-alternative forced-choice paradigm. Ears with SSOAEs (filled squares). Ears without SSOAEs (open diamonds). (c) OAE response levels to a 50 dB SPL 4000-Hz toneburst as a function of Q_{10} in ears with SSOAEs (filled squares) and ears without SSOAEs (open diamonds). Solid trend lines shown for subjects with SSOAEs, dotted trends lines shown for subjects without SSOAEs. Asterisks indicate significant regression analyses.

were tested with thresholds above approximately 10 dB SPL, and none of these had SSOAEs. Q_{10} was significantly correlated with threshold from the group without SSOAE ($r = 0.645$, $df = 19$, $p < 0.01$). Statistical significance was maintained with removal of the outlying data point at 29 dB SPL ($r = 0.460$, $df = 18$, $p < 0.05$). An opposite trend was observed in ears with SSOAEs, that is higher Q_{10} was observed with lower thresholds ($r = -0.031$; $df = 17$; $p = 0.904$).

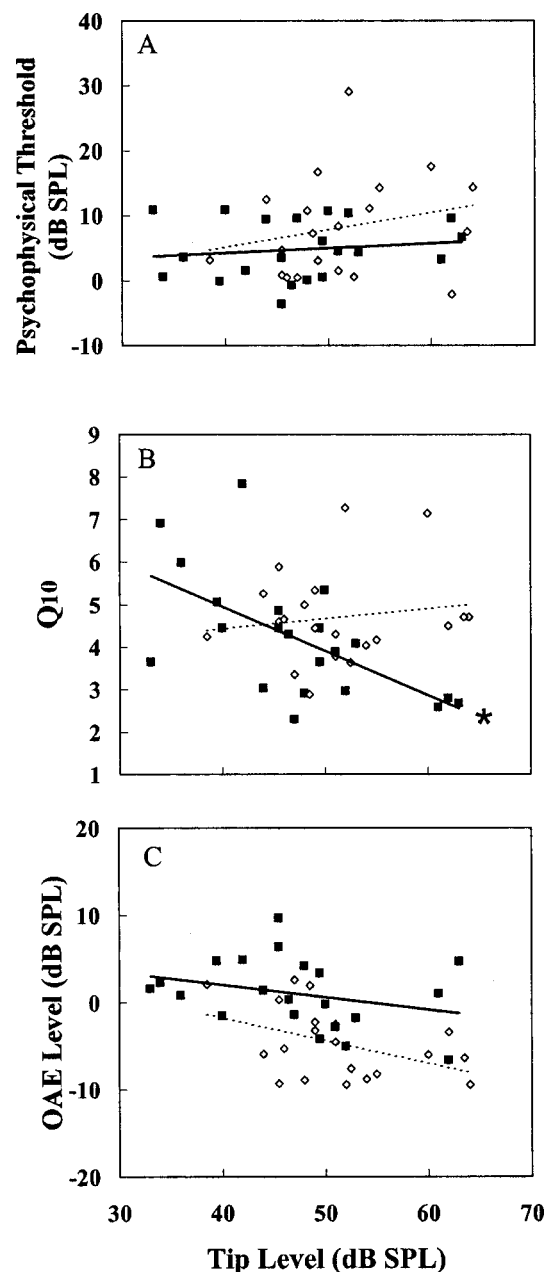


FIG. 8. Suppression tuning curve tip level in relation to (a) psychoacoustic threshold, (b) Q_{10} , and (c) OAE response level. Data from ears with synchronized spontaneous otoacoustic emissions (SSOAE) are shown with filled squares; ears without SSOAEs are shown with open diamonds. Asterisks indicate where significant regression analyses were obtained.

H. Q_{10} versus OAE response level

Figure 7(c) shows that the pattern for Q_{10} versus OAE response level also seemed to depend on presence or absence of SSOAEs. Without SSOAEs, Q_{10} values were significantly higher when OAE response levels were low ($r = -0.503$, $df = 19$; $p < 0.05$). This contrasted with ears with SSOAEs where low Q_{10} values tended to occur with low OAE response levels ($r = 0.302$; $df = 20$; $p = 0.184$).

I. Tip level

Tuning curve tip level was examined in relation to three parameters: psychoacoustic threshold, Q_{10} , and OAE re-

sponse level (Fig. 8). Tip level and psychoacoustic threshold were not related in a meaningful way in either group [Fig. 8(a)]. However, tip level was significantly related to Q_{10} in ears with SSOAEs, such that lower tip levels were associated with higher Q_{10} values ($r = -0.599$; $df = 20$; $p < 0.01$) [Fig. 8(b)]. A clear relationship between Q_{10} and tip level in ears without SSOAEs was not found. Lastly, the data from ears without SSOAEs suggest that lower tip levels occurred when OAE response levels were high [Fig. 8(c)] ($r = -0.443$; $df = 19$; $p = 0.051$). No relation was seen between OAE response level and tip levels in ears with SSOAEs ($r = -0.297$; $df = 20$; $p = 0.192$).

IV. DISCUSSION

A. General characteristics of TEOAE STCs

Tuning curves obtained by ipsilateral simultaneous suppression of toneburst evoked OAE were qualitatively similar to previously published neural and psychoacoustic tuning curves (e.g., Liberman and Dodds, 1984; Vogten, 1974). TEOAE STCs were asymmetrical, V-shaped, and in many cases, exhibited a low-frequency tail. Characteristics of the STCs in this study are comparable to human TEOAE STCs previously published and are similar to STCs using other types of emissions as well (Kemp and Chum, 1980; Wilson, 1980; Kummer *et al.*, 1995; Abdala *et al.*, 1996). In addition, the shapes of the suppression rate curves were consistent with previously published TEOAE suppression experiments (e.g., Zwicker and Wesel, 1990), with DPOAE suppression experiments (e.g., Abdala, 1998; Kummer *et al.*, 1995) and SOAE suppression experiments (e.g., Rabinowitz and Widin, 1984). Overall, however, the rate of suppression at each frequency appeared lower (Fig. 5) compared to previously published data especially below the toneburst frequency (Abdala, 1998; Kummer *et al.*, 1995).

The average Q_{10} for the 6-dB isosuppression tuning curves in this study was $4.44 (\pm 1.3)$ and ranged from 2.88 to 7.84. These values compare well with previously published works in humans at similar frequencies and levels to those used in the current study. Folsom *et al.* (1995) reported an average Q_{10} in five adults of 3.52 (range 2.56–4.44) for TEOAE STCs using a nonlinear 70-dB SPL toneburst at 4000 Hz. There are a few studies that offer Q_{10} measures of DPOAE STCs in humans using frequencies similar to those in the current study. Specifically, Kummer *et al.* (1995) reported Q_{10} values between about 3.0 and 5.3 at 4000 Hz (f_1 and f_2 levels of 55 and 40 dB SPL, respectively). Abdala (1998) and Abdala *et al.* (1996) reported Q_{10} values between approximately 2.5 and 4.25 at 3000 Hz and between 3.0 and 4.5 at 6000 Hz (65 and 50 dB SPL for f_1 and f_2 stimulus levels). And, Harris *et al.* (1992) obtained an average Q_{10} of 2.97 using slightly higher stimulus levels in the 4000-Hz region. Studies using laboratory animals, including those of the barn owl and rabbit, have also shown Q_{10} values in the same range at this frequency (Martin *et al.*, 1998; Taschenberger *et al.*, 1998).

Seventy-one percent (71%) of the TEOAE STCs using a 50 dB SPL stimulus had tip levels within 10 dB of the stimulus level. A decrease in the stimulus level by 10 dB was

matched by an average downward shift in tip level by 11.19 dB without a significant change in tuning (Fig. 6). This contrasts with findings from Kummer *et al.* (1995) and Harris *et al.* (1992) who reported a decrease in DPOAE tuning curve width with lower primary tone levels. In the current study, tip frequencies of the 6-dB isosuppression curves occurred at and sometimes as much as $\frac{1}{4}$ oct above the stimulus frequency (4000 Hz). And, increasing the suppression criterion from 1 to 10 typically resulted in a downward frequency shift in the tuning curve tip [see Fig. 1(b) for an example], a trend that Kummer *et al.* (1995) has also reported. This shift down in tip frequency with increasing suppression criterion has been noted in basilar membrane vibration measures as well (Rhode, 1978; Sellick *et al.*, 1982).

B. Contributions of SSOAEs to TEOAE STCs

One objective of this study was to examine effects of SOAE on TEOAE STC attributes since SOAEs are known to affect amplitude, spectral, and latency characteristics of TEOAEs (Kulawiec and Orlando, 1995; Prieve and Falter, 1995). Furthermore, STCs obtained by suppressing DPOAEs and SOAEs often exhibit a multi peaked and irregular fine structure (Bargones and Burns, 1988; Kummer *et al.*, 1995; Zettner *et al.*, 1996). This fine structure may be the result of complex interactions between the suppressor tones and emissions. For example, multiple SOAEs in an ear have been shown to interact with each other by generating additional DPOAE in addition to demonstrating reciprocal suppression (Burns *et al.*, 1984). In addition, externally presented tones can release SOAEs from suppression (Rabinowitz and Widin, 1984). Whether presence of additional SOAEs specifically influence characteristics of DPOAE STCs is not clear. While Kummer *et al.* (1995) did not find a consistent effect of SOAEs on STC shape, they did report a number of shape irregularities when SOAEs were present. And, Abdala *et al.* (1996) reported a greater proportion of multi peaked DPOAE STC in infant ears, which can exhibit greater numbers of SOAE (Burns *et al.*, 1992). In contrast, Harris *et al.* (1992) reported that presence of SOAE did not appear to affect rate of suppression nor characteristics of the tuning curves. The TEOAE STCs in the current study did not demonstrate shape irregularities often demonstrated by DPOAE and SOAE generated STCs.

Though TEOAE STCs obtained in this study appeared unaltered in shape by SSOAEs, their presence did influence the relation of Q_{10} with OAE response. Ears without SSOAEs showed significantly greater tuning when OAE response level was low. This was in contrast to ears with SSOAEs tended to exhibit greater STC sharpness when OAE response level was high. This suggests that presence of any SSOAE may influence tuning curve sharpness. Since finer tuning and high-level emissions are both expected to be associated with low thresholds, it was reasonable to expect that high level OAEs would be associated with sharper suppression tuning, regardless of presence or absence of SOAEs. Thus, the results found for subjects without SOAEs are unexpected and an adequate explanation for this finding is lacking.

An error in categorizing subjects on the basis of presence or absence of SSOAE may have occurred for subjects who did not have readily identifiable SSOAE. It is possible that some of these subjects had low level SSOAE below the noise floor of the recording. The likelihood for this kind of misclassification increased in subjects with lower psychoacoustic thresholds as this could have indicated a minimum in their threshold microstructure often associated with SOAE (Long and Tubis, 1988).

C. Relation of psychoacoustic threshold to TEOAE STC parameters

Psychophysical measures of frequency resolution have shown a wide range of frequency resolution characteristics across ears with similar thresholds (Tyler, 1986; Bergman *et al.*, 1992). As a peripheral response, the OAE STC excludes processing across much of the auditory system that could contribute to the variability of psychophysical measures (reviewed by Moore, 1986). Nevertheless, a wide range in tuning was also observed for the subjects in this study with relatively similar psychoacoustic thresholds. Unlike neural or psychophysical methods, OAE STCs are shaped by individual output impedances of the middle ear. However, as output impedance functions are broadly tuned, it is not probable that this was a significant source of variability on the TEOAE STC width values observed in this study.

The methodology used for calibration of the suppressor tones was also a possible source for variability in tuning curve characteristics seen in this study. Since the stimuli were calibrated in each ear at the plane of the probe rather than at the tympanic membrane, individual frequency-dependent ear canal properties would have resulted in suppressors at different frequencies reaching the cochlea at different levels. While the same error would have existed for both the 4000-Hz tone burst and the suppressors at and close to 4000 Hz, the amount of error at other frequencies may have varied across subjects and contributed to measurable differences in suppression tuning.

It was also reasonable to speculate that the participants with similar thresholds in this study did not represent a homogeneous group. West and Evans (1990), for example, showed that psychophysical tuning curves are sometimes broader before measurable changes in behavioral thresholds occur. Thus, some of the Q_{10} variability across subjects with similar threshold may have indicated a real difference in frequency resolution characteristics of individual ears.

Despite evidence for a common active cochlear process underlying evoked otoacoustic emission generation and auditory frequency selectivity, little work has been published that directly tested this association. Micheyl and Collet (1994) investigated the relation between frequency resolution and the level of evoked OAE in subjects without SOAEs and found sharper psychoacoustic tuning curves (at 2000 Hz) in ears with lower OAE response levels. This finding is similar to results in the present study at 4000 Hz. That is, in ears without SSOAEs, sharper STCs were obtained when OAE response level to the toneburst was low [Fig. 7(c)]. This contradicts with the notion that large OAEs and sharper tuning result from a more active cochlear amplifier and that smaller

OAEs result when there is decreased activity. This may reflect the same process influencing the paradoxically sharper tuning that some researchers have reported during salicylate ototoxicity, temporary noise-induced hearing loss, and application of furosemide (Howard *et al.*, 2002; Martin *et al.*, 1998; Zettner *et al.*, 1996).

D. Site of suppressive effects

There is evidence to suggest that TEOAE are composed of intermodulation distortion energy generated over a widely distributed area of the cochlear partition. Withnell *et al.* (2000), for example, found that damage to the basal region of the cochlea produced changes in the amplitude and latency characteristics of low frequency TEOAE. Temporal patterns of TEOAE from normal and noise-induced impaired human cochlea studied by Avan *et al.* (1993) also indicate that energy may be generated by regions as far as 1 to 1.5 oct from its tuned location on the basilar membrane and contribute to the overall response. Therefore, intermodulation distortion may add to TEOAE energy generated locally. The extent to which one component dominates the emitted energy may be determined by individual physical cochlear characteristics. On the other hand, a puretone suppressor stimulates a discrete location and its suppressive effects are local. While TEOAE energy may come from a distributed region, the suppression-tuning curve reflects a suppression effect (or synchronization effect as described by Neumann *et al.*, 1997) occurring at a characteristic frequency. Therefore, even though TEOAE may consist of energy generated over a widely distributed area of the cochlear partition, STC can be interpreted as indicating tuning at a localized place on the partition.

V. CONCLUSIONS

OAE suppression seems to provide a direct and noninvasive means for studying initial auditory stages of frequency resolution otherwise relatively inaccessible in humans. Many of the TEOAE STC characteristics described in this study were qualitatively similar to tuning curves obtained using DPOAEs and with other auditory responses. Yet, findings in this study and DPOAE STCs from other studies contradict with expected results. That is, the quantitative characteristics were not always related to psychoacoustic threshold and OAE response level in a clear and meaningful way. The presence of SSOAEs often existed with results distinct from the results from ears without SSOAEs. Specifically, in ears without SSOAEs, poorer psychoacoustic thresholds were associated with sharper tuning, whereas poorer thresholds were associated with broader tuning when SSOAEs were present. Further, the relation between Q_{10} and OAE response level and tuning curve tip level also depended on the presence or absence of SSOAEs. At least in ears without SSOAEs, frequency selectivity appeared better in ears with lower OAE response levels. These issues must be resolved in order to interpret the information gained from the suppression of otoacoustic emissions as measures of peripheral frequency resolution.

ACKNOWLEDGMENTS

This research was supported by NIDCD Grant No. PO1-DC00520 and NIDCD Training Grant No. T32-DC00033-06. The authors gratefully acknowledge two anonymous reviewers for their careful reading and useful suggestions. The authors would also like to thank Li Xu, Ph.D., for graphical assistance, numerous discussions, and readings of early versions of this document.

- Abdala, C. (1998). "A development study of distortion product otoacoustic emission (2f1-f2) suppression in humans," *Hear. Res.* **121**, 125–138.
- Abdala, C. (2001). "Maturation of the human cochlear amplifier: Distortion product otoacoustic emission suppression tuning curves recorded at low and high primary tone levels," *J. Acoust. Soc. Am.* **110**, 1465–1476.
- Abdala, C., Sininger, Y. S., Ekelid, M., and Zeng, F.-G. (1996). "Distortion product otoacoustic emission suppression tuning curves in human adults and neonates," *Hear. Res.* **98**, 38–53.
- Avan, P., Bonfils, P., Loth, D., and Wit, H. P. (1993). "Temporal patterns of transient-evoked otoacoustic emissions in normal and impaired cochleae," *Hear. Res.* **70**, 109–120.
- Bargones, J. Y., and Burns, E. M. (1988). "Suppression tuning curves for spontaneous otoacoustic emissions in infants and adults," *J. Acoust. Soc. Am.* **83**(5), 1809–1816.
- Bergman, M., Najenson, T., Korn, C., Harel, N., Erenthal, P., and Sachartov, E. (1992). "Frequency selectivity as a potential measure of noise damage susceptibility," *Br. J. Audiol.* **26**, 15–22.
- Bonding, P. (1979). "Critical bandwidth in patients with a hearing loss induced by salicylates," *Audiology* **18**, 133–144.
- Bright, K. E. (1985). "Microstructure audiograms and psychophysical tuning curves from subjects with spontaneous otoacoustic emissions," unpublished doctoral dissertation, University of Arizona, Tucson, AZ.
- Brown, A. M., Williams, D. M., and Gaskill, S. A. (1993). "The effect of aspirin on cochlear mechanical tuning," *J. Acoust. Soc. Am.* **93**, 3298–3307.
- Burns, E. M., Hoberg-Arehart, K. H., and Campbell, S. L. (1992). "Prevalence of spontaneous otoacoustic emissions in neonates," *J. Acoust. Soc. Am.* **91**, 1571–1575.
- Burns, E. M., Strickland, E. A., Tubis, A., and Jones, K. (1984). "Interactions among spontaneous otoacoustic emissions. I. Distortion products and linked emissions," *Hear. Res.* **16**, 271–278.
- Cooper, N. P., and Rhode, W. S. (1992). "Basilar mechanics in the hook region of the cat and guinea pig cochlea: Sharp tuning and nonlinearity in the absence of baseline position shifts," *Hear. Res.* **63**, 163–190.
- Festen, J. M., and Plomp, R. (1983). "Relations between auditory functions in impaired hearing," *J. Acoust. Soc. Am.* **73**, 652–662.
- Folsom, R. C., Burns, E. M., Morrison, R. B., and Zettner, E. M. (1995). "Comparison of peripheral vs. central tuning in human adults and infants," *Abstr. Assoc. Res. Otolaryngol.* **18**, 120.
- Harris, F. P., and Glattko, T. J. (1992). "The use of suppression to determine the characteristics of otoacoustic emissions," *Semin. Hear.* **13**(1), 67–80.
- Harris, F. P., Probst, R., and Xu, L. (1992). "Suppression of the 2f1-f2 otoacoustic emission in humans," *Hear. Res.* **64**, 133–141.
- Howard, M. A., Stagner, B. B., Lonsbury-Martin, B. L., and Martin, G. K. (2002). "Effects of reversible noise exposure on the suppression tuning of rabbit distortion-product otoacoustic emissions," *J. Acoust. Soc. Am.* **111**, 285–296.
- Kemp, D. T., and Chum, R. A. (1980). "Observations on the generator mechanism of stimulus frequency acoustic emissions—two tone suppression," in *Psychophysical, Physiological and Behavioral Studies in Hearing: Proceedings of the 5th International Symposium on Hearing*, Noordwijkerhooft, The Netherlands, 8–12 April 1980, edited by G. van den Brink and F. A. Bilsen.
- Kulawiec, J. T., and Orlando, M. S. (1995). "The contribution of spontaneous otoacoustic emission to the click evoked otoacoustic emissions," *Ear Hear.* **16**(5), 515–520.
- Kummer, P., Janssen, T., and Arnold, W. (1995). "Suppression tuning characteristics of the 2f1-f2 distortion product otoacoustic emission in humans," *J. Acoust. Soc. Am.* **98**, 197–210.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Lieberman, M. C., and Dodds, L. W. (1984). "Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves," *Hear. Res.* **16**, 55–74.
- Long, G. R., and Tubis, A. (1988). "Investigations into the nature of the association between threshold microstructure and otoacoustic emissions," *Hear. Res.* **36**, 125–138.
- Long, G. R., Tubis, A., and Jones, K. L. (1991). "Modeling synchronization and suppression of spontaneous otoacoustic emissions using Van der Pol oscillators: Effects of aspirin administration," *J. Acoust. Soc. Am.* **89**, 1201–1212.
- Martin, G. K., Jassir, D., Stagner, B. B., and Lonsbury-Martin, B. L. (1998). "Effects of loop diuretics on the suppression tuning of distortion-product otoacoustic emissions in rabbits," *J. Acoust. Soc. Am.* **104**, 972–983.
- Michéyl, C., and Collet, L. (1994). "Interrelations between psychoacoustical tuning curves and spontaneous and evoked otoacoustic emissions," *Scand. Audiol.* **23**, 171–178.
- Moore, B. C. (1986). *Frequency Selectivity in Hearing* (Academic, London).
- Neumann, J., Uppenkamp, S., and Kollmeier, B. (1997). "Interaction of otoacoustic emissions with additional tones: suppression or synchronization?" *Hear. Res.* **103**, 19–27.
- Norton, S. J., and Neely, S. T. (1987). "Tone-burst-evoked otoacoustic emissions from normal-hearing subjects," *J. Acoust. Soc. Am.* **81**, 1860–1872.
- Patuzzi, R. B., Sellick, P. M., and Johnstone, B. M. (1984). "The modulation of the sensitivity of the mammalian cochlea by low frequency tones. III. Basilar membrane motion," *Hear. Res.* **13**, 19–27.
- Prieve, B. A., and Falter, S. R. (1995). "COAEs and SSOAEs in adults with increased age," *Ear Hear.* **16**, 521–528.
- Probst, R., Coats, A. C., Martin, G. K., and Lonsbury-Martin, B. L. (1986). "Spontaneous, click- and toneburst-evoked otoacoustic emission from normal ears," *Hear. Res.* **21**, 261–275.
- Rabinowitz, W. M., and Widin, G. P. (1984). "Interaction of spontaneous otoacoustic emissions and external sounds," *J. Acoust. Soc. Am.* **76**, 1713–1720.
- Rhode, W. S. (1978). "Some observations on cochlear mechanics," *J. Acoust. Soc. Am.* **64**, 158–176.
- Robles, L., Ruggero, M. A., and Rich, N. C. (1991). "Two-tone distortion in the basilar membrane of the cochlea," *Nature (London)* **349**, 413–414.
- Ruggero, M. A., Robles, L., Rich, N. C., and Recio, A. (1992). "Basilar membrane responses to two-tone and broadband stimuli," *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **336**(1278), 307–315.
- Sellick, P. M., and Russell, I. J. (1979). "Two-tone suppression in cochlear hair cells," *Hear. Res.* **1**, 227–236.
- Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (1982). "Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique," *J. Acoust. Soc. Am.* **72**, 131–141.
- Siegel, J. H., Borneman, A. L., and Dreisbach, L. E. (2000). "Suppressor conditions for optimal separation of distortion product otoacoustic emission sources," *Assoc. Res. Otolaryngol. Abstr.* **23**, 283.
- Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P., and Mott, J. (1985). "Speech perception ability and psychophysical tuning curves in hearing-impaired listeners," *J. Acoust. Soc. Am.* **77**, 620–627.
- Stover, L. J., Neeley, S. T., and Gorga, M. P. (1996). "Latency and multiple sources of distortion product otoacoustic emissions," *J. Acoust. Soc. Am.* **99**, 1016–1024.
- Talmadge, C. L., Long, G. R., Tubis, A., and Dhar, S. (1999). "Experimental confirmation of the two-source interference model for the fine structure of distortion-product otoacoustic emissions," *J. Acoust. Soc. Am.* **105**, 275–292.
- Taschenberger, G., and Manley, G. A. (1998). "General characteristics and suppression tuning properties of the distortion-product otoacoustic emission 2f1-f2 in the barn owl," *Hear. Res.* **123**, 183–200.
- Tavartkiladze, G. A., Frolenkov, G. I., Kruglov, A. V., and Artamasov, S. V. (1994). "Ipsilateral suppression effects on transient evoked otoacoustic emission," *Br. J. Audiol.* **28**, 193–204.
- Taylor, M. M., and Creelman, C. D. (1967). "PEST: Efficient estimates on probability functions," *J. Acoust. Soc. Am.* **41**, 782–787.
- Tyler, R. S. (1986). "Frequency resolution in hearing impaired listeners," in *Frequency Selectivity in Hearing*, edited by C. J. Moore (Academic, London).
- Vogten, L. L. M. (1974). "Pure tone masking: a new result from a new method," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer-Verlag, Berlin).
- West, P. D. B., and Evans, E. F. (1990). "Early detection of hearing damage

- in young listeners resulting from exposure to amplified music," *Br. J. Audiol.* **24**, 89–103.
- Wilson, J. P. (1980). "Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus," *Hear. Res.* **2**, 233–252.
- Withnell, R. H., Yates, G. K., and Kirk, D. L. (2000). "Changes to low-frequency components of the TEOAE following acoustic trauma to the base of the cochlea," *Hear. Res.* **139**, 1–12.
- Xu, L., Probst, R., Harris, F. P., and Roede, J. (1994). "Peripheral analysis of frequency in human ears revealed by tone burst evoked otoacoustic emissions," *Hear. Res.* **74**, 173–180.
- Zettner, E., Folsom, R., and Burns, E. (1996). "DPOAE suppression tuning curves in human subjects during salicylate ototoxicity," *Abstr. Assoc. Res. Otolaryngol.* **96**, 27.
- Zwicker, E., and Wesel, J. (1990). "The effect of "addition" in suppression of delayed evoked oto-acoustic emissions and in masking," *Acustica* **70**, 189–196.

Modulation masking in cochlear implant listeners: envelope versus tonotopic components

Monita Chatterjee^{a)}

Department of Auditory Implants and Perception, House Ear Institute, 2100 West Third Street, Los Angeles, California 90057

(Received 27 December 2001; revised 31 December 2002; accepted 6 January 2003)

It is hypothesized that channel-interaction in cochlear implant listeners as measured in a modulation-masking experiment would be influenced by both the tonotopic overlap between masker and signal as well as an interaction between their envelopes. Two experiments were conducted to measure the effects of maskers with noisy and steady-state envelopes on modulation detection by four adult Nucleus-22 cochlear implant listeners, as a function of tonotopic distance between the masker and the signal. In the first experiment, we measured detection thresholds for a 50-Hz modulation in the envelope of a 500-Hz carrier pulse train, in the presence of a masker stimulating regions basal and apical to, as well as overlapping with, the signal. The maskers had two kinds of envelopes: (i) amplitude-modulated by flat-spectrum noise (NAM) and (ii) steady-state (SS_{peak}) at a level corresponding to the maximum of the noise fluctuation range. In general, modulation thresholds obtained in the presence of the NAM maskers significantly exceeded thresholds obtained with the corresponding SS_{peak} maskers. The ratio ρ of the threshold modulation depth m obtained with the NAM masker to that obtained with the SS_{peak} masker was defined as a conservative index of “envelope masking.” In the second experiment, ρ was determined for two different tasks: the detection of modulation at 20 Hz and steady-state intensity increment detection. Compared to the 50-Hz modulation detection results, the ratio ρ was reduced for the 20-Hz modulation detection task and even more so for the steady-state increment detection task. It is concluded that channel-interaction can be significantly increased in cochlear implant listeners when dynamic stimuli are used in place of steady-state stimuli. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1555613]

PACS numbers: 43.66.Ts, 43.66.Ba, 43.66.Dc [NFV]

I. INTRODUCTION

Present-day cochlear implants are capable of transmitting a great deal of information to the brain. The electrical signals stimulating the auditory nerve are delivered to multiple channels at a rapid rate. The perceptual consequences of multi-channel, concurrent stimulation are likely to arise from interactions at various levels of the auditory system. At the most peripheral level, summations and cancellations will occur instantaneously between overlapping electric fields. One commonly used method to avoid this is to stagger pulses in time across channels, or “interleave” them. Even with interleaving, if the pulse rate is fast enough, residual charge on the auditory-nerve membrane from prior stimuli is likely to interact with the charge from future stimuli. For auditory neurons responding to stimuli on adjacent channels, recovery from stimulation on one channel will influence the response to the second. This kind of peripheral interaction is expected to accumulate over time for very rapid stimulation rates. At a more central processing stage, interactions may occur between stimuli that do not necessarily generate overlapping excitation patterns, but have coherent fluctuations or similar temporal envelopes. The net interaction between concurrent stimuli on two channels of a cochlear implant is likely to be

a composite of their mutual distance in tonotopic (more peripheral) and temporal (more central) domains.

Speech processors in cochlear implant devices deliver acoustic information in the form of amplitude modulations in carrier signals stimulating tonotopically appropriate electrodes in the implanted array. It is therefore of some importance to understand the processing of such modulation by cochlear implant listeners. We are particularly interested in quantifying how modulation is processed in the presence of competing fluctuations, a common everyday challenge for cochlear implant listeners.

In a task involving the detection of envelope modulation on a “signal” channel in the presence of a competing fluctuation on a “noise” channel, masking may occur in both tonotopic and envelope domains. A masker having more tonotopic overlap with the signal may be expected to produce more masking than a tonotopically remote masker: we define this as “tonotopic masking.” In the second kind of interaction considered here, a masker with a temporal envelope that is similar to the signal envelope will have a greater effect on modulation detection than a masker with a perceptually distinct temporal envelope: we define this as “envelope masking.” Whereas tonotopic interactions can be explained by overlap at the level of the auditory nerve, temporal envelope interactions are likely to involve more central processes.

^{a)}Electronic mail: monita@hei.org

In the psychoacoustics literature, envelope interactions have been referred to as “modulation masking” (Bacon and Grantham, 1989) or as “modulation detection interference” (MDI) (Yost and Sheft, 1989). When the carriers have been broadband, the term “modulation masking” has been used, while “MDI” refers more to envelope interactions between narrow-band carriers. However, the distinction is unclear: MDI may be attributed to both within-channel and across-channel interactions (Bacon and Grantham, 1989; Yost and Sheft, 1989; Houtgast, 1989; Bacon and Konrad, 1993). Across-channel MDI has been demonstrated in hearing-impaired listeners (Grose and Hall, 1994), suggesting that the central processing necessary to produce such effects is preserved in listeners with cochlear loss. Recent experiments in acoustic hearing have implicated auditory grouping mechanisms in across-channel MDI (Hall and Grose, 1991; Oxenham and Dau, 2001; Gockel and Carlyon, 2000). It is possible that within-channel MDI effects are due to the auditory system essentially performing a temporal pattern discrimination task (a modulation depth discrimination task) between the signal-modulated and the signal-unmodulated intervals. Auditory grouping of the masker and signal may also play a role in within-channel MDI, as both stimuli would excite the same peripheral channel.

Several researchers working with cochlear implant listeners have studied channel-interactions in the tonotopic domain using forward-masked excitation patterns, across-channel gap-detection, and pitch comparisons across electrodes (Shannon, 1983; Lim *et al.*, 1989; Tong and Clark, 1986; Hanekom and Shannon, 1998; Chatterjee and Shannon, 1998; Throckmorton and Collins 1999). McKay and McDermott (1996) studied spatio-temporal interactions by measuring discriminability and perceptual fusion of the temporal patterns of interleaved pulse trains on different electrode pairs. A general result of these studies is that information is integrated across electrode locations spanning at least 3 mm on the cochlea (for the Nucleus-22 device, this means a distance of four electrode separations). More recently, Richardson *et al.* (1998) have shown the presence of masking in the temporal domain in cochlear implant listeners in an MDI experiment. They found that the masking due to a modulated masker exceeded that due to an unmodulated masker by amounts that depended on the individual subject, the tonotopic distance between the masker and the signal, and the modulation depth in the masker envelope. In general, the magnitude of MDI depended on masker-signal tonotopic distance, being larger for maskers that were closer to the signal.

In experiments with cochlear implants, “channels” are not as easily defined as in normal hearing. To sidestep the ambiguity, we will refer to the net masking of modulation, within- or across-channel, as “modulation masking.” We speculate that both “tonotopic masking” and “envelope masking” contribute to modulation masking. We tentatively suggest that “tonotopic masking” and “envelope masking” respectively correspond to within- and across-channel processes yielding MDI in acoustic experiments.

Here, we explore in some detail the relative modulation-masking effects of steady-state maskers versus dynamic

maskers with noisy envelopes, as a function of the masker-signal tonotopic distance. We initially hypothesized that tonotopic masking would be maximum when this distance was zero, and that it would decrease as the masker-signal electrode distance was increased. It was less easy to formulate a hypothesis about envelope masking. We began with the assumption that envelope masking would be independent of masker-signal electrode distance: this would predict that masked modulation thresholds with the fluctuating maskers would be elevated by a constant amount over thresholds obtained with the steady-state maskers. However, the results of the experiments showed that, in fact, envelope masking depends on masker-signal distance in a complex manner.

In these experiments, the cochlear implant microphone and speech processor were bypassed and the stimuli were delivered through a custom research interface. This made it possible to control the “tonotopic” and “envelope” aspects of the stimuli precisely, without any intervening spectral processing. When fluctuations were applied to a carrier pulse train, they changed only the temporal envelope of the signal at the desired electrode location: no sidebands were generated on the excitation pattern in the electrode array. This is an important factor of such experiments to keep in mind when comparing the experimental conditions described here with those in similar acoustic experiments.

However, fluctuations in the excitation pattern might accompany envelope fluctuations at the auditory-nerve level, i.e., the peaks of the envelope might excite more neurons than the valleys. To ensure that our measure of envelope masking did not stem from such spread of excitation at the peaks of the masker envelope, the amplitude of the steady-state masker we used as our reference was fixed at the peak of the corresponding noisy envelope. A comparison of modulation masking with these two kinds of maskers is expected to yield a conservative measure of envelope masking.

II. EXPERIMENT 1

In experiment 1, we measured the effects of noisy maskers (i.e., masker pulse trains with noisy envelopes) on the detection of a 50-Hz modulation in the envelope of the signal. The 50-Hz modulation frequency was selected as one that is in the most sensitive region of the modulation transfer function for these subjects at the carrier level used in these experiments.

A. Subjects

The subjects (S1, S2, S3, and S4) were four postlingually deafened adult users of the Nucleus-22 cochlear implant device. They have participated in numerous psychophysical experiments and may be considered to be highly trained. They also participated in the preliminary experiments leading up to this study and were specifically trained in modulation masking tasks. These subjects perform well on speech recognition tests in the laboratory: scores range from 80% to 100% in easy sentence recognition tests (CUNY, sound-only). Table I lists other potentially relevant information.

TABLE I. Potentially relevant information about subjects.

Subject	Age	Gender	Age at onset of profound hearing loss	Duration of implant use on the side tested
S1	60	M	45	11 years
S2	59	F	38	11 years
S3	59	M	47	6 years
S4	44	M	35	9 years

B. Stimuli

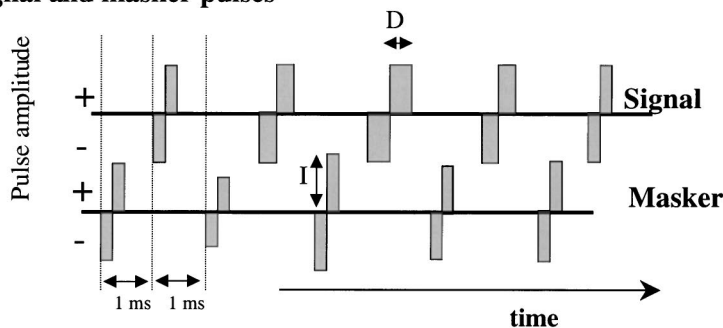
Both masker and signal were 200-ms-long, 500-Hz trains of biphasic current pulses (200 μ s/phase). Masker and signal pulses were interleaved in time with a 1-ms delay between the onset of each masker pulse and the onset of the next signal pulse [see Fig. 1(a)] to achieve concurrent but nonsimultaneous stimulation. A custom-built research interface (Shannon *et al.*, 1990) was used to deliver stimuli. The exact pulse amplitudes were based upon the manufacturer's calibration table for each subject's device. In the Nucleus-22

device, amplitudes increase from the lowest ($\sim 20 \mu$ A) to the highest level (~ 1500 – 1800μ A) in 238 discrete steps, also known as clinical units. The calibrated microamperes of current at each step vary from device to device. This variation is small at lower levels and large at higher levels.

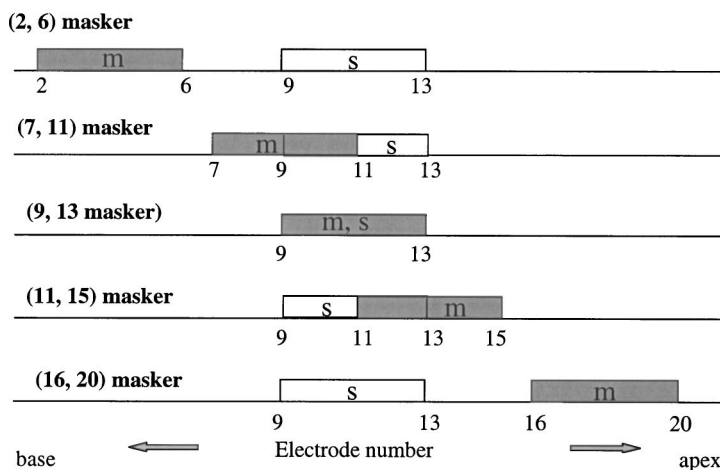
The masker and signal electrode pairs were chosen to stimulate moderately broad regions of the cochlea. At the same time, the selected stimulation mode (BP+3, or a separation of 3 mm on the cochlear length between the active and return electrodes of the pair) allowed the dynamic range to be measured without having to exceed currents of 1 mA. This was a concern because in the Nucleus-22 device, compliance limits of the current sources may be approached at high current amplitudes. The implanted electrode array has 22 electrodes, spaced 0.75 mm apart and numbered 1–22 (base to apex). We refer to bipolar electrode pairs in the form (X,Y), where X and Y are the numbers of the basal and apical members of the pair respectively.

The signal electrode pair (9,13) stimulated the central region of the electrode array. The masker electrode pairs

a. Signal and masker pulses



b. Signal and masker electrode locations



c. Masker envelopes

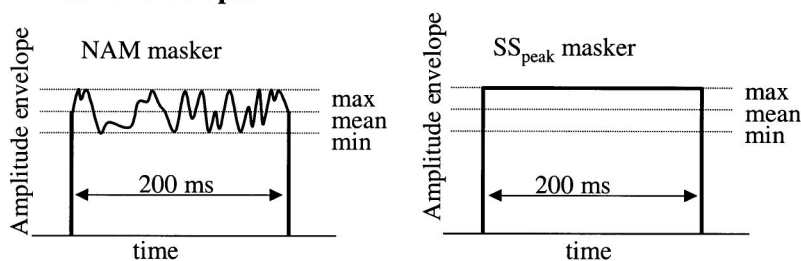


FIG. 1. (a) Schematic to illustrate interleaving of biphasic pulses of the masker and signal pulse trains. Note that the signal pulse phase duration (D) is the modulated or adapted variable, while changes in the masker envelope are applied to masker pulse amplitudes (I) in this study. (b) Schematic of idealized tonotopic distributions of masker (m) and signal (s) electrical fields. (c) Schematic of masker envelope types used.

(2,6), (7,11), (9,13), (11,15), and (16,20) were chosen to stimulate regions basal to, overlapping basally, overlapping completely, overlapping apically, and apical to the signal electrical field. The distances between the masker and signal electrode pairs were 5.25, 1.5, 0, -1.5 , and -5.25 mm along a distance scale centered on the signal electrode 9. Figure 1(b) shows a schematic of the electrode locations for the different maskers and the signal.

Modulation detection thresholds were measured for various levels of noise on each masker electrode pair. “Noise” was applied by scaling successive pulses of a periodic pulse train by numbers drawn from a pseudorandom uniform distribution with a mean of 1.0 and a range $1.0 \pm r$ ($0 < r < 1$), where r is the factor that determines the noise level (Chatterjee and Robert, 2001). We refer to noise level in percent, thus, when r is 0.20, noise level is 20%, and the pulses in the pulse train have amplitudes ranging from 80% to 120% of the reference amplitude.¹ For each level of noise on the masker, a measurement was made of modulation threshold in the presence of a similar masker pulse train presented to the same electrode pair, but with a fixed steady-state amplitude at the peak of the noise fluctuation. Thus, when the level of noise was 20%, the corresponding steady-state masker had a fixed pulse amplitude corresponding to 120% of the reference amplitude. The two maskers are labeled “NAM” (noise amplitude modulated) and “SS_{peak}” [see Fig. 1(c)]. The difference in dB between modulation detection thresholds obtained with each NAM masker to thresholds obtained with the corresponding SS_{peak} masker gives a conservative estimate of the contribution of the noise envelope to the net masking.

Dynamic range for each electrode pair was determined as follows. Detection thresholds were obtained for each unmodulated carrier as described under “Procedure for threshold measurements” below. Maximum acceptable levels (MALs) were measured using loudness estimation methods described previously (Chatterjee, 1999). Dynamic range is defined as the difference between the MAL and threshold in microamperes.

The reference levels of the masker and signal were fixed at 50% of their individual dynamic ranges in microamperes. Table II lists the actual microampere levels for threshold, maximum acceptable level, and the 50% dynamic range level for each stimulus and subject. This level corresponded to a comfortable listening level per channel. When the two channels of stimulation were combined, loudness of the composite increased. The 50% dynamic range level was set based upon pilot studies conducted to ensure that at the maximum noise level for the masker, the overall loudness for the composite masker+signal was still within the comfortably loud range.

Modulation was applied to the pulse phase duration of the signal. This is because the discrete amplitude steps increase logarithmically in the Nucleus device; as cochlear implant listeners are very sensitive to modulation above soft listening levels, in some cases the size of the amplitude step may be larger than the resolution of the subject. In contrast, the device allows fine resolution in the pulse phase duration (steps of 0.4 μ s).

TABLE II. Threshold, maximum acceptable level, and the 50% dynamic range level for the carrier stimuli used in these experiments, for each subject.

Subject	Electrode pair	Threshold (μ A)	MAL (μ A)	50% DR (μ A)
S1	(2,6)	161	667	413
	(7,11)	161	892	525
	(9,13)	139	938	540
	(11,15)	166	938	547
	(16,20)	122	794	459
S2	(2,6)	85	337	213
	(7,11)	93	445	269
	(9,13)	134	535	337
	(11,15)	176	630	401
	(16,20)	254	789	525
S3	(2,6)	225	564	396
	(7,11)	271	591	430
	(9,13)	254	581	423
	(11,15)	234	581	408
	(16,20)	293	596	445
S4	(2,6)	105	449	278
	(7,11)	95	476	286
	(9,13)	110	535	320
	(11,15)	107	633	371
	(16,20)	134	564	344

The noise, however, was applied to the amplitudes of the pulses and not to the pulse phase durations. This was done to ensure that masker and signal pulses could be interleaved in time without reducing the interpulse interval excessively. The noise was applied at levels that exceeded the restrictions imposed by the amplitude steps.

Although the scaling factors of the noise were uniformly distributed, the actual pulse amplitudes delivered were constrained to the finite number of microampere levels allowed by the device. This means that the actual distribution of amplitudes deviated from the theoretical uniform distribution to some extent. Figure 2 illustrates different aspects of the noise using subject S2’s amplitude calibration table as an example. Subject S2 was selected because the 50% dynamic range level for maskers spanned a large range of amplitudes across electrodes in her case (see Table II); as the range of amplitudes is smallest for the lowest level carrier, we can illustrate the worst case of deviation from uniformity (the 213 μ A level). The left- and right-hand panels of Fig. 2(a) show scatter plots of the actual pulse amplitudes for one pulse train at each of the five carrier amplitudes, for 10% and 20% noise levels, respectively. A close observation of Fig. 2(a) reveals that the current amplitudes are restricted to discrete levels. At the lowest carrier level, the steps are smaller in number, and the error due to quantization is larger. This is further shown in Fig. 2(b), which plots histograms of pulse amplitudes for the different carrier amplitudes, for 10% and 20% noise levels (left and right hand columns). Each of the five bins spans 20% of the amplitude range in microamps for the specific signal. The heights of the bars and the error bars show the means and standard deviations of the number of pulses falling into each bin, collected from 20 independent samples of the 100-pulse sequence. Thus, for a uniform distribution, the

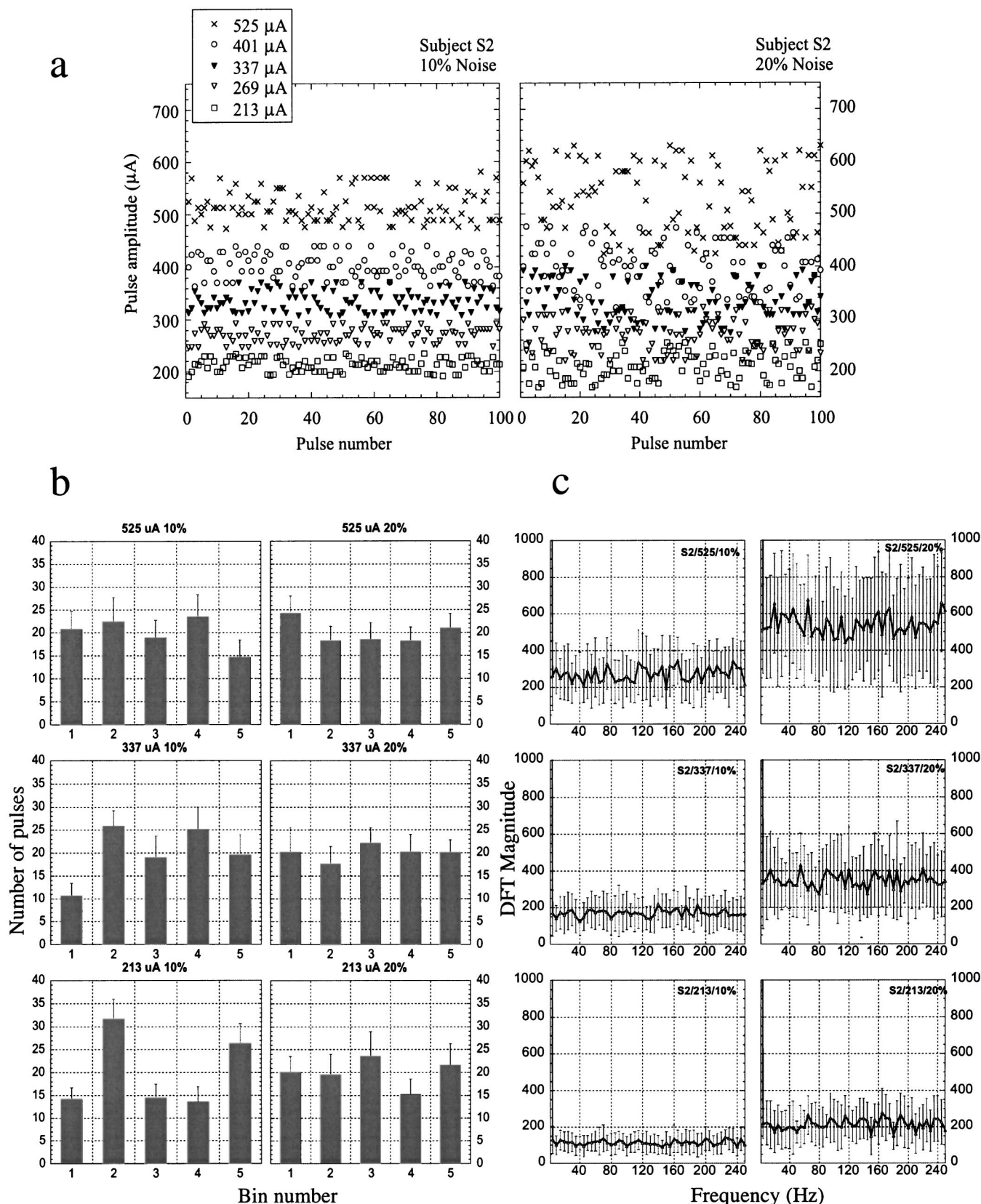


FIG. 2. (a) The actual pulse amplitudes (μA) for one 100-pulse series, for each of the five current levels corresponding to the 50% dynamic range point for the five masker electrode pairs in subject S2, obtained with 10% (left-hand panel) and 20% noise (right-hand panel). (b) Histograms showing the number of pulses falling into each of five bins, each spanning successive one-fifths of the range of amplitudes used for each carrier and noise level. Each bar shows the mean and standard deviation of the pulse counts from 20 100-pulse sequences of pulse amplitudes. Left- and right-hand panels show histograms for the 10% and 20% noise levels, respectively. Three of the five carrier amplitudes used with subject S2 are shown (top to bottom: 525, 337, 213 μA respectively). (c) Mean and standard deviation of 20 100-point DFTs conducted on the same pulse amplitude data shown in (b). No windowing was applied. The full extent of the dc component cannot be seen because of truncation of the ordinate scale.

mean should correspond to 20 pulses overall for each bin. It is apparent that the lowest carrier level and the lowest noise level results in the least uniform distribution: when the noise depth is increased or when the carrier level is increased, the number of amplitude steps available increases, and the distribution becomes more uniform. Figure 2(c) shows the means and standard deviations of 20 100-point discrete Fourier transforms (DFTs) applied to the amplitude series of the same pulse trains shown in Fig. 2(b): it is apparent that the noise has a flat spectrum up to the Nyquist frequency (250 Hz).

C. Procedure for threshold measurements

The task in experiment 1 was to detect modulation at 50 Hz in the signal carrier. For each of the five maskers, we varied the noise level from 0% to 20% of the reference amplitude, in steps of 5%. When noise was applied to the masker, within a given trial, all three intervals contained the same masker noise. The noise was refreshed from trial to trial.

All thresholds were measured using three-down, one-up, adaptive procedures, theoretically converging at the 79.4% correct point on the psychometric function (Levitt, 1971). Each run consisted of a maximum of ten reversals and a minimum of eight, and was terminated after 55 trials. The step size was reduced after the first four reversals. The mean and standard deviation of the last five to seven reversals were calculated (i.e., the first three reversals were discarded) to obtain the result of each run. In all cases, feedback was provided to the subject (correct/incorrect) on a trial by trial basis.

- (i) Absolute detection thresholds: A 2IFC procedure was used. Initial and final step sizes were 2 and 0.5 dB, respectively. The mean of two such runs was calculated to obtain the final threshold. Table II shows the mean values of the thresholds, corrected to the individual subject's amplitude calibration table.
- (ii) Modulation detection thresholds: A 3IFC procedure was used to obtain modulation detection thresholds. The (linearly) adapted variable was modulation index (m). Initial and final step sizes depended on the subject's sensitivity based on pilot data. Typically, if m at threshold was less than 0.1, initial and final steps were 0.02 and 0.005. If m exceeded 0.1, steps were typically 0.04 and 0.01. The reference per-phase pulse duration was fixed at 200 μ s/phase (the minimum measurable value of m for this reference is 0.002, corresponding to the 0.4- μ s minimum increment in duration allowed by the device). The mean and standard deviation of the results of at least four runs were calculated to obtain the final threshold. These procedures were used to measure both masked and unmasked thresholds.

D. Results

Each of the 20 panels of Fig. 3 shows modulation thresholds obtained with the NAM (squares) and SS_{peak} maskers (circles) on a particular electrode pair, expressed in

dB [$20 \cdot \log(m)$, where m is the modulation index at threshold]. The abscissa shows the noise level in the masker envelope or the corresponding increment in the SS_{peak} masker amplitude, referred to mean level. The four columns show results obtained with the four subjects. The five rows show data obtained with the five masker electrode pairs. In each panel, those modulation thresholds obtained with the NAM masker that significantly exceed the corresponding thresholds obtained with the SS_{peak} masker at the $p < 0.05$ level (Students t -test) are indicated by a pattern fill in the squares.

It is apparent that in most conditions, the NAM masker is more effective than the SS_{peak} masker. On the dB axis used in Fig. 3, the vertical distance between the two plots corresponds to $20 \cdot \log(\rho)$, where ρ is the ratio between the two masked thresholds (m obtained with the NAM and the SS_{peak} maskers). We define this ratio, ρ , as a conservative index of the strength of envelope masking over tonotopic masking. When ρ exceeds 1.0 (i.e., the dB difference between the two masked thresholds exceeds 0), we infer that envelope masking is dominant. It is apparent from a visual inspection of Fig. 3 that ρ is not always monotonically increasing with the noise level. When the masker is not overlapping with the signal [electrode pairs (2,6) and (16,20)], increasing the amplitude of the SS_{peak} masker has little effect on masked threshold, but increasing the noise in the masker has a more significant effect. As overlap between masker and signal is increased [electrode pair (7,11) or (11,15)], the SS_{peak} masker becomes more effective. When the overlap is complete [electrode pair (9,13)], the SS_{peak} masker and the NAM masker both have greater effect with increasing level, and in some cases the difference in masking between the two is insignificant. Regardless of the individual variations, thresholds with the NAM masker are either equal to (i.e., not significantly different from) or greater than thresholds with the SS_{peak} masker in all cases.

The same data are replotted in Fig. 4 to illustrate the tonotopic distribution of the masking, with noise level as the parameter. For each of the four subjects, Fig. 4 shows the modulation index m at threshold (in dB) obtained with the NAM masker (left-hand panels) and the SS_{peak} masker (middle panels), and the "envelope masking" $20 \cdot \log(\rho)$, the dB difference between the two (right-hand panels), all as a function of masker electrode location. The different symbols in each panel show thresholds obtained with different levels of noise in the masker (from 0% to 20%). The horizontal line in each plot shows the unmasked threshold for detecting a 50-Hz modulation on the signal electrode pair. Each subject's results are plotted in a different row. Recall that the signal is presented to electrode pair (9,13): its location is indicated by the downward pointing arrows in the top three panels.

Tonotopic masking is reflected in the systematic relation between masker electrode location and masked modulation threshold. While the NAM-masked thresholds (left hand panels) generally exceed thresholds obtained with the SS_{peak} maskers (middle panels), the overall shapes of the patterns in the two panels are not dissimilar. Strong peaks in the left-hand panels are generally mirrored in the middle panels, indicating that the net effect of the noise is a composite of the

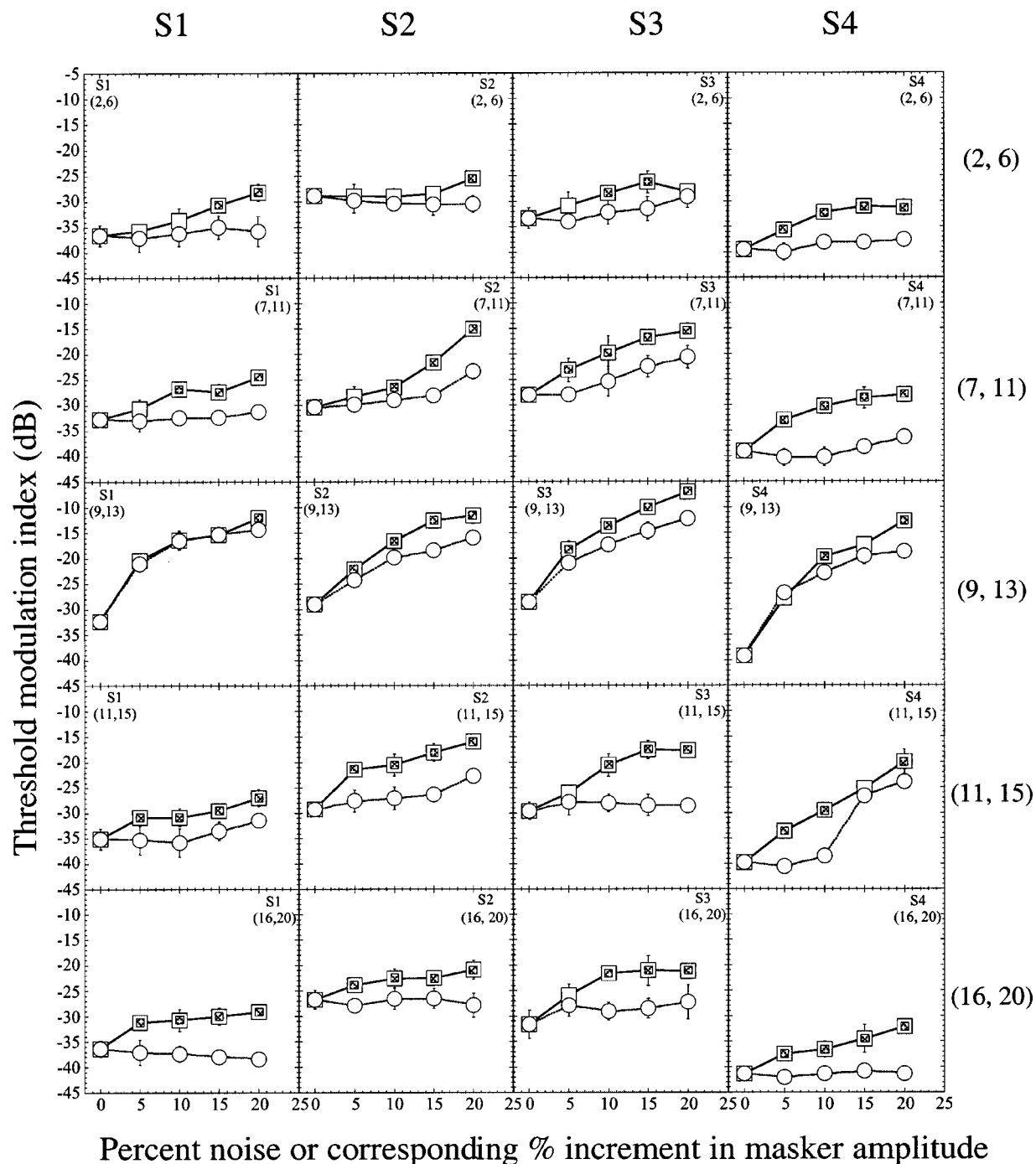


FIG. 3. Masked modulation thresholds obtained with the four subjects (S1–S4, left to right) and the five masker electrode locations (top to bottom, basal to apical). Squares show modulation thresholds obtained with NAM maskers. Unfilled circles show modulation thresholds obtained with corresponding SS_{peak} maskers. The abscissa represents increasing noise level or increment from the reference for the SS_{peak} masker. Error bars show \pm one standard deviation from the mean. Pattern fill in squares indicates a statistically significant difference from the corresponding circle ($p < 0.05$, Student's t -test).

tonotopic and the envelope contributions. (Note that the 0% noise condition is identical for the two kinds of maskers.)

In contrast, the relationship between the ratio ρ (the relative strength of envelope masking over tonotopic masking) and the tonotopic distance between the masker and signal electrode pairs is not simple, as is evident in the right-hand column in Fig. 4, which plots ρ in dB for each of the noise levels. In each case, the horizontal axis shows the basal member of the masker electrode pair, and the parameter is the noise level. The function relating ρ and masker electrode

location changes with noise level and is dependent upon the individual subject. The strong peaks in the two left-hand panels in each row are not mirrored in the corresponding right hand panels: the peaks in ρ generally occur at different masker electrode locations than the peaks in the absolute modulation thresholds. In fact, ρ often shows a minimum when the masking by the NAM and SS_{peak} maskers is maximum. This pattern is consistent with our hypothesis that masking by the SS_{peak} masker reflects the component due to tonotopic overlap, and that the ratio ρ reflects the contribu-

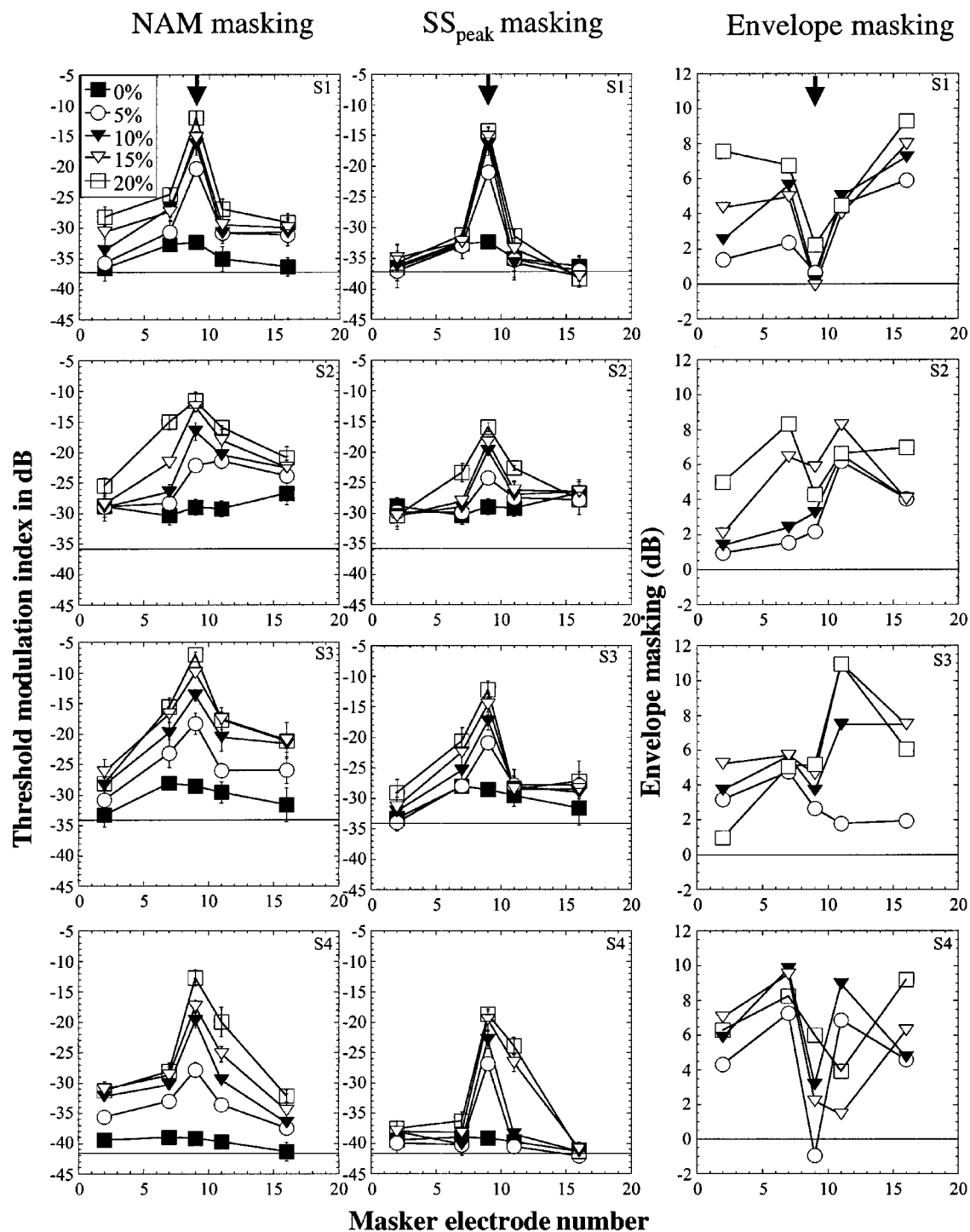


FIG. 4. Tonotopic distribution of masked thresholds with the NAM masker (left-hand panels) and the SS_{peak} masker (middle panels) for each of the four subjects. The right-hand panels show “envelope masking” [dB difference between NAM- and SS_{peak} -masked thresholds, or $20 \cdot \log(p)$]. The symbols show the noise level. The downward pointing arrows in the top three panels show the location of the signal. Solid horizontal lines in each panel of the two left-hand columns show unmasked threshold for detecting the 50-Hz modulation. Solid horizontal line in each panel in the right hand column shows 0 dB of envelope masking.

tion of envelope interaction. We would expect the SS_{peak} masking pattern to have peaks near the signal electrode location: however, the envelope component is not necessarily expected to show the same pattern. The results show strong dependence on tonotopic distance between the two. Although the nature of this dependence is complex, subject-dependent, and also dependent on the level of the noise, the pattern

suggests that the masking by the NAM masker can be largely attributed to envelope masking when the masker is apical to or basal to the signal [electrode pairs (2,6) and (16,20)] and tonotopic masking is minimal. In the case of the other three, more overlapping maskers, the relative strengths of the tonotopic and envelope components change in a more complex manner with noise level.

III. EXPERIMENT 2

One question that arises from the results of experiment 1 concerns the possibility that some of the masking effect of the noise may be simply explained by greater uncertainty in the task in the presence of random fluctuations in the masker. If this is the case, we would expect to observe identical or similar effects when the task is changed (making the reasonable assumption that increasing the uncertainty by a certain amount results in similar elevation in thresholds across tasks). On the other hand, if the effect is truly due to temporal interactions between the envelopes, we might expect to observe some task-related differences in the effect, especially if we changed the task along a temporal dimension such as the modulation rate.

The specific objective of this experiment was to determine whether the greater masking effect of the NAM masker over the SS_{peak} masker would persist when the task was changed to the detection of a slower modulation frequency (20 Hz) or to the detection of a steady-state increment in the charge per pulse in an unmodulated carrier. We did not parametrically vary the masker noise level, but rather chose a fixed noise level of 20% for this experiment.

A. Subjects, stimuli, and procedures

The same four subjects S1, S2, S3, and S4 participated in experiment 2. Masker and signal stimuli were identical to those in experiment 1. Modulation detection thresholds for 20-Hz modulation were measured in the same way as in the previous experiment. When measuring the just detectable intensity increment, the task was to detect a static increment in the pulse phase duration in the signal over the standard 200 μ s/pulse phase duration. Experimental methods were identical to those used in experiment 1.

B. Results

Each panel in Fig. 5 shows the thresholds obtained in the presence of the NAM, SS_{peak} , and the reference steady-state masker (the 0% noise masker), as a function of masker electrode location. Each row corresponds to data obtained with a different subject. The three columns show results obtained with the three tasks: the left-hand one shows the 50-Hz modulation detection thresholds (identical to experiment 1), the middle one shows the 20-Hz modulation detection thresholds, and the right-hand one shows the increment detection thresholds. Within each panel, filled circles show modulation thresholds obtained with the 0% noise masker; unfilled squares show thresholds obtained with the 20% noise masker, and the unfilled circles show thresholds obtained with the SS_{peak} masker corresponding to the 20% noise level. The squares with pattern fill represent modulation thresholds obtained with the NAM masker that significantly exceed modulation thresholds obtained with the SS_{peak} masker at the $p < 0.05$ level of significance. Horizontal lines show the unmasked threshold (on the signal electrode pair) for the respective tasks. Once again, the downward pointing arrows in the top panels show the location of the signal.

In general, steady-state maskers at the reference amplitude have the least effect on these tasks (thresholds shown by

the filled circles, labeled 0% noise). The SS_{peak} masker has a considerably larger effect in all cases, especially when the masker electrodes approach or overlap the signal electrodes. Compared to the results obtained for the 50-Hz modulation detection task, we observe fewer significant increases in modulation thresholds with the NAM masker over the SS_{peak} masker when the task is 20-Hz modulation detection, and even fewer when the task is increment detection. It is clear that the vertical distance between the squares (NAM masker) and the unfilled circles (SS_{peak} masker) diminishes as we go from left to right in Fig. 5. When the task is 20-Hz modulation detection or intensity increment detection, the effect of the SS_{peak} masker is sometimes stronger than the NAM masker. This is clearly seen in the behavior of the ratio ρ . For each subject, the tonotopic distribution of $20 \cdot \log(\rho)$ (the dB difference between the noise and SS_{peak} maskers) for the three different tasks are plotted in Fig. 6. In each panel, the arrows indicate the signal location. For all subjects, ρ is generally largest for 50-Hz modulation detection (squares). We observe much less envelope masking for 20-Hz modulation detection (filled circles). For the increment detection task (unfilled circles), values of ρ are consistently closer to 1.0 (the dB value is close to 0.0), and in some cases much less than 1.0 (negative dB values), indicating that tonotopic masking is the dominant underlying factor. If the masking were mostly due to tonotopic overlap, we would indeed expect the SS_{peak} masker to be dominant, as it is the more intense of the two maskers.

Thus, the effect of envelope masking by noise is greatest at the higher modulation frequency, less at the lower modulation frequency, and becomes negligible or is even reversed in the intensity increment detection task.

IV. DISCUSSION

We have observed here at least two separable components of masking: one arising from overlap in the tonotopic excitation patterns evoked by the masker and the signal, and a second, arising from a potentially more central interaction, due to the envelope of the noise masking the signal envelope. It is likely that the envelope interaction is identical to the “modulation detection interference” phenomenon observed with modulated maskers in acoustic and electric hearing. In preliminary work with sinusoidally amplitude-modulated maskers, we find comparable amounts of modulation masking in the same subjects (Chatterjee, 2002). The results also suggest that modulation masking measured in the conditions when masker and signal overlap is dominated by tonotopic masking by the peaks of the masker envelope, and the weight of envelope masking is diminished. It is interesting to note that the unmodulated masker at the reference amplitude (0% noise) had little effect on modulation detection, even in the condition of complete overlap. From a visual inspection of Fig. 4, it is evident that if we had used the difference in modulation thresholds obtained with the NAM masker and the unmodulated masker (0% noise level) as our index of envelope interaction, we would have found a peak in this index in the “within-channel” condition [consistent with results obtained with sinusoidally modulated maskers by Rich-

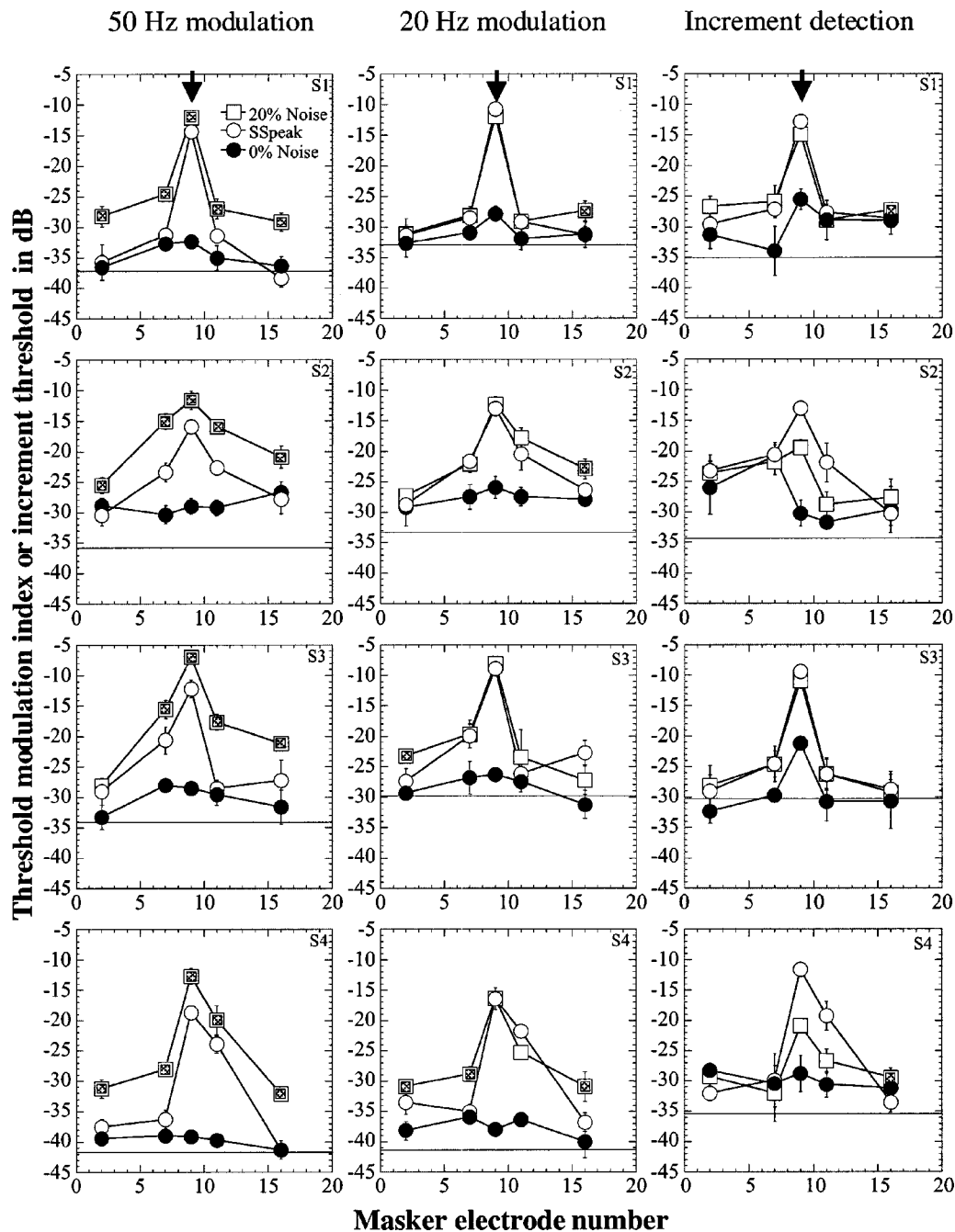


FIG. 5. A comparison of masking in three kinds of tasks. Masker noise level fixed at 20%. For each subject, the left-hand panel replots results obtained for 50-Hz modulation detection, the middle panel plots results obtained for 20-Hz modulation detection, and the right-hand panel plots results obtained for pulse phase duration increment detection. In each panel, the filled circles show results obtained with the unmodulated masker at the reference level (0% noise), the unfilled circles show results obtained with the unmodulated SS_{peak} masker for the 20% noise level, and the squares show results obtained with the 20% noise masker. Pattern fill in squares indicates that the NAM-masked thresholds significantly exceeded the SS_{peak} -masked thresholds ($p < 0.05$, Student's t -test). Horizontal lines indicate unmasked thresholds for each task. The location of the signal is again shown by the downward pointing arrows in the top panels.

ardson *et al.* (1998)]. The use of the SS_{peak} masker as the reference yields an essentially different, and we would argue truer, pattern of envelope interaction.

There seems to be a hierarchy in the effect of the noise, which is strongest on the 50-Hz modulation detection task, weakest on the increment detection task, and weak to moderate on the 20-Hz modulation detection task. This suggests that the effect of the NAM masker is not due to increased uncertainty, since greater uncertainty should have produced similar elevations in thresholds in all three tasks in experi-

ment 2. The fact that the interaction of the masker and signal envelopes is greatest for the more dynamic signal is interesting in light of the flat envelope spectrum of the noise (Fig. 2). It is also interesting to observe that masking by the SS_{peak} masker exceeds masking by the NAM masker (i.e., the ratio ρ is less than 1.0) in several cases when the task is intensity increment detection. This would be expected if all or most of the masking were due to tonotopic masking (recall that the pulse amplitude in the SS_{peak} masker is 20% greater than the average pulse amplitude in the NAM masker).

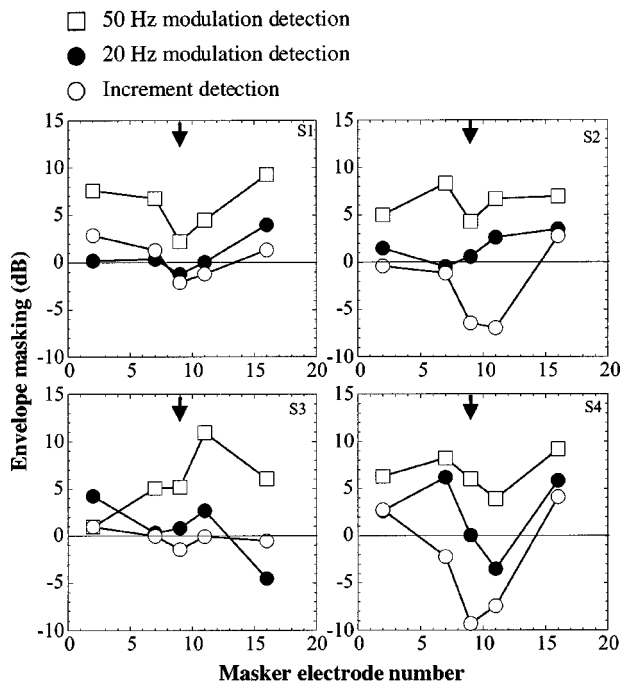


FIG. 6. A comparison of the amount of envelope masking (the ratio ρ in dB units) obtained in the three kinds of tasks in experiment 2, plotted as a function of masker electrode location. Each panel shows results obtained with a different subject. In each panel, the signal location is shown by the downward pointing arrow.

It seems reasonable to speculate that the two kinds of masking observed here originate in different mechanisms. Tonotopic masking can be explained in the classical way—as due to the overlap of neural populations (relatively peripheral in locus) responding to the masker and the probe. Envelope masking by the noise, as a phenomenon that does not require tonotopic proximity between masker and signal, is likely to arise in more central processes.

Two aspects of the noise might contribute to the envelope masking that we observe. One contributing factor may be that the intrinsic envelope fluctuations in the noise masker interact in some way with the signal modulation at some convergent point of processing. It is also possible that such interaction is due to spectral processing of the envelope—the noise might excite hypothetical modulation detectors tuned to several modulation frequencies (Dau *et al.*, 1997). The results of experiment 2 are consistent with the modulation filterbank model and related experimental results of Dau *et al.* (1997), assuming more sharply tuned (along a linear modulation frequency scale) modulation filters centered around lower modulation frequencies and more broadly tuned modulation filters centered around higher modulation frequencies. The NAM masker would excite the 20-Hz modulation filter less than the 50-Hz modulation filter, and thereby cause more interference with the latter.

Whatever the precise mechanism—a temporal interaction between the envelopes or a specific spectral analysis of the envelopes—consistent with results obtained in normal hearing listeners, it seems that a model incorporating converging inputs from peripheral channels onto some central “envelope” processor is necessary to account for the results obtained in cochlear implant listeners. How this convergence

might occur is unclear. In experiments in across-channel MDI, it has been shown that asynchronous gating can reduce the amount of MDI in normal hearing listeners (Hall and Grose, 1991), suggesting that auditory grouping may be involved in the phenomenon. Recently, Oxenham and Dau (2001) have confirmed these results and further shown that when the masker is one of a “stream” of identical precursors, MDI is reduced significantly in normal-hearing subjects. This finding strengthens the argument for the involvement of central, grouping phenomena in modulation detection interference. Whether such phenomena also involve the kind of processing suggested by Dau *et al.*’s (1997) modulation filterbank model is not known.

Recent work by Gockel and Carlyon (2000) suggests that such “grouping” may not always require the masker and signal to be concurrent. In experiments with FM detection (15 Hz), the authors have shown that some modulation detection interference persists even when the masker and signal are not simultaneous, but the signal is flanked in time by the masker: further, the authors found that increasing the gaps between masker and signal up to 200 ms did not influence the effect. In contrast, steady-state maskers had no effect on the FM detection thresholds. The authors hypothesized that the greater perceptual salience of fluctuating signals contributes to the strong masking effect, which persists even when masker and signal are nonsimultaneous.

In recent preliminary experiments, we have observed strong streaming effects in cochlear implant listeners, for sequential patterns of carriers that are either tonotopically identical, but very different in temporal envelope, or for carriers that have identical envelopes but are different in tonotopic location (Chatterjee and Galvin, 2002). These results indicate that mechanisms for perceptual organization are preserved in auditory perception through a cochlear implant. It would therefore not be surprising to find that envelope masking in cochlear implant perception arises from such mechanisms.

The majority of cochlear implant listeners have great difficulty understanding speech in noisy or multi-speaker environments. The present results suggest that competing fluctuations on concurrently stimulated channels may compromise the listener’s ability to process temporal information, even when the electrodes are tonotopically distant. Thus, “channel-interaction” can be much larger in the case of dynamic stimuli than steady-state stimuli. It is apparent that future work needs to consider the dynamic aspects of component stimuli in multi-channel stimulation of cochlear implants.

V. CONCLUSION

This study has shown that channel-interaction in cochlear-implant listeners is altered significantly in the case of dynamic stimuli. The use of the SS_{peak} masker in these experiments was important in separating the effects of tonotopic masking from true envelope masking. The results suggest that envelope interactions are likely to arise from processes that are distinct from tonotopic interactions between signals. We speculate that these processes are central in origin. Much further work needs to be done to explore more

precisely the mechanisms underlying such interactions. Finally, it appears that measures of channel-interaction in cochlear implant listeners should incorporate dynamic aspects of signals: steady-state stimuli do not tell the entire story, even in a relatively simple experimental paradigm such as the one employed in this study.

ACKNOWLEDGMENTS

We thank Mark E. Robert for programming the interface used in these experiments, and John J. Galvin III for his help with pilot-data collection. We thank Sandra Oba for her assistance in running experiments and her help with manuscript preparation, editing, and proofing, and Cochlear Corporation for providing us with the calibration tables necessary to conduct these experiments. Neal Viemeister and two anonymous reviewers are thanked for their constructive criticisms of an earlier version of this work, resulting in substantial changes to the manuscript. We are very grateful to the cochlear implant listeners who participated in these experiments for their continued support of this research. This work was supported by NIDCD (Grant Nos. R03DC03519-03, N01DC92100-03, R01DC01526-S1, and R01DC004786-01).

¹Note that with uniformly distributed noise, the mean is constant regardless of the level of the noise. However, the variance of the pulse amplitudes changes with changing noise level.

- Bacon, S. P., and Grantham, D. W. (1989). "Modulation masking: effects of modulation frequency, depth and phase," *J. Acoust. Soc. Am.* **85**, 2575–2580.
- Bacon, S. P., and Konrad, D. L. (1993). "Modulation detection interference under conditions favoring within- or across-channel processing," *J. Acoust. Soc. Am.* **93**, 1012–1022.
- Chatterjee, M. (1999). "Effects of stimulation mode on threshold and loudness growth in multielectrode cochlear implants," *J. Acoust. Soc. Am.* **105**, 850–860.
- Chatterjee, M. (2002). "Modulation masking in cochlear implant listeners: effects of masker envelope," *J. Acoust. Soc. Am.* **111**, 2338(A).
- Chatterjee, M., and Galvin III, J. J. (2002). "Auditory streaming in cochlear implant listeners," *J. Acoust. Soc. Am.* **111**, 2429(A).
- Chatterjee, M., and Robert, M. E. (2001). "Noise enhances modulation sensitivity in cochlear implant listeners: stochastic resonance in a prosthetic sensory system?" *J. Assoc. Res. Otolaryngol.* **2**(2), 159–171.

- Chatterjee, M., and Shannon, R. V. (1998). "Forward masked excitation patterns in multielectrode electrical stimulation," *J. Acoust. Soc. Am.* **103**, 2565–2572.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," *J. Acoust. Soc. Am.* **102**, 2892–2905.
- Gockel, H., and Carlyon, R. P. (2000). "Frequency modulation detection interference produced by asynchronous and nonsimultaneous interferers," *J. Acoust. Soc. Am.* **108**, 2329–2336.
- Grose, J. H., and Hall III, J. W. (1994). "Modulation Detection Interference (MDI) in listeners with cochlear hearing loss," *J. Speech Hear. Res.* **37**(3), 680–686.
- Hall III, J. W., and Grose, J. H. (1991). "Some effects of auditory grouping factors on modulation detection interference (MDI)," *J. Acoust. Soc. Am.* **90**, 3028–3035.
- Hanekom, J. J., and Shannon, R. V. (1998). "Gap detection as a measure of electrode interaction in cochlear implants," *J. Acoust. Soc. Am.* **104**, 2372–2384.
- Houtgast, T. (1989). "Frequency selectivity in amplitude-modulation detection," *J. Acoust. Soc. Am.* **85**, 1676–1680.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am. Suppl. 1* **49**, 467.
- Lim, H. H., Tong, Y. C., and Clark, G. M. (1989). "Forward masking patterns produced by intracochlear electrical stimulation of one and two electrode pairs in the human cochlea," *J. Acoust. Soc. Am.* **86**, 971–980.
- McKay, C. M., and McDermott, H. J. (1996). "The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites," *J. Acoust. Soc. Am.* **100**, 1081–1092.
- Oxenham, A. J., and Dau, T. (2001). "Modulation detection interference: Effects of concurrent and sequential streaming," *J. Acoust. Soc. Am.* **110**, 402–408.
- Richardson, L. M., Busby, P. A., and Clark, G. M. (1998). "Modulation detection interference in cochlear implant subjects," *J. Acoust. Soc. Am.* **104**, 442–452.
- Shannon, R. V. (1983). "Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction," *Hear. Res.* **12**, 1–16.
- Shannon, R. V., Adams, D. D., Ferrel, R. L., Palumbo, R. L., and Grandgenett, M. (1990). "A computer interface for psychophysical and speech research with the Nucleus cochlear implant," *J. Acoust. Soc. Am.* **87**, 905–907.
- Throckmorton, C. S., and Collins, L. M. (1999). "Investigation of the effects of temporal and spatial interactions on speech-recognition skills in cochlear-implant subjects," *J. Acoust. Soc. Am.* **105**, 861–873.
- Tong, Y. C., and Clark, G. M. (1986). "Loudness summation, masking, and temporal interaction for sensations produced by electric stimulation of two sites in the human cochlea," *J. Acoust. Soc. Am.* **79**, 1958–1966.
- Yost, W. A., and Sheft, S. (1989). "Across-critical-band processing of amplitude-modulated tones," *J. Acoust. Soc. Am.* **85**, 848–857.

A practical method of predicting the loudness of complex electrical stimuli

Colette M. McKay,^{a)} Katherine R. Henshall, Rebecca J. Farrell,^{b)} and Hugh J. McDermott
*Department of Otolaryngology, The University of Melbourne, 384-388 Albert Street,
East Melbourne 3002, Australia*

(Received 14 June 2002; revised 13 December 2002; accepted 13 January 2003)

The output of speech processors for multiple-electrode cochlear implants consists of current waveforms with complex temporal and spatial patterns. The majority of existing processors output sequential biphasic current pulses. This paper describes a practical method of calculating loudness estimates for such stimuli, in addition to the relative loudness contributions from different cochlear regions. The method can be used either to manipulate the loudness or levels in existing processing strategies, or to control intensity cues in novel sound processing strategies. The method is based on a loudness model described by McKay *et al.* [J. Acoust. Soc. Am. **110**, 1514–1524 (2001)] with the addition of the simplifying approximation that current pulses falling within a temporal integration window of several milliseconds' duration contribute independently to the overall loudness of the stimulus. Three experiments were carried out with six implantees who use the CI24M device manufactured by Cochlear Ltd. The first experiment validated the simplifying assumption, and allowed loudness growth functions to be calculated for use in the loudness prediction method. The following experiments confirmed the accuracy of the method using multiple-electrode stimuli with various patterns of electrode locations and current levels. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1558378]

PACS numbers: 43.66.Ts, 43.66.Cb, 43.66.Mk [MRL]

I. INTRODUCTION

The output of speech processors for cochlear implants consists of a complex pattern of current waveforms on multiple electrode positions. All present-day processors have one design feature in common. That is, the amplitude of the current on each electrode is determined by the output amplitude of the acoustic filter that is assigned to that electrode. An underlying assumption of this design feature is that the resultant loudness of the electrical output signal will be related in a fixed way to the loudness of the acoustic input signal. However, previous work has shown that this assumption is not fully justified. McKay *et al.* have shown that loudness summation in the spatial domain (across electrodes) (McKay *et al.*, 2001) and in the temporal domain (interpulse intervals) (McKay and McDermott, 1998) produce effects that are inconsistent with this assumption. These effects can lead to inaccuracies in amplitude envelope cues, and inaccurate representation of the relative loudness of acoustic signals which differ in bandwidth (and hence result in concurrent electrical stimulation on differing numbers of electrodes).

This paper presents a practical method for estimating the loudness of complex electrical stimuli consisting of sequential current pulses. As well as being a useful tool for understanding loudness effects in psychophysical stimuli, the method has possible application in sound processing. The method could be used for adapting current levels in real-time in existing processing schemes, to specifically take into consideration loudness summation effects, or to control ampli-

tude cues in novel sound processing strategies. The loudness method steps are described in Sec. II.

In a previous paper (McKay *et al.*, 2001), a model of electrical loudness was proposed which accounted for the pattern of temporal and spatial loudness summation effects seen in psychophysical data. This perceptual model proposed a set of central auditory processing steps which used the pattern of peripheral neural spike activation as the input. The steps in this model are the following.

- (1) The spike activity at each position along the cochlea is summed in a temporal integration window with equivalent rectangular duration (ERD) of approximately 7 ms.
- (2) The resultant excitation density function is transformed to a specific loudness function via a nonlinear relationship, such as a power function with exponent greater than 1.
- (3) The overall loudness is calculated by integrating the specific loudness over position in the cochlea.

Although this model successfully explained the patterns of loudness change observed when temporal and spatial patterns of activation were systematically varied, its practical usefulness is limited in clinical situations. This is because the input to the model is at the level of neural spike initiation, and the spatial pattern of this neural activity for each stimulus pulse, and how it is affected by the parameters of the current pulse and neural refractory factors, are difficult to objectively determine for individual implantees. Although current research is addressing all these factors, it is not yet a practical option to determine them for individual implantees for clinical applications.

A key finding of the previous study, which was important to the development of the practical loudness prediction

^{a)}Electronic mail: colette@unimelb.edu.au

^{b)}Current address: Australian Hearing, 2 Lonsdale St., Melbourne 3000, Australia.

method described in this paper, was that the effect on loudness of the spatial separation of the two active electrodes in dual-electrode stimuli was very small in comparison to the overall size of the loudness summation effect. Although there was always a significant reduction of current required to match the loudness of the dual-electrode stimulus to that of its single-electrode component stimuli, the size of this current reduction remained almost constant as the electrode separation was varied. However, in some stimulus conditions (depending on overall stimulation rate and level), there were small but significant effects of electrode separation which varied in direction. The model proposed two effects of electrode separation which work in opposite directions to largely offset each other: as electrode separation is decreased neural refractory factors work to reduce the overall spike initiation and loudness; at the same time, as electrode separation is decreased, the overall spatial pattern of neural spike initiations becomes more place-specific, and this increases the loudness due to the nonlinear relationship between neural spike activity and specific loudness.

A reasonable approximation for this previous data was that there was no effect of electrode separation, even for very large separations and low current levels. This approximation forms the basis of the practical method for loudness estimation presented below, together with the assumption that individual components of a complex stimulus which do not produce spatially overlapping effects in the cochlea contribute independently to the overall loudness. The latter assumption is based on acoustic models of loudness (e.g., Moore and Glasberg, 1997; Zwicker and Scharf, 1965). Since this part of the acoustic models refers to loudness processing above the cochlear level, it is reasonable to presume that the same process applies in electrical hearing. If non-overlapping neural excitation patterns contribute independently to loudness, and the overall loudness does not change with the degree of overlap, then electrical signals must always behave *as if* the loudness contributions from different electrodes are independent, regardless of the electrode positions or the amount of overlap. It is stressed here that we are not asserting that the loudness contributions *are* independent, but that, given specific conditions in which electrode separation effects offset each other, the approximation of treating the contributions as independent will lead to a reasonable estimate of overall loudness.

In this paper, the above assumption is used to develop a practical loudness estimation method. The accuracy of the estimation method is then tested in three experiments.

II. A PRACTICAL METHOD OF LOUDNESS ESTIMATION

The following proposition is the basis of the proposed practical loudness method: *Given a particular range of stimulus conditions, a running estimate of the loudness of a complex stimulus is given by adding the loudness contribution of each pulse within a small reference time interval. The loudness contribution of each pulse is determined from the current and a loudness growth function for the particular electrode.*

First, a reference time interval is chosen. In order to provide estimates with temporal resolution at least comparable to intensity resolution in normal hearing, the window duration must be of the same order of magnitude as or less than that of the temporal integration window in the acoustic and electric loudness models (an ERD of about 7 ms). In this paper a reference period of 2 ms was chosen.

Second, the loudness growth functions for each electrode are determined. This process is specifically illustrated for six subjects in experiment 1. The principle upon which the loudness growth function is based is that, if the number of (equally loud) pulses per stimulus period is doubled, and the stimulus is then loudness balanced to the original stimulus, then each pulse in the doubled-pulse stimulus contributes half the loudness compared to each pulse in the original stimulus. Thus the reduction in current in the doubled-pulse stimulus compared to that in the original stimulus is the reduction required for halving the effective loudness contribution per pulse. A similar method in acoustic hearing has been used by Buus and Florentine (2002) to estimate the loudness growth slope at low intensities for hearing-impaired listeners. If a number of such loudness-balanced pairs of stimuli are established, using different reference rates and different loudness-levels (including threshold measurements if desired), many estimates of the slope of the loudness growth function [$\log(L)$ versus current] can be obtained at different current levels. These data can then be fitted with a function, $f(i)$, where i is the current parameter. The loudness growth function is then the integral of this slope function over current:

$$\log(L) = \int f(i) di + k. \quad (1)$$

The constant k allows the absolute loudness contribution of a pulse to be determined relative to a reference loudness value. For example, loudness-balanced 500-Hz rate pulse trains on the individual electrodes at a comfortably loud level may be arbitrarily assigned the loudness of 100. These stimuli have only one pulse per reference period (2 ms) and so 100 represents the loudness contribution per pulse for these reference stimuli. The constants k can then be determined independently for each electrode from Eq. (1).

Third, for any arbitrary complex stimulus, the loudness relative to this reference loudness of 100 can then be determined by inserting the current value of each pulse in a 2-ms reference time window into Eq. (1) to find the individual loudness contribution of each pulse, and adding these to find the overall loudness estimate. The loudness contribution for each electrode position in such a stimulus is given by adding together the loudness contributions for all pulses which occurred on that electrode in the time window.

The simplest example of a complex stimulus (and one used in the experiments reported below) is a steady one, in which there is a group of n pulses that are repeated in each repetition period. If one initially supposes that the steady complex stimulus has the same repetition period as the reference period (in the example, 2 ms), then the overall loudness of the complex stimulus (L) is just the sum of the indi-

vidual loudness contributions (L_n) from each of the n pulses in the stimulus period, i.e.,

$$L=\sum_n L_n. \tag{2}$$

More generally, for a steady complex stimulus which has m pulses in a repetition period P_C , which is different from the one in the reference stimuli P_R , the loudness of the complex stimulus (relative to the reference stimuli) will be given by

$$L=P_R/P_C\sum_m L_m. \tag{3}$$

The case of a nonsteady complex stimulus is less straightforward. A normal-hearing listener can monitor intensity changes within an acoustic signal to a degree governed by a temporal integration window similar to that used in the electrical loudness model of McKay *et al.* (2001) (see, for example, Moore *et al.*, 1996). This practical loudness method, however, does not specifically model intensity perception or resolution for dynamic nonperiodic electrical stimuli.

In a cochlear implant speech processor, the short-term spectral shape of an acoustic signal is usually first estimated using a sliding analysis window within which the signal is assumed to be steady, or by using the smoothed output of analog filters. Sequential pulses are then output across the electrode array to represent this short term spectrum, which is updated in each stimulation cycle. For these stimuli, a running estimate of equivalent static loudness over time might be obtained by summing the loudness contributions from each electrode in each stimulus period [as in Eq. (2) or (3)]. These estimates can then be smoothed to estimate the equivalent static loudness of steady portions of the electric signal, or those portions which change only slowly over time.

In the following experiments, the loudness calculation method was tested for steady complex electrical stimuli (i.e., those that repeated the same set of pulses in each fixed stimulus period). Loudness growth functions were measured in six implantees using the method described in the second step above. The method was then applied to predict the loudness of steady complex sounds in two further experiments.

III. EXPERIMENTAL METHODS

A. Subjects and stimuli

Six implantees who use the CI24M implant manufactured by Cochlear Ltd. participated in the experiments. The details of their etiology and implant use are listed in Table I. All were postlingually deaf adults, who had no useful hearing before implantation.

The psychophysical experiments were carried out using specifically designed software. The computer which controlled the experiment was interfaced with the implant via a Spear processor. All stimuli used in these experiments were of 500-ms duration and consisted of trains of biphasic current pulses, with groups of up to eight pulses presented in

TABLE I. Details of etiology and implant use for the subjects who participated in the experiments.

Subject	Length of profound deafness (years)	Implant date	Etiology
S1	2	1999	Chronic infection
S2	2	2000	Trauma and infection
S3	15	1998	Unknown
S4	14	1999	Progressive/Unknown
S5	1	1978/ 1998 (reimplanted)	Trauma
S6	4	1998	Progressive/Unknown

each stimulus period of 2 ms. Within each group of pulses, the time between onset of pulses was 0.25 ms, with the group of pulses placed at the beginning of each 2-ms period (see Fig. 1 for examples from experiment 1). All biphasic pulses used a phase duration of 25 μ s and an interphase gap of 45 μ s. The mode of stimulation was always monopolar. These parameters were chosen to be representative of a temporal pattern produced by a speech processor (for example, an ACE or CIS strategy running at a rate of 500 Hz per electrode). The number of pulses in each stimulus cycle, and the electrodes which activated those pulses were parameters which were varied in the experiments.

The current values and current adjustments are reported in this paper in clinical current level units. For the CI24M implant, each current step is a 0.176 dB change in current. The conversion between current step value, c , and current in microamps, i , is given by the following formula:

$$i=10*175^{c/255}. \tag{4}$$

B. Experiment 1

In this experiment, the effect on loudness of varying the number of pulses in each period was measured. The four stimulus temporal patterns illustrated in Fig. 1 were used. Pattern 1 presented one stimulus pulse per period, and hence is a 500-Hz pulse train on a single electrode. One such stimulus was used for each subject, using the middle active

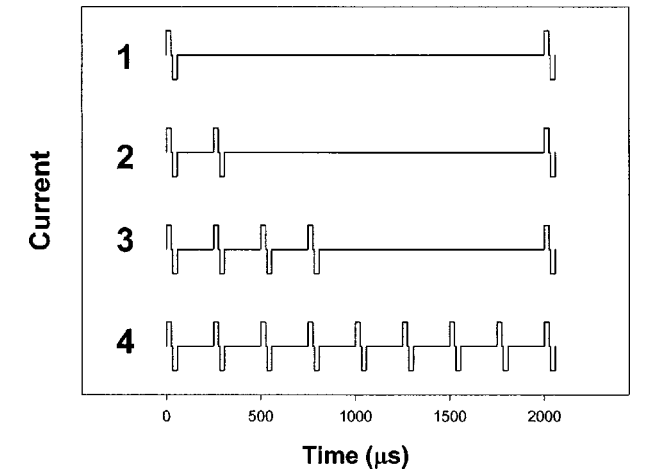


FIG. 1. The temporal patterns of electric current pulses used in experiment 1. The figure shows one period of each of the four stimulus patterns (a 2-ms portion of the total stimulus).

electrode in their array (electrode 11, except for subject S2 whose middle electrode was 14). This was the *reference stimulus* against which all other stimuli were loudness balanced. Pattern 2 presented two pulses in each period. Two such stimuli were constructed, one which presented both pulses on the reference electrode, and one which presented one pulse on each of the electrodes immediately adjacent to the reference electrode (for example, electrodes 10 and 12, giving an electrode separation of 1.5 mm). Pattern 3 presented four pulses in each period. Four such stimuli were constructed, one in which all four pulses were on the reference electrode, and three in which the four pulses were presented on four different electrodes. The four electrodes in the latter three stimuli were chosen to have three different spatial separations of one, two, and four electrode rings (for example, electrodes {9, 10, 11, 12}, {8, 10, 12, 14}, and {4, 8, 12, 16}). Pattern 4 presented eight pulses in each period. Two such stimuli were constructed, one in which all eight pulses were on the reference electrode (making a 4-kHz pulse train), and the other presented the eight pulses on eight different electrodes, each separated by two rings (for example, {4, 6, 8, 10, 12, 14, 16, 18}). The currents used in the multiple-electrode stimuli were those that produced equally loud sensations for the component electrodes at three different overall loudnesses: comfortably loud, a soft level, and threshold loudness. These current values were determined as described below.

Initially, a 500-Hz pulse train was constructed for each of the ten electrodes used in the above stimuli. The current in the reference stimulus (the 500-Hz pulse train on the center electrode) was set to produce a comfortably loud sensation, and the currents in the other nine single-electrode stimuli were adjusted in a loudness balance procedure so that each of these single-electrode stimuli were the same loudness as the reference stimulus. The balancing procedure (used throughout the three experiments) involved the reference and a second stimulus being presented continuously alternating (with 500-ms silent periods between). The subject adjusted the current of the second stimulus using an up/down toggle switch until they were satisfied that the two sounds were equally loud. The procedure was repeated with the subject adjusting the reference stimulus current, and then the two procedures were repeated (four comparisons in all). The average current difference between the two stimuli after these four trials was used to set the current value in the second stimulus. Second, the thresholds of all the single-electrode stimuli were measured using an adaptive four-interval forced choice procedure. Third, the set of currents that produced equal loudness on the different electrodes at a soft loudness were determined by balancing all the electrodes to the reference stimulus set to a current level which was 60% of the range, measured in current steps, between the threshold and comfortably loud levels.

The loudness change produced by adding multiple pulses in the stimulus period (patterns 2, 3, and 4) compared to only a single pulse per period (pattern 1, or the reference stimulus) was investigated by determining the reduction in current required in the multiple-electrode stimuli to make them equally loud to the reference stimulus. The same loud-

ness balance procedure as described above was used. When adjusting the multiple-electrode stimuli, all electrodes were adjusted by the same number of current steps from the predetermined equally loud currents. The thresholds of the multiple-electrode stimuli (patterns 2, 3, and 4) were determined by an adaptive four-interval forced choice procedure. In this procedure, the current on each individual electrode in the multiple-electrode stimuli was adjusted by an equal number of current steps from the single-electrode threshold values determined previously.

A loudness growth function for each electrode was then determined from this data, in the way described in Sec. II. This process is illustrated later in Sec. IV.

C. Experiment 2

In this experiment, a dual-electrode stimulus (temporal pattern 2) was used. The purpose of the experiment was to determine the set of currents on the two electrodes in the dual-electrode stimulus which produced the same overall loudness. That is, unlike experiment 1, the stimuli tested in this experiment had unequal contributions of loudness from the two electrodes to varying degrees. The two electrodes were two mid-array electrodes separated by one ring (for example 11 and 12). First, the currents which produced equally loud sensations for 500-Hz pulse trains on the individual electrodes were determined at threshold, soft, and comfortably loud levels, with the methods described in experiment 1. These two equally loud single-electrode stimuli were used as references to determine the equal-loudness functions for the dual electrode stimuli. Initially, the currents in the dual-electrode stimulus were set to be equal to the currents in the single-electrode stimuli at comfortably loud or soft levels. The currents in the dual-electrode stimulus were then adjusted until the dual-electrode stimulus was equally loud as the single-electrode stimuli. The dual-electrode stimulus was balanced separately to both single-electrode stimuli (eight trials in all) and the results averaged. Similarly, the threshold of the dual-electrode stimulus (with currents initially set to threshold currents of the reference stimuli) was determined. In all adjustments of the dual-electrode stimulus, both currents were adjusted in equal steps, maintaining a constant current-step difference. Next, the current difference on the two electrodes in the dual electrode stimulus was systematically changed and each such stimulus was again balanced to the two single-electrode reference stimuli (or threshold obtained). The two currents in the dual-electrode stimuli were plotted against each other to form three equal-loudness contours.

The loudness prediction method (with loudness growth parameters derived from experiment 1) was then used to predict the individual loudness of each dual-electrode stimulus. The model, if accurate, should predict a constant loudness for all stimuli on an equal-loudness contour, with varying relative loudness contributions from the two electrodes. The single-electrode comfortably loud stimuli in this experiment were designated to have a loudness rating of 100. All of the dual-electrode stimuli which were balanced to these references should also be predicted by the model to have a loud-

ness of 100. Similarly, the dual-electrode soft-level and threshold-level stimuli should be predicted by the model to have constant but lower values of loudness.

D. Experiment 3

This experiment tested the predictions of the model for more arbitrary stimuli. For each subject, nine stimuli were constructed which had from two to eight pulses repeated every stimulus cycle. The temporal patterns were analogous to those shown in Fig. 1, with the addition of patterns with three, five, six, or seven pulses per period. The stimuli were constructed so as to have a range of predicted overall loudness from threshold to comfortably loud. For example, to create a stimulus at the comfortably loud level, the number of pulses per period was arbitrarily selected (between two and eight), and each pulse was arbitrarily assigned an electrode position. An electrode was permitted to have more than one pulse assigned to it in one stimulus period. The relative loudness contribution from the different pulses was also arbitrarily set (for example, for three pulses per period, the contributions of loudness may be 60, 30, and 10 to create a stimulus with overall loudness of 100). The current values required to produce the selected loudness contributions were calculated using the model. The nine stimuli were constructed so that they had a range of electrodes, numbers of pulses, and relative loudness contributions for each subject. In addition, the stimuli were constructed independently for each subject, so that, in effect there were 54 different stimuli. (Subjects S5 and S6 used the same temporal patterns and electrode allocations, but different currents.) The electrode and current parameters of the 54 individual stimuli are listed in Table II.

The predictions of loudness for these 54 stimuli were tested in a loudness-balancing task, in which the current in the reference stimulus from experiment 1 (a 500-Hz pulse train on a central electrode) was adjusted so that the reference stimulus was equally loud to each test stimulus. This current was then compared to the current predicted from the model to produce the same loudness in the reference stimulus as that produced by the test stimulus.

IV. RESULTS

A. Experiment 1

Figure 2 shows the amount of current reduction required to make the multiple-pulse-per-period stimuli the same loudness as the reference single-pulse-per-period stimulus. The plot shows the means and standard errors across all six subjects. Three-way analysis of variance with subject as a random factor, and number of pulses per period and electrode spacing as fixed factors, was carried out at the three levels, using the data for zero and 1.5-mm electrode spacings. A significant effect of electrode spacing (more summation with zero separation) was found at the threshold level only ($p = 0.003$). A significant effect of number of pulses per period ($p < 0.001$) was found at all levels. Three-way analysis of variance with subject as a random factor, and electrode spacing and level as fixed factors, was carried out on the set of data with four pulses per period. There was a significant

TABLE II. Details of the nine stimuli used for each subject in experiment 3. The electrodes (and currents in parentheses) of each pulse within the 2-ms repetition interval are detailed, and the braces surround the set of pulses for each stimulus.

Subject	Electrodes and currents for the arbitrarily constructed stimuli in experiment 3
S1	{6(163) 10(158)}, {8(163) 16(140) 9(152)}, {11(146) 16(128) 8(146) 11(146) 18(128)}, {4(135) 4(135) 4(135) 4(135)}, {14(145) 4(148) 16(129)}, {9(150) 4(138) 18(107) 8(146) 11(104) 11(105) 14(104) 6(151)}, {12(114) 16(135) 6(153)}, {10(130) 14(135)}, {8(118) 10(99) 12(81) 16(81) 9(113)}, {6(104) 9(98) 10(99) 14(104)}.
S2	{15(176) 11(163)}, {9(160) 9(162) 14(167) 11(150)}, {7(157) 9(149) 19(168) 17(168) 15(165) 11(145)}, {13(167) 15(166) 13(148)}, {21(144) 17(155) 19(170) 19(163) 13(149) 15(139) 13(122)}, {13(167) 13(149)}, {14(163) 15(157) 9(145) 7(120)}, {21(140) 14(138) 15(143) 17(151) 11(127)}, {15(108) 13(97) 12(97)}.
S3	{9(187) 16(184) 8(187)}, {6(185) 6(167) 12(181) 14(176)}, {10(184) 11(187)}, {4(132) 9(163) 10(182) 16(182) 14(156) 9(161)}, {9(163) 16(164) 9(175) 16(163) 12(158) 18(170) 12(158) 18(158)}, {11(175) 11(184)}, {8(174) 10(175) 12(158) 14(171)}, {6(148) 16(153) 11(154) 16(153) 12(148)}, {9(131) 18(107) 14(111)}.
S4	{10(185) 8(177) 18(177)}, {11(181) 16(166) 12(181) 18(169) 12(174)}, {9(181) 6(176)}, {8(172) 8(158) 12(175) 18(152) 4(141) 11(171) 9(159) 18(168)}, {6(156) 10(183) 8(151) 14(158)}, {10(171) 10(171) 10(164) 12(179)}, {8(175) 18(160) 4(148)}, {9(168) 6(148) 14(159) 18(156)}, {11(124) 12(121) 6(103) 12(121) 9(117)}.
S5	{8(187) 6(189) 16(191)}, {12(185) 14(179) 12(179) 11(186)}, {4(169) 6(168) 10(188) 12(177) 16(162) 18(152)}, {9(191) 9(183)}, {14(184) 6(183)}, {9(174) 12(161) 6(168) 12(161) 14(161) 14(161) 10(173) 9(174)}, {9(186) 10(175) 11(163)}, {16(166) 9(161) 9(169) 4(168)}, {4(134) 6(129)}.
S6	{8(179) 6(184) 16(182)}, {12(179) 14(168) 12(171) 11(184)}, {4(157) 6(156) 10(183) 12(168) 16(135) 18(150)}, {9(183) 9(177)}, {14(173) 6(177)}, {9(160) 12(152) 6(146) 12(152) 14(146) 14(146) 10(160) 9(160)}, {9(176) 10(164) 11(155)}, {16(135) 9(141) 9(151) 4(157)}, {4(129) 6(116)}.

effect of level ($p < 0.001$) but no significant effect of electrode spacing, or interaction between these two variables.

The data can be summarized with three points:

- (1) Increasing the number of pulses per period systematically increased the amount of current reduction required to maintain a loudness equal to the reference stimulus.
- (2) More current reduction was required at lower levels in the perceptual dynamic range of the subjects.
- (3) The separation of the electrodes used in the multiple-pulse-per-period stimuli had very little impact on the current reduction. The only significant effect was at the threshold level, where the current reduction required was greater for zero electrode separation than for 1.5-mm separation. The size of this effect at threshold was very small (4.5 current steps average) compared to the effects of doubling the numbers of pulses per period (15 current steps average).

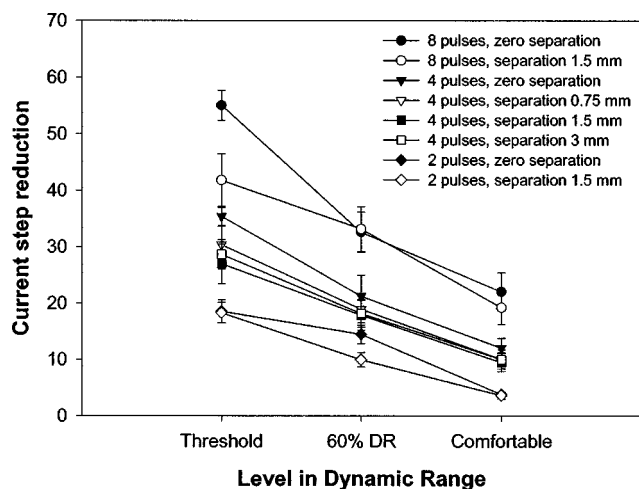


FIG. 2. The results of experiment 1. The vertical axis represents the current reduction (in current steps) for each multiple-pulse-per-period stimulus to make it the same loudness as the reference single-pulse-per-period stimulus. The horizontal axis (category scale) represents the level in the subjects' dynamic ranges at which the measurements were carried out. Each data point shows the mean and standard error for the six subjects. The stimulus temporal pattern and electrode separation are denoted by the symbol type, as specified in the legend.

The latter two points are consistent with and extend the findings for dual-electrode stimuli of McKay *et al.* (2001). In particular, the finding that electrode separation has very little impact on loudness summation in sequential pulsatile stimulation, even when all of the pulses are on a single electrode, supports the validity of the approximation used in the simplified loudness method (that the pulses act as if they contribute essentially independently to the overall loudness, regardless of the electrode position).

For each subject, the effective loudness growth functions were determined by the method described in Sec. II above. Figure 3 illustrates the process for one subject (S5). Each data point in this figure represents the loudness slope calculated from the difference in current between two

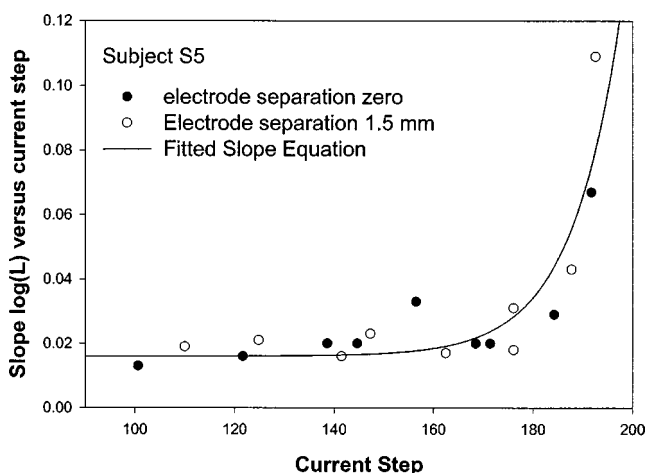


FIG. 3. The estimations for subject S5 of the slope of the log(loudness) versus current step function, derived from the data obtained in experiment 1. The estimations from the single-electrode data and for the multiple-electrode data (with electrode separation of 1.5 mm) are shown separately as different symbols. The solid line is the function fitted to the combined data as described in Eq. (5), the parameters of which are listed in Table III.

TABLE III. The fitted parameters from Eq. (5) using data from experiment 1 for each subject.

Subject	a	c_0	b
S1	0.015	153	10
S2	0.016	169	20
S3	0.021	181	7
S4	0.020	177	19
S5	0.016	185	10
S6	0.017	180	20

loudness-balanced stimuli that only differ by a doubling of the number of pulses per period. As a specific example, the average current levels for one-pulse-per-period and two-pulse-per-period stimuli on electrode 11 at mid-dynamic range were 176 and 161, respectively. Thus, an adjustment of 15 current steps halved the loudness per pulse (0.3 reduction of loudness on a log scale). From this, a slope estimate of 0.02 is derived for the log(loudness) versus current step function at a current level of 168.5. All six subjects produced results of a similar pattern to that shown in Fig. 3, with a relatively constant slope for lower currents and an increasing slope for higher currents. Note that the slope estimates were consistent for similar current levels, regardless of the loudness of the loudness-balanced pair. For example, the four slope estimates in Fig. 3, which are clustered around 0.02 for current step values near 140, are derived from threshold two- and four-pulse stimuli (on a single or multiple electrodes) and medium loud four- and eight-pulse stimuli. Thus the slope estimates are consistent for a particular current value, but across a range of stimuli loudness. This consistency is required for the use of the fitted loudness function in the method to be valid.

It can be seen in Fig. 3 that the slopes derived from stimuli on a single electrode (the filled symbols) and from stimuli with multiple electrodes (open symbols) do not differ significantly. We therefore combined the data to derive a single loudness slope function that applies to all electrodes. This simplifying assumption (that a single rule for loudness slopes applies across the electrode array) reduces the amount of psychophysical data that needs to be collected for individual subjects in order to apply the loudness estimation method. It is not a necessary assumption, however (see Sec. V).

The combined data were well-fit (using non-linear regression) for each subject with a function consisting of a constant plus an exponential:

$$\text{slope} = a + 0.03 * e^{(c-c_0)/b}, \quad (5)$$

where the constants a , c_0 , and b are fitting constants, and c is the current step value. This equation (with $a=0.016$, $c_0=185$, and $b=10$) is shown for subject S5 in Fig. 3. The derived values of these constants for all the subjects are shown in Table III. The loudness growth function can now be derived by integrating Eq. (5) over c :

$$\log(L) = a * c + 0.03 * b * e^{(c-c_0)/b} + k. \quad (6)$$

The constant k is used to set the absolute loudness of a reference stimulus to an appropriate arbitrary value. For ex-

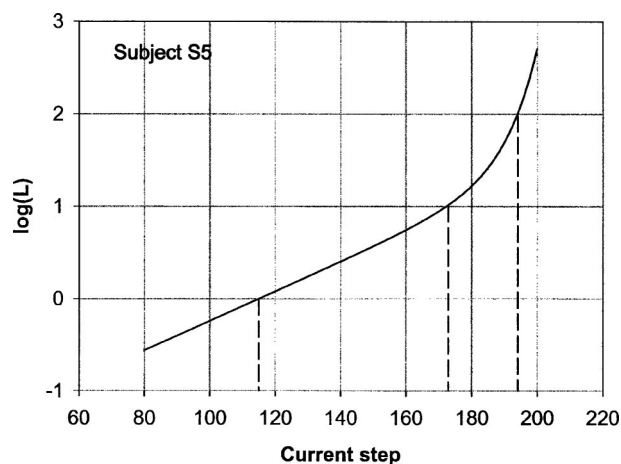


FIG. 4. The loudness growth function for subject S5, derived from the slope function illustrated in Fig. 3. The graph shows the separate loudness contributions predicted for each current pulse within a 2-ms reference time interval, relative to a reference loudness of 100. These loudness contributions, when summed, give an overall loudness for the stimulus (relative to the reference loudness).

ample, in the applications illustrated below, the loudness of the 500-Hz pulse train on the reference electrode at a comfortably loud level was set to 100, leading to a value of k for that electrode. Similarly, a loudness of 100 was set for the 500-Hz pulse trains on the other electrodes which were balanced to the same reference stimulus, leading to different values of k for each electrode (depending on the currents in these stimuli).

The loudness function for subject S5 and electrode 11 is illustrated in Fig. 4. It should be stressed that, since we used a 500-Hz pulse train reference stimulus, the function gives the loudness contribution *per pulse within a 2-ms time interval*. If the stimulus whose loudness is to be predicted was also a 500-Hz pulse train, then the loudness of the stimulus can be read directly from the graph. For the example in Fig. 4, a current of 173 is predicted to have a loudness of 10 (middle drop line), and a current of 116 to give a loudness of 1 (left drop line), compared to the reference stimulus current of 194 (right drop line) which has a loudness of 100. On the other hand, if the stimulus whose loudness is to be predicted had ten pulses with equal current of 173 in each 2-ms period, then each pulse contributes a loudness of 10, making a total loudness prediction of 100 for that stimulus.

B. Experiment 2

The results of loudness-balancing the dual-electrode stimuli with varying relative current on the two electrodes are shown for the six subjects in Fig. 5. The ordinate and abscissa of each data point represent the currents on the two electrodes which together make the dual-electrode stimulus equal in loudness to the two reference single-electrode stimuli. The three symbol types represent the three overall loudness levels.

For each of the data points in Fig. 5, a loudness prediction was made using the loudness growth curves as measured in experiment 1. Since there were two pulses per 2-ms period, one on each electrode, the loudness prediction for these stimuli is the sum of the loudness per pulse for the currents

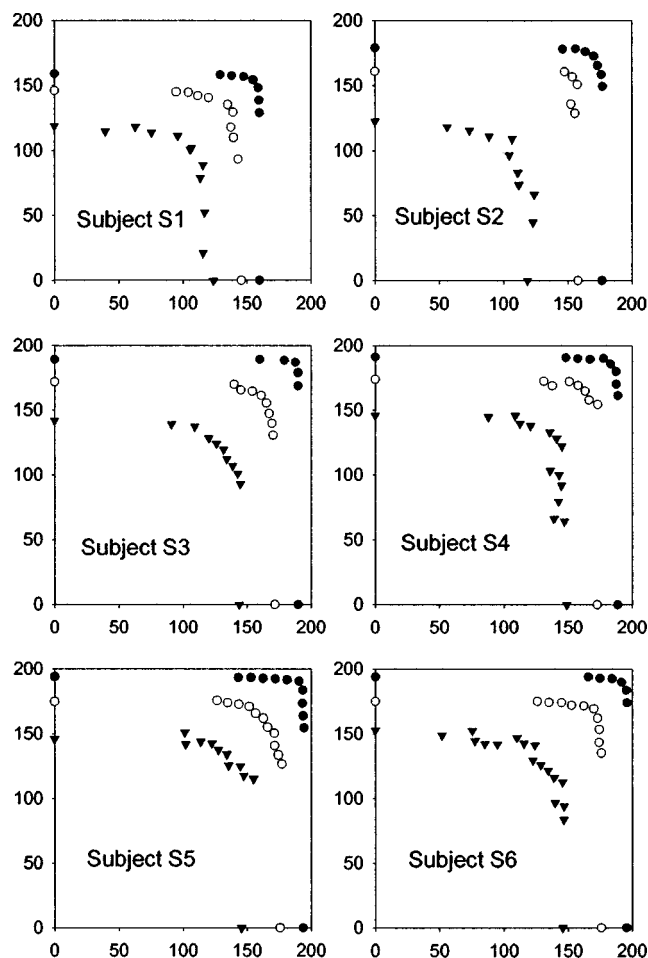


FIG. 5. The results of experiment 2 for each of the six subjects. The two axes represent the current step values on each electrode in a dual-electrode stimulus, so that the overall loudness is maintained at a constant value. Each data point represents a stimulus with a different relative current difference on the two electrodes. Three overall loudnesses are shown, distinguished by symbol type: comfortably loud (\bullet), threshold (\blacktriangledown), and a soft level (60% of the range of current steps on the reference electrodes) (\circ).

on each of the two electrodes. As one moves around the equal-loudness contours in Fig. 5, the relative contribution of each electrode changes systematically, but the total predicted loudness should remain constant. These loudness predictions are shown in Fig. 6. The data points are arranged along the horizontal axis so that the relative loudness contributions from each electrode vary systematically from left to right, as indicated by the arrows. On the left, the overall loudness is only contributed to by electrode 11 (or 14), with insignificant contribution from electrode 12 (or 15). Moving to the right along the row of data points, the contribution from electrode 12 (or 15) increases to 50% of the total loudness in the middle, and to 100% of the total loudness on the right. It can be seen that the predicted total loudness values, which are the sums of the predicted loudness contributions from the two electrodes, remain essentially constant along each equal-loudness contour for each subject.

The average ratio of threshold to comfortable loudness predicted for the six subjects was 1:44. This ratio is equivalent to a loudness change due to an increase of approximately 55 dB in sound level of a pure tone for a normal-

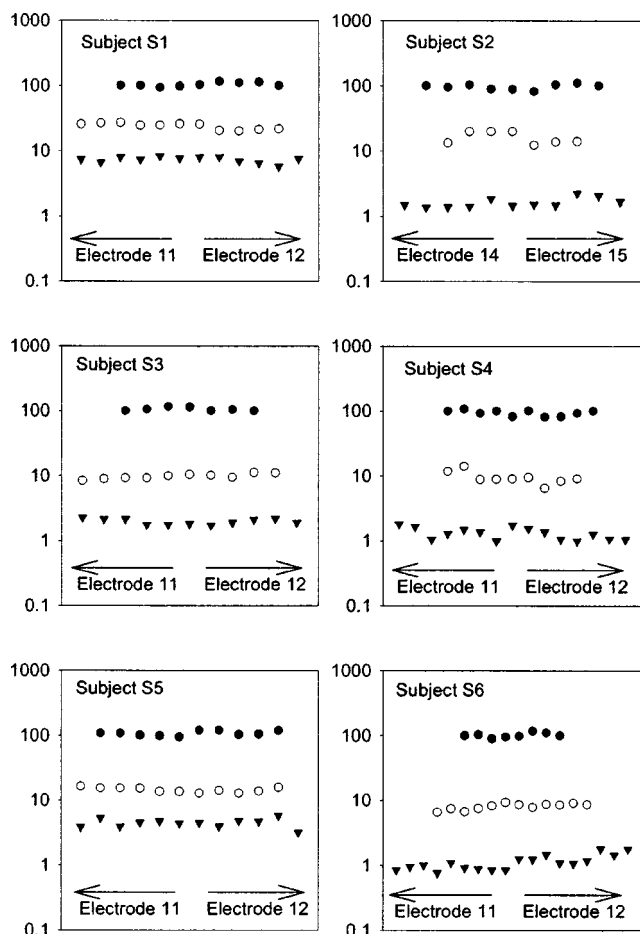


FIG. 6. The prediction of loudness for all of the stimuli in Fig. 5. Loudness (vertical axis) is predicted relative to that of the single-electrode reference stimuli at comfortable loudness (given a value of 100). The stimuli are arranged along the horizontal axis in order of their relative loudness contribution from each electrode in the dual-electrode stimuli.

hearing listener (assuming a 10-dB increase for a doubling of the loudness). The comfortable listening level range of pure tones for normal listeners ranges from 30 to 60 dB SL approximately, where 60 dB SL represents the upper limit of comfortable loudness (Pollack, 1952). Since the electrical comfortably loud levels were set to be near but below each patient's upper comfortable limit, the predicted loudness change between threshold and comfortable loudness from the experiment closely resembles what would be predicted from normal hearing.

C. Experiment 3

The results of experiment 3 are shown in Fig. 7. Each data point represents 1 of the 54 arbitrarily constructed test stimuli. The vertical axis represents the measured current in the reference stimulus when it was equally loud to the test stimulus. The horizontal axis is the current value of the reference stimulus that is predicted by the loudness method to evoke the same loudness as the test stimulus. As can be seen from the figure, the predicted and measured current values are very close. The average difference between measured and predicted values was 0.2 current steps with a standard error of 0.4 current steps. The average absolute error across all test stimuli was 2.3 current steps.

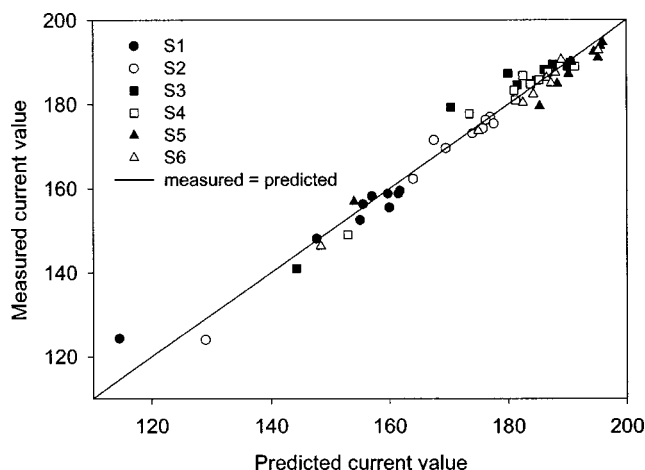


FIG. 7. The results of experiment 3. Each symbol represents 1 of 54 arbitrarily constructed stimuli (detailed in Table II). The symbols refer to different subjects, as shown in the legend. For each stimulus, the measured current level in the equally loud reference stimulus is plotted against the predicted current level. The line represents the measured=predicted situation.

V. DISCUSSION

As detailed in Secs. I and II, the above practical method of loudness prediction is based upon a model of electrical loudness proposed by McKay *et al.* (2001), with the addition of a simplifying approximation. This approximation (that the individual pulses within a set time interval act as if their contributions to the overall loudness are independent) allows a loudness growth function to be derived from a set of loudness-balanced stimuli. The function can then be used to predict the loudness of arbitrary complex stimuli for the same subject.

The form of the loudness growth function that fitted the data for these subjects was a combination of a power and exponential function. The data were consistent with loudness growing as a power function of current for low current values and growing at faster than an exponential rate for high current values. The nonlinear aspect of the growth function (on log scales) has also been noted by other investigators (Chatterjee *et al.*, 2000; Zeng and Shannon, 1992). It should be noted that the loudness growth with current for any one stimulus is represented by only a portion of this loudness function and the portion that applies is dependent on the overall rate of the stimulus. For the example shown in Fig. 4 for subject S5, the dynamic range (between threshold and comfortable) for a 500-Hz pulse train is represented by the portion between the current steps of 154 ($L=5$) and 194 ($L=100$). In this portion, the growth function is quite curved and is not well described by a power function. For a 2500-Hz pulse train, the equivalent perceptual loudness range is predicted to be represented by the portion between $L=20$ and $L=1$. These loudness values give the same overall loudness for the threshold and comfortable stimuli with 500-Hz rate, since there are five pulses in the 2-ms period for the 2500-Hz rate. Thus, reading from the graph, the threshold and comfortable currents for the 2500-Hz stimulus are predicted to be 86 and 182, respectively. For a 2500-Hz pulse train, the predicted loudness growth function for subject S5

approximates a power function more closely and the dynamic range is predicted to be much larger when compared to the 500-Hz rate pulse train.

The subjects studied in these experiments produced very similar loudness growth functions (Table III). The exponent of the power function term (a) had an average value of 0.0175 for the six subjects, which is equivalent to an exponent of approximately 2 when current difference is expressed on a log scale instead of current step scale [see Eq. (5)]. It is interesting to note that the second step in the loudness model proposed by McKay *et al.* (2001) involves a nonlinear transform between neural excitation and specific loudness, and this was approximated by those authors as a power function with exponent of 2. The similarity of the two power exponents suggests that neural excitation itself may grow approximately linearly with current for low current values, and that the power function relating loudness to current at low currents is essentially governed by the perceptual translation of neural spike activity into the loudness domain.

For higher current values, the loudness growth function becomes highly nonlinear (on the log-log scale). The parameter which defines the current at which the loudness slope starts to significantly increase is c_0 (see Table III). The variation among subjects of the kneepoint in the loudness function was correlated with the variations in absolute current values for corresponding stimuli. For example, the highest correlation was between the kneepoint currents and those in the comfortably loud 500-Hz pulse train on the reference electrode ($r^2=0.99$). These correlations suggest that factors such as the degree of neural survival and/or the distance of the electrode from the modiolus may be important in determining the kneepoint. One possible physiological correlate of the kneepoint current is the activation of neural elements central to the cell bodies in the modiolus. The density of neural elements in the internal auditory meatus is greater than in more-peripheral locations, and this would lead to a rapid recruiting of neurons with current, resulting in the fast loudness growth seen above the kneepoint. Further research is needed to understand the relationship of the loudness function to subject-dependent or electrode-dependent differences in patterns of neural survival or electrode placement. In addition, it should be noted that all of the subjects in this experiment had the same type of electrode array and used the same stimulation parameters (mode, pulse duration, and interphase gap). Further research is needed to establish how these stimulation parameters influence the loudness growth patterns.

In these experiments, the possible differences in loudness functions among the electrode positions used for each subject were not examined. Instead, the data in experiment 1 were combined to obtain an average estimate of the slope of the loudness function to be used for all electrodes. This approximation did not lead to significant errors in the loudness predictions made in experiments 2 and 3. However, given the discussion above about the possible correlates of the loudness function parameters, this approximation is not likely to be a good one in every stimulation condition or subject. The subjects in this experiment all used the monopolar mode and had a "straight" electrode array (i.e., not the modiolus-

hugging Contour array). These conditions tend to produce very even current levels for threshold and comfortable loudness across electrodes compared to other stimulation modes (such as bipolar) and other electrode arrays (such as the Contour array). More variation of loudness functions with electrode position would be expected in such conditions. To apply the loudness prediction method in these cases, it may be necessary to use single electrodes or small groups of adjacent electrodes to first obtain loudness functions specific to electrode positions.

In experiments 2 and 3 the loudness of complex stimuli was successfully predicted. These stimuli had overall rates between 500 and 4000 Hz and per-electrode rates between 500 and 2000 Hz. Some stimuli had multiple pulses of different levels in each period on a single electrode. Most stimuli activated multiple electrodes per period, and in many of these the number of pulses per period differed across the activated electrodes. There are likely to be limitations on the stimuli to which the method can be accurately applied, however. The method is based on an approximation that there is no overall effect of electrode separation. The data shown in this article (experiment 1) and in McKay *et al.* (2001) have shown this approximation to be justified for repetition rates between 250 and 1000 Hz and for overall rates between 500 and 4000 Hz, and for monopolar and bipolar stimulation. According to the loudness model, the maintaining of constant loudness for varying electrode separations requires that the effect of neural refractoriness is compensated by the effect of the excitation-to-loudness transformation. In the example discussed in McKay *et al.* (2001) this condition was satisfied if a doubling of rate (on a single electrode) led to the neural excitation per pulse dropping by a factor of 0.7. There is likely to be an upper limit of rate (on a single electrode or at a single cochlear location) for which this may be approximately true. For example, if a pulse train were eliciting neural spikes at the saturation spike rate of the activated neurons, then doubling the rate of stimulation would lead to a halving of neural spikes per pulse, on average. Further research is needed to establish the upper limit of rate for which the loudness prediction method can give accurate estimates. For steady stimuli in which the repetition rate is lower than about 250 Hz, or for nonsteady stimuli, the overall loudness of a stimulus cannot be usefully estimated from a single reference period. Further research is needed to find the best way to model the perception of fine temporal information.

Given the possible limits discussed above, there is no theoretical reason why the loudness method cannot be applied to implanted devices other than the Nucleus device used in these experiments, provided that sequential pulsatile stimulation is used. Differences in electrode design and spacing would produce changes analogous to the use of different modes and electrode spacing in the Nucleus device. That is, these features would be expected to change the degree of overlap of neural excitation patterns elicited by different electrodes. However, this feature has been shown by our experiments to have very little effect on loudness summation (and is the basis for this loudness approximation method).

One stimulation condition in which the approximation of

independent contributions of pulses to loudness will not hold if simultaneous activation of current pulses on more than one electrode. In this situation, if there is overlap of activation areas from the electrode positions, there will be loudness effects due to direct summation of electric fields, which will increase or decrease loudness depending on the relative phases of the pulses. In theory, the method should work reasonably well if these electrode interactions are minimized by increased electrode separation or by appropriate electrode-array design. However, it remains to be tested whether existing array designs and strategies would meet these conditions. The general case of simultaneous stimulation on multiple electrodes could be modeled using the full loudness model of McKay *et al.* (2001) in combination with a separate model to predict the summed pattern of electrical currents along the cochlea and the resultant neural excitation pattern over time.

The clinical importance of the effects of loudness summation is underlined by the results shown in Fig. 2. The degree of loudness summation effects can be illustrated by looking at the current adjustment required to maintain a threshold percept when eight electrodes are activated at their individual thresholds. This stimulus needs a reduction of about 50 current steps to bring the percept back to threshold. This current reduction is the same order of magnitude as the electrical dynamic range of typical stimuli, implying that the sound was quite loud before the current reduction was applied. Such a stimulus may result from the activation of an ACE or CIS speech processing scheme if the acoustic input was a low-level wideband signal. At first thought, it would seem a reasonable solution to reduce the threshold current levels applied in the speech-processor “map” to correct the loudness of wideband signals. However, this would lead to narrow-band acoustic signals being too soft or inaudible. In summary, the loudness of a speech-processor output depends to a large degree on the bandwidth of the acoustic input if loudness summation effects are not specifically accounted for. The conventional way of controlling electrical levels on each electrode (as a direct map from the intensity levels in corresponding band-pass filters) does not account for the effects of input signal bandwidth, and can lead to perceived loudness relationships for implantees which are unnatural.

The development of the practical method of loudness estimation was motivated by the need for improved processing schemes which are able to take into direct consideration the acoustic spectral pattern and its effects on loudness summation in real time. The method described here can theoretically be applied to any processing scheme provided that the processor utilizes sequential current pulses and that the stimulation conditions do not lead to gross violations of the underlying assumptions. Such processing schemes may be based on the present ones (SPEAK, ACE, or CIS). For example, the current levels may be adapted in real time to produce outputs in each cycle which evoke a loudness related to the short-term loudness of the acoustic input. Alternatively, new schemes may be proposed which deviate from the conventional principle of mapping acoustic levels to electric levels independently in separate frequency bands.

Such new schemes would have a major goal of providing loudness relationships which are more natural, and may also improve the perception of amplitude envelope cues in speech.

VI. CONCLUSIONS

In these experiments, a method of predicting loudness for complex electrical stimuli was evaluated in six subjects. In experiment 1, the parameters of the loudness growth function for each subject were derived from loudness-balanced sets of stimuli in which all pulses contributed equally to loudness. In experiments 2 and 3, these parameters were used to successfully predict the loudness in conditions in which the number of pulses per period, the relative currents on different electrodes, and the electrode positions were varied or arbitrarily selected.

The method (described in Sec. II) enables the prediction of the loudness of a steady stimulus, compared to the loudness of an arbitrary reference stimulus. It also enables the relative contribution of different electrodes to the overall loudness to be determined. The data from experiment 1 show that the conventional way of mapping acoustic intensities onto electrical current levels in cochlear implant speech processing can lead to serious misrepresentation of the relative loudness of sounds differing in bandwidth and level. The incorporation of the loudness prediction method into speech processing algorithms may lead to alternative speech processing strategies which limit the effects of loudness summation and improve the perception of sounds by implantees.

ACKNOWLEDGMENTS

This work was supported by the Garnett Passe and Rodney Williams Memorial Foundation and the Human Communication Research Center. The authors wish to thank Peter Blamey for helpful discussions, and the six cochlear implantees who generously provided their time.

- Buus, S., and Florentine, M. (2002). “Growth of loudness in listeners with cochlear hearing losses: recruitment reconsidered,” *J.A.R.O.* **3**, 120–139.
- Chatterjee, M., Fu, Q.-J., and Shannon, R. V. (2000). “Effects of phase duration and electrode separation on loudness growth in cochlear implant listeners,” *J. Acoust. Soc. Am.* **107**, 1637–1644.
- McKay, C. M., and McDermott, H. J. (1998). “Loudness perception with pulsatile electrical stimulation: the effect of interpulse intervals,” *J. Acoust. Soc. Am.* **104**, 1061–1074.
- McKay, C. M., Remine, M. D., and McDermott, H. J. (2001). “Loudness summation for pulsatile electrical stimulation of the cochlea: Effects of rate, electrode separation, level, and mode of stimulation,” *J. Acoust. Soc. Am.* **110**, 1514–1524.
- Moore, B. C., Peters, R. W., and Glasberg, B. R. (1996). “Detection of decrements and increments in sinusoids at high overall levels,” *J. Acoust. Soc. Am.* **99**, 3669–3677.
- Moore, B. G., and Glasberg, B. R. (1997). “A model of loudness perception applied to cochlear hearing loss,” *Aud. Neurosci.* **3**, 289–311.
- Pollack, I. (1952). “Comfortable listening levels for pure tones in quiet and noise,” *J. Acoust. Soc. Am.* **24**, 158–162.
- Zeng, F. G., and Shannon, R. V. (1992). “Loudness balance between electric and acoustic stimulation,” *Hear. Res.* **60**, 231–235.
- Zwicker, E., and Scharf, B. (1965). “A model of loudness summation,” *Psychol. Rev.* **72**, 3–26.

Speech recognition under conditions of frequency-place compression and expansion

Deniz Baskent^{a)}

Department of Biomedical Engineering, University of Southern California, Los Angeles, California 90089

Robert V. Shannon

*Department of Biomedical Engineering, University of Southern California, Los Angeles, California 90089
and House Ear Institute, 2100 West Third Street, Los Angeles, California 90057*

(Received 19 December 2001; accepted for publication 20 December 2002)

In normal acoustic hearing the mapping of acoustic frequency information onto the appropriate cochlear place is a natural biological function, but in cochlear implants it is controlled by the speech processor. The cochlear tonotopic range of the implant is determined by the length and insertion depth of the electrode array. Conventional cochlear implant electrode arrays are designed for an insertion of 25 mm inside the round window and the active electrodes occupy 16 mm, which would place the electrodes in a cochlear region corresponding to an acoustic frequency range of 500–6000 Hz. However, some implant speech processors map an acoustic frequency range from 150 to 10 000 Hz onto these electrodes. While this mapping preserves the entire range of acoustic frequency information, it also results in a compression of the tonotopic pattern of speech information delivered to the brain. The present study measured the effects of such a compression of frequency-to-place mapping on speech recognition using acoustic simulations. Also measured were the effects of an expansion of the frequency-to-place mapping, which produces an expanded representation of speech in the cochlea. Such an expanded representation might improve speech recognition by improving the relative spatial (tonotopic) resolution, like an “acoustic fovea.” Phoneme and sentence recognition was measured as a function of linear (in terms of cochlear distance) frequency-place compression and expansion. These conditions were presented to normal-hearing listeners using a noise-band vocoder, simulating cochlear implant electrodes with different insertion depths and different number of electrode channels. The cochlear tonotopic range was held constant by employing the same noise carrier bands for each condition, while the analysis frequency range was either compressed or expanded relative to the carrier frequency range. For each condition, the result was compared to that of the perfect frequency-place match, where the carrier and the analysis bands were perfectly matched. Speech recognition in the matched conditions was generally better than any condition of frequency-place expansion and compression, even when the matched condition eliminated a considerable amount of acoustic information. This result suggests that speech recognition, at least without training, is dependent on the mapping of acoustic frequency information onto the appropriate cochlear place. © 2003 Acoustical Society of America.

[DOI: 10.1121/1.1558357]

PACS numbers: 43.66.Ts, 43.71.Es, 43.71.Pc [K RK]

I. INTRODUCTION

Speech recognition is adversely affected if the spectral information of speech is presented to an inappropriate cochlear location. For example, several studies have shown a reduction in speech recognition when the speech spectrum is shifted up to higher frequencies (e.g., Daniloff *et al.*, 1968; Fu and Shannon, 1999), or if the frequency-to-cochlear place mapping is distorted nonlinearly (Shannon *et al.*, 1998). Changes in speech recognition as a result of such distortions in the frequency-to-cochlear place mapping are of theoretical interest as an indication of the mechanisms by which speech patterns are stored and retrieved in the central nervous system. In addition, understanding the potential effect of an ap-

propriate frequency-place mapping is of critical importance for the design and programming of cochlear implants and hearing aids. These prosthetic devices can stimulate the peripheral auditory system with a tonotopic pattern of information that is distorted relative to the normal acoustic pattern. Such stimulation raises several questions: In the case of hearing loss, can the patient's speech recognition be improved by adjusting the spectral range of speech to match the frequency region of her residual hearing (Braida *et al.*, 1979; Reed *et al.*, 1983)? Or can the resulting frequency-place distortion actually interfere with speech understanding, as shown in some listeners by Turner *et al.* (1999)? Which of these spectral manipulations are more detrimental and so should be avoided?

Several previous studies have addressed the question of how spectral manipulations affect speech understanding. Fu and Shannon (1999) measured vowel recognition in normal hearing (NH) and cochlear implant (CI) listeners when the

^{a)} Author to whom correspondence should be addressed. Department of Auditory Implants and Perception, House Ear Institute, 2100 W. Third St., Los Angeles, CA 90057. Electronic mail: baskent@usc.edu

acoustic spectral information was mapped to cochlear locations that were shifted apically or basally relative to the acoustic location for that information (for NH listeners) or relative to each listener's clinical frequency-to-electrode map (for CI listeners). They found that vowel recognition was robust to tonotopic shifts up to 3 mm, but dropped significantly for larger shifts. This result matches well with classical studies on frequency shifting (Daniloff *et al.*, 1968; Nagafuchi, 1976; Tiffany and Bennett, 1961).

In a similar study, Dorman *et al.* (1997) measured the effect of a shift in mm between the acoustic frequency range presented and the cochlear range to which it was presented. In the acoustic simulations the analysis filter bands were fixed, and sine wave carriers were shifted in mm along the cochlea relative to the normal acoustic place for that information. Speech recognition performance dropped as the stimulated electrode locations were shifted basally from the normal tonotopic location.

Shannon *et al.* (1998) measured speech recognition under conditions that produced a nonlinear warping of the frequency-place mapping. They used a noise-band vocoder to implement a logarithmic or exponential transformation between acoustic frequency and the normal cochlear place for that frequency. Although four spectral channels of information were presented, listeners' performance with the warped mapping dropped to the same level as that seen with a single-channel noise vocoder. This result suggests that nonlinear frequency-place warping can eliminate listeners' ability to utilize spectral cues.

As an extension to previous studies dealing with frequency-place distortions, the present study explored the effects on speech recognition when the acoustic frequency range delivered is larger or smaller than the normal cochlear range.

Note that neither the present study nor the previous studies discussed above address the potential effects of learning. Research by Rosen *et al.* (1999) showed that, following a short training process, listeners could partially adapt to basalward spectral shifts of as much as 6.5 mm. Another study (Fu *et al.*, 2002) showed significant improvement over the first few days by cochlear implant patients using a 2–4-mm apically shifted frequency-place map, but only little change was observed over the following 3 months. At the end of the 3-month training period consonant and HINT sentence recognition scores were comparable to the baseline scores, while vowel and TIMIT sentence recognition scores were still significantly lower than the baseline scores obtained with the patient's own clinical map before the beginning of the test. In the present experiments the emphasis is on speech pattern recognition without any training. The intention was to test the ability of central pattern recognition mechanisms to accommodate alterations in the peripheral pattern of information with no time to adapt.

Frequency compression has historically been used in an attempt to increase the performance of hearing aids. Most hearing aid users have hearing loss at high frequencies with residual hearing at lower frequencies. To make better use of this residual hearing the spectrum of the speech was lowered and compressed so that the entire speech information was

delivered to the audible range of the patient. The main techniques used for this purpose were slow playback, frequency shifting, vocoding, and zero-crossing-rate division. In terms of frequency-to-place mapping most of these manipulations consisted of a compressed apical cochlear shift. Braida *et al.* (1979) reviewed frequency compression/shifting studies and concluded that frequency lowering did not result in any substantial improvement in speech recognition, and often decreased the performance compared to simple amplification. Reed *et al.* (1983) evaluated the effect of frequency lowering on consonant recognition in a more systematic way, parametrically varying the frequency compression scheme from linear compression to nonlinear frequency-place warping. The results from this study confirmed that frequency lowering did not improve consonant recognition. Linear frequency compression, where the whole frequency range was compressed, resulted in worse consonant recognition scores than a frequency warping compression in which only higher frequencies were spectrally compressed and lowered.

These studies provide insight into the mechanisms used by the central nervous system for storing and retrieving tonotopic patterns of speech. If speech patterns were stored in a "positionally relocatable" fashion, then a tonotopic shift that maintained the overall spatial distribution should still be intelligible. This is clearly not the case, because frequency shifting usually reduces speech recognition. If only the relative order of spectral features were important, then monotonic alterations in the tonotopic pattern would still be intelligible. This is also not the case, because frequency shifting usually reduces speech recognition. The present experiment was designed to further quantify the importance of linear compression or expansion of the tonotopic pattern of information (in cochlear mm). If the central pattern recognition system stores and retrieves information in terms of the relative tonotopic pattern, then it might be able to tolerate a substantial amount of linear compression or expansion.

These issues are not only noteworthy in terms of understanding the relative importance of peripheral vs central pattern recognition for speech, but are of critical importance for the design and fitting of cochlear implants and hearing aids. In a cochlear implant, the electrode array is typically inserted into the scala tympani, reaching a depth of 20–30 mm inside the round window. The average insertion depth from 20 Nucleus implant patients was estimated to be 20 mm by Ketten *et al.* (1998). However, newer electrode designs are intended to achieve array insertions as deep as 30 mm (Gstoettner *et al.*, 1999). The active stimulation range is typically 16 mm in length for Clarion I and Clarion II, and 16.5 mm for Nucleus 22 and Nucleus 24 devices. According to Greenwood's (1990) frequency-to-place equation, and assuming a 35-mm cochlear length in humans, this stimulated cochlear region corresponds to an acoustic frequency range of 500–6000 Hz in humans for a 25-mm insertion depth, and an acoustic frequency range of 1–12 kHz for a 20-mm insertion depth of the electrode array. Present cochlear implants

offer only a limited choice of analysis filters, which cannot be changed individually to match a given patient's electrode location. Most commercial implant speech processors assign a wider fixed acoustic frequency range to this limited cochlear region regardless of the length or the insertion depth of the electrode array. For example, Clarion II assigns an acoustic range of 350–8000 Hz (Advanced Bionics Corporation, 2001) and the default frequency allocation of the Nucleus-22 implant (SPEAK strategy Table 9) assigns a frequency range of 150 Hz–10 kHz to the electrodes (Cochlear Corporation, 1995). The latter acoustic range would normally cover a 25-mm range in the cochlea, specifically from 5 to 30 mm from the round window, rather than the 16.5 mm covered by the electrode array. Thus, mapping the larger acoustic frequency range onto the electrode array results in a compression of the frequency-to-place mapping. In some cochlear implant patients there may also be a tonotopic shift due to the discrepancy between the actual electrode location and the acoustic information assigned. The present experiment evaluated the effect of frequency-place compression on speech recognition in normal-hearing listeners in conditions that simulated two implant electrode insertion depths.

In addition to compression, frequency-place expansion was also evaluated. In this condition, the midfrequency region was expanded in terms of its representation in the cochlea, effectively increasing the sensory resolution within this frequency range. This frequency-place expansion is analogous to the “acoustic fovea” in bats or cetaceans, where a large portion of the cochlea is devoted to the small frequency region used for echolocation (e.g., Echteler *et al.*, 1994). While this type of expansion results in the loss of some acoustic information, the most critical spectral information is presented to a larger cochlear region, resulting in better neural resolution (increased mm/Hz) within that range.

II. METHOD

A. Subjects

Six normal-hearing listeners, aged 26 to 34, participated in the study. All subjects were native speakers of American English and had thresholds better than 20 dB HL at audiometric frequencies between 125 and 8000 Hz. One subject was excluded from the sentence recognition test because she was already familiar with the sentences in the database. Another 32-year-old subject was added to the sentence recognition test to maintain six subjects for each test.

B. Stimuli

The speech recognition tasks consisted of medial vowel and consonant discrimination, and sentence recognition. All stimuli were presented via a loudspeaker in a sound field at 70 dB on an A-weighted scale.

Consonant stimuli were taken from materials recorded by Turner *et al.* (1992, 1999) and Fu *et al.* (1998) at a 44.1-kHz sampling rate. Six presentations (three male and three female talkers) were made of 14 medial consonants /b d f g k m n p s ʃ t θ v z/, presented in an /a/-consonant-/a/ context. Tokens were presented in random order by custom software (Robert, 1998), and subjects were instructed to select the

consonant they heard from the set of 14 possible consonants displayed on the screen. The resulting consonant confusion matrices were analyzed for information received on the production-based categories of voicing, manner, and place of articulation (Miller and Nicely, 1955). Chance performance level for this test was 7.14% correct, and the single-tailed 95% confidence level was 11.77% correct based on a binomial distribution.

Vowel stimuli were taken from the phoneme set recorded by Hillenbrand *et al.* (1995) at a 32-kHz sampling rate. The tokens were presented to the listeners in random order via custom software (Robert, 1998), and subjects were instructed to select the vowel they heard from the set of 12 possible vowels displayed on the screen. Ten presentations (five male and five female talkers) were made of 12 medial vowels, including 10 monophthongs and 2 diphthongs presented in an /h/-vowel-/d/ context (heed, hid, head, had, hod, hawed, hood, who'd, hud, heard, hayed, hoed). Chance level on this test was 8.33% correct, and the single-tailed 95% confidence level was 12.48% correct based on a binomial distribution.

Recognition of words in sentences was measured using the custom software (TIGER SPEECH RECOGNITION SYSTEM developed by Qian-Jie Fu) with Texas Instruments/Massachusetts Institute of Technology (DARPA/TIMIT) corpus of sentence materials (National Institute of Standards and Technology, 1990). The sentences were of moderate-to-hard difficulty, such that individual words were difficult to predict from the context of the sentence, and sentences were spoken by multiple talkers. For each condition, the percent-correct score was acquired for 20 sentences of varying length from each listener. The length of the sentences varied from 3 words to 12 words. The groups of 20 sentences were prepared such that the average word length per sentence was 6–8 words. They were presented without any context information and no sentences were repeated to an individual listener. Sentences were not balanced for difficulty, so 20 sentences were used for each condition to obtain a sample that included varying levels of difficulty. In addition, the order of the presentation of sentences was completely randomized using a random number generator for each subject so that all subjects heard different sentences for different conditions. Consequently differences arising from varying difficulty of the sentences were randomly distributed across different conditions and different subjects. Subjects were asked to repeat what they had heard. The percent-correct score was obtained by counting the percentage of words repeated correctly by the subject.

This study concentrates on effects of frequency-place compression and expansion on speech recognition without any learning effects. We typically observe a short-term adaptation to the test procedure by inexperienced subjects where all scores (regardless of the condition) increased slightly over the first 3 days of the testing. However, the scores remained more or less stable after this initial adaptation period. This was not observed with subjects who already had experience in similar experiments. To minimize any learning effects for the specific experimental conditions no practice was provided on any conditions prior to data collection, even for

TABLE I. Frequency-place mismatch conditions for the 4-channel processor at the simulated 20-mm electrode insertion depth. The name of each condition represents the change in frequency range expressed in mm between the analysis and carrier bands. For each condition the table lists the following information for the analysis bands: cochlear location in mm from the round window, cutoff frequencies for a 4-band processor, and total frequency range. Because the simulated electrode location was fixed, the noise carrier bands covered the frequency range from 1168 to 11 837 Hz in all conditions, and the frequency partition of carrier bands was as shown in the center entry.

Frequency-place mismatch condition	Cochlear location of analysis bands (mm)	Bandpass filter cutoff frequencies for 4 channels (Hz)						Frequency range of analysis bands (Hz)
− 5 mm (expansion)	15–9	2476	3080	3822	4736	5 860		2476–5 860
− 3 mm (expansion)	17–7	1843	2663	3822	5459	7 771		1843–7 771
− 1 mm (expansion)	19–5	1363	2301	3822	6289	10 290		1363–10 290
0 mm (matching)	20–4	1168	2138	3822	6749	11 837		1168–11 837
+ 1 mm (compression)	21–3	999	1985	3822	7243	13 612		999–13 612
+ 3 mm (compression)	23–1	722	1710	3822	8337	17 990		722–17 990
+ 5 mm (compression)	25–0	513	1471	3822	9594	23 762		513–23 762

new subjects, and no feedback was given in any part of the testing. In addition, to reduce a possible adaptation to a particular condition, all conditions with all stimuli were presented to subjects in a completely random order. Therefore, any effects of learning on scores would be distributed across different conditions with different subjects.

C. Signal processing

The noise-band vocoder technique described by Shannon *et al.* (1995) was implemented in MATLAB to generate the stimuli. First, speech materials were bandpass filtered into a number of contiguous frequency bands by sixth-order Butterworth filters. The − 3-dB cutoff frequencies were determined according to Greenwood's (1990) frequency-to-place mapping equation. The exact frequency ranges and cutoff frequencies were determined depending on the specific experimental conditions (see Tables I and II). The speech envelope was extracted from each band by half-wave rectification and low-pass filtering using a third-order Butterworth filter whose output was 3 dB down at a cutoff frequency of 160 Hz. The noise carrier bands representing the

stimulation region in the cochlea were obtained from white noise by sixth-order Butterworth filters where the cutoff frequencies (− 3-dB points) were determined by the condition. The extracted speech envelopes were used to modulate the noise carrier bands, and all modulated noise bands were combined to form the processed speech. The amplitude level of the processed speech was adjusted such that the original and processed tokens had the same overall rms energy.

D. Mapping conditions

Speech materials were processed using 4, 8, or 16 frequency bands. Two different electrode array locations were simulated, representing insertion depths of 20 and 25 mm from the round window. For all expansion and compression conditions the stimulation region covered by the simulated electrode array was fixed at 16 mm (comparable to the typical length of the electrode array for many implant devices). The 20-mm insertion depth condition simulated an electrode array located between 4 and 20 mm from the round window, and the 25-mm insertion depth condition simulated a location between 9 and 25 mm from the round window. Because

TABLE II. Frequency-place mismatch conditions for a 4-channel processor at the simulated 25-mm electrode insertion depth. For each condition the table lists the cochlear locations of the analysis bands, cutoff frequencies of the bandpass filters, and the total analysis frequency range. The noise carrier bands were fixed between 513 and 5860 Hz with the partition shown as in the center of the table. The + 5-mm condition is the one most similar to frequency-to-electrode assignment used in a cochlear implant with a full electrode insertion.

Frequency-place mismatch condition	Cochlear location of analysis bands (mm)	Bandpass filter cutoff frequencies for 4 channels (Hz)						Frequency range of analysis bands (Hz)
− 5 mm (expansion)	20–14	1168	1471	1843	2300	2 864		1168–2 864
− 3 mm (expansion)	22–12	851	1262	1843	2663	3 822		851–3 822
− 1 mm (expansion)	24–10	611	1081	1843	3080	5 085		611–5 085
0 mm (matching)	25–9	513	999	1843	3310	5 860		513–5 860
+ 1 mm (compression)	26–8	428	922	1843	3557	6 750		428–6 750
+ 3 mm (compression)	28–6	290	785	1843	4106	8 944		290–8 944
+ 5 mm (compression)	30–4	184	665	1843	4736	11 837		184–11 837

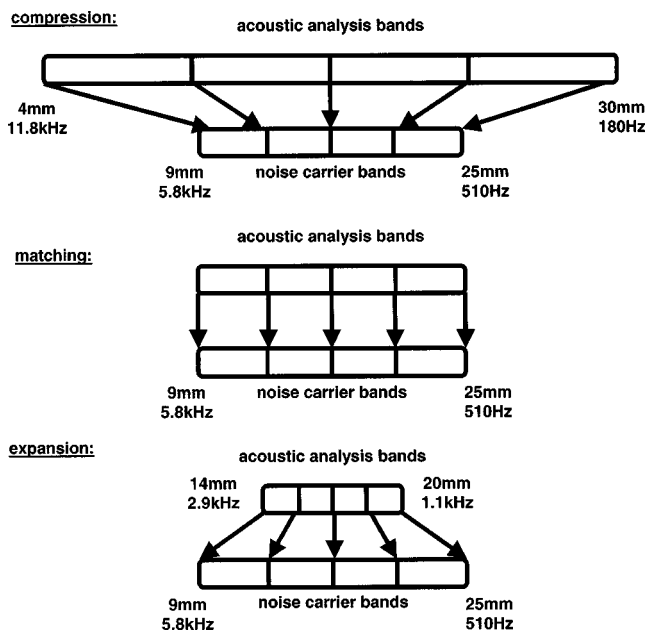


FIG. 1. Frequency-place mapping conditions for 4-channel processor at the simulated 25-mm electrode insertion depth. For this condition the noise carrier bands were fixed (9–25 mm; 510–5800 Hz). The speech envelope was extracted from the analysis bands and used to modulate the noise carrier bands. The top panel shows the +5-mm compression condition schematically: the analysis bands are mapped onto narrower carrier bands. The middle panel shows the 0-mm condition schematically, in which the analysis and carrier bands are matched. The lower panel shows the –5-mm expansion condition schematically, in which analysis bands are mapped onto wider carrier bands.

the stimulation region was fixed at 16 mm, electrode locations were represented by noise bands that were 4 mm wide in terms of cochlear location for the 4-band condition, 2 mm wide for the 8-band condition, and 1 mm wide for the 16-band condition.

In the simulation, the noise carrier bands determine the cochlear location stimulated. The “acoustic analysis bands” are the filters used to process and extract the acoustic envelope information used to modulate the carrier bands (Fig. 1). The distribution of carrier bands was kept fixed for each simulated insertion depth while the analysis bands were systematically altered to create the conditions of frequency-place expansion or compression.

The cutoff frequencies of each analysis band and the related cochlear locations for the 4-band processors are summarized in Table I for the 20-mm simulated insertion depth, and Table II for the 25-mm simulated insertion depth. Filters for the 8- and 16-band conditions were determined by dividing the four bands into two or four equal parts (in mm) and using Greenwood’s (1990) formula to determine the acoustic frequencies for the band edges. The 0-mm condition refers to a perfect match between analysis and carrier bands. Thus, the cochlear location and frequency cutoffs listed for the 0-mm condition specify the fixed locations and frequencies of both analysis and carrier bands. Cochlear locations are all specified in terms of mm from the round window, using Greenwood’s (1990) formula, assuming a 35-mm cochlear length for the human. In the +5-mm condition, the analysis band range was 5 mm wider than the carrier band range on both

apical and basal ends, causing a frequency-place compression of approximately 2 octaves. Similarly, in the –5-mm condition, the analysis band range was 5 mm shorter on each end, causing a frequency-place expansion of about 2 octaves. The +5-mm compression condition in Table I most closely simulates the typical frequency-place compression observed in the standard clinical map of the Nucleus speech processors. Also, +5-mm compression and –5-mm expansion conditions for the simulated 25-mm insertion depth are shown schematically in Fig. 1, as well as the matching case. Here, one can see that the simulated electrode locations are from 9 to 16 mm from the round window and the frequency range used to modulate each noise carrier band is different for every condition.

As the analysis filters were changed from –5-mm to +5-mm conditions, the amount of acoustic information delivered was changed. An important control condition was included to evaluate the effect of the varying amount of acoustic information. In these control conditions the analysis bands and noise carrier bands were always matched in frequency place. These baseline conditions were not intended to simulate any electrode insertion depth or spacing, because in a cochlear implant the electrode position and length are fixed after the implant surgery. Rather, the baseline conditions only assess the effect of changing the overall acoustic bandwidth. Performance in the baseline condition indicates the effect of the gain or loss of acoustic information resulting from the expansion or truncation of the analysis frequency range. The difference in performance between the baseline condition and the compression–expansion condition is due to the frequency-place distortion only.

III. RESULTS

Percent-correct scores for consonants, vowels, and sentences were obtained with 4-, 8-, and 16-channel processors at simulated insertion depths of 20 and 25 mm. In Figs. 2–9, the number of channels increases from 4 to 8 to 16 in the left, middle, and right panels, respectively, of each figure. The average percent-correct scores of six subjects, corrected for chance [$p = 100 * (\text{score} - \text{chance}) / (100 - \text{chance})$], are plotted for consonant and vowel recognition. The average score of six subjects is plotted for sentence recognition. Within each panel the filled symbols present results from the baseline conditions in which analysis and carrier bands were always matched, and the open symbols present results from the experimental conditions in which the frequency-place mapping was expanded or compressed.

A. Consonants

The consonant recognition scores are presented in Fig. 2 for the 20-mm simulated insertion depth and Fig. 3 for the 25-mm simulated insertion depth.

First, consider the data from the baseline conditions (filled symbols), where the analysis and carrier bands were always matched. Consonant recognition increased only slightly as the analysis (and matched carrier) bands were widened (+5-mm baseline condition) from the simulated electrode range. However, as the analysis and carrier bands were narrowed (–5-mm baseline condition) there was a loss

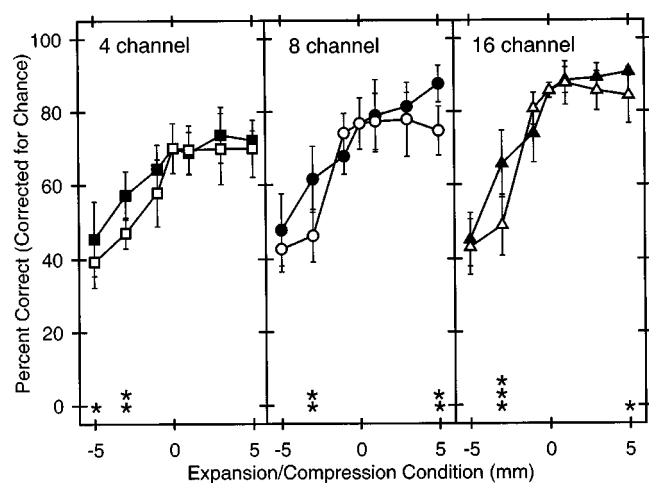


FIG. 2. Consonant recognition percent scores for the simulated 20-mm electrode insertion depth, as a function of compression or expansion in the frequency-place mapping. The number of spectral bands increases from 4 to 8 to 16 in the left, middle, and right panels, respectively. The percent-correct scores represent the average performance of 6 normal-hearing subjects, corrected for chance, and the error bars represent the standard deviation. Filled symbols denote the baseline condition where the carrier bands were always matched to the analysis band frequency range. Open symbols denote the compression–expansion conditions where the carrier bands were fixed and the analysis bandwidth was varied. The stars at the bottom of the figure denote significant differences between the baseline (filled symbols) and mismatch (open symbols) conditions: one star indicates $p < 0.05$, two stars for $p < 0.01$, and three stars for $p < 0.001$.

of acoustic information (due to the reduced acoustical bandwidth) that resulted in lower consonant recognition. This reduction in consonant recognition was more severe in the simulated 20-mm insertion depth condition (Fig. 2) because more low-frequency information was eliminated.

Next, consider the data from the experimental mismatch conditions (open symbols), where the analysis filter frequency range was smaller, equal to, or larger than the simulated electrode length, resulting in frequency-place expansion, matching, or compression, respectively. Note that performance in the mismatched conditions was always equal to or poorer than the baseline conditions. Thus, there are two contributing factors to the reduced performance: (1) the reduction in the amount of information delivered, and (2) the distortion in frequency-place mapping.

A one-way repeated-measures ANOVA test was used to assess the significance of the drop in the performance with

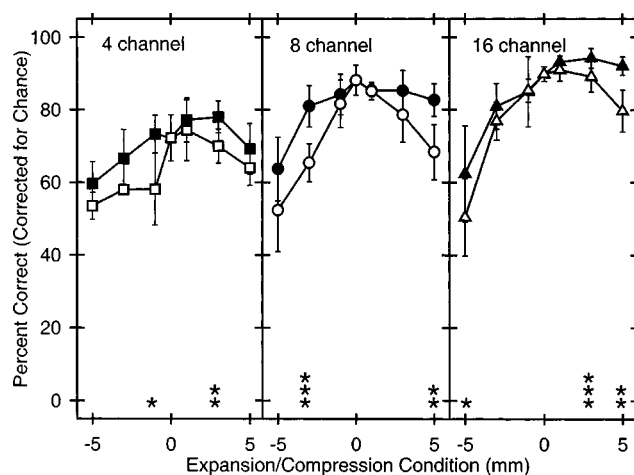


FIG. 3. Consonant recognition, similar to Fig. 2, but for carrier bands simulating a 25-mm insertion depth.

expansion/compression mismatch conditions from the matched condition. Each run included only the expansion or compression percent scores in addition to the 0-mm matched percent score at a particular insertion depth with a specific number of channels. The baseline scores were not included in the ANOVA to isolate the effect of expansion or compression only on speech recognition. The baseline condition was compared to the corresponding mismatch condition with paired t-tests.

The analysis revealed that all frequency expansion conditions reduced performance significantly from the 0-mm matched condition. Corresponding F and p values are shown in Table III. A substantial amount of this drop was due to the loss of acoustic information, as indicated by the filled symbols. An additional drop was observed for some conditions when the frequency-place mapping was expanded (open symbols).

A paired t-test analysis compared the baseline (filled symbols) and expansion (open symbols) performance. Conditions that are significantly different are indicated with stars on the bottom of the figure. One star denotes significant difference with a p value of $p < 0.05$, two stars denote $p < 0.01$, and three stars denote $p < 0.001$. The analysis shows a significant difference for -3 -mm expansion for most processors at both insertion depths. The difference was generally not significant for -5 -mm expansion, possibly because the

TABLE III. F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for consonant recognition at 20-mm and 25-mm simulated insertion depths.

Expansion at 20-mm insertion	$F(3,15)$	p	Compression at 20-mm insertion	$F(3,15)$	p
4 channel	54.60	<0.001	4 channel	0.01	1
8 channel	89.05	<0.001	8 channel	0.23	0.87
16 channel	117.79	<0.001	16 channel	0.66	0.59
Expansion at 25-mm insertion	$F(3,15)$	p	Compression at 25-mm insertion	$F(3,15)$	p
4 channel	15.10	<0.001	4 channel	7.40	<0.01
8 channel	62.60	<0.001	8 channel	26.99	<0.001
16 channel	90.26	<0.001	16 channel	15.38	<0.001

performance was limited by a floor effect (previous studies have found 30%–40%-correct consonant recognition even for single-channel noise processors, indicating that this level of performance is possible using only temporal cues: Van Tasell *et al.*, 1987; Shannon *et al.*, 1995). The difference was also not significant for –1-mm expansion, which produced performance similar to the 0-mm matched condition.

Thus, even though the cochlear tonotopic representation of the spectral information was expanded, resulting in improved spectral resolution within the pattern, performance was poorer than the matched condition. This result suggests that improved resolution in the spectral domain does not necessarily improve speech recognition, probably because the information is not in the appropriate cochlear place. In the present experiment, which did not provide any practice or time to accommodate to the new mapping, expansion in frequency-place mapping always resulted in poorer consonant recognition. It is possible that additional practice with the experimental processors would have resulted in improved performance.

Frequency-place compression did not have a significant effect on consonant recognition for the 20-mm simulated insertion depth, yet there was significant decrease from the matched 0-mm condition (10%–20% drop at +5-mm compression) in performance for the 25-mm simulated insertion depth (as shown in Table III). For the extreme condition of frequency-place compression (+5 mm) there was a significant reduction in performance of 10%–15% relative to the baseline condition for 8- and 16-channel processors.

There was no clear difference between the pattern of results for 4, 8, and 16 channels other than the overall improvement in performance with more channels.

B. Consonant feature analysis

Information transmitted on the consonant features of place, manner, and voicing is plotted in Figs. 4 and 5 for 20-mm and 25-mm simulated insertion depths, respectively. Within each figure the top, middle, and lower panels show the percent of information transmitted on place, manner, and voicing, respectively.

Information transmission percent scores are calculated from the confusion matrices where the diagonal entries are the numbers of correct answers and the off-diagonals are the confusions. A measure for the transmission of information is the covariance between the input and the output, as given by

$$\text{Cov}(x,y) = - \sum_{i,j} p_{ij} \log \frac{p_i p_j}{p_{ij}},$$

where p_i , p_j , p_{ij} are directly related to the frequencies of occurrences of stimulus i , response j , and joint occurrence of stimulus i and response j , respectively. Next, the covariance is converted to the information transmission percent score by normalization such that it yields 100% when the subject identifies all phonemes accurately (Miller and Nicely, 1955).

Note that information received on manner at 20-mm simulated insertion and on voicing at both insertion depths was not affected by frequency-place compression, and man-

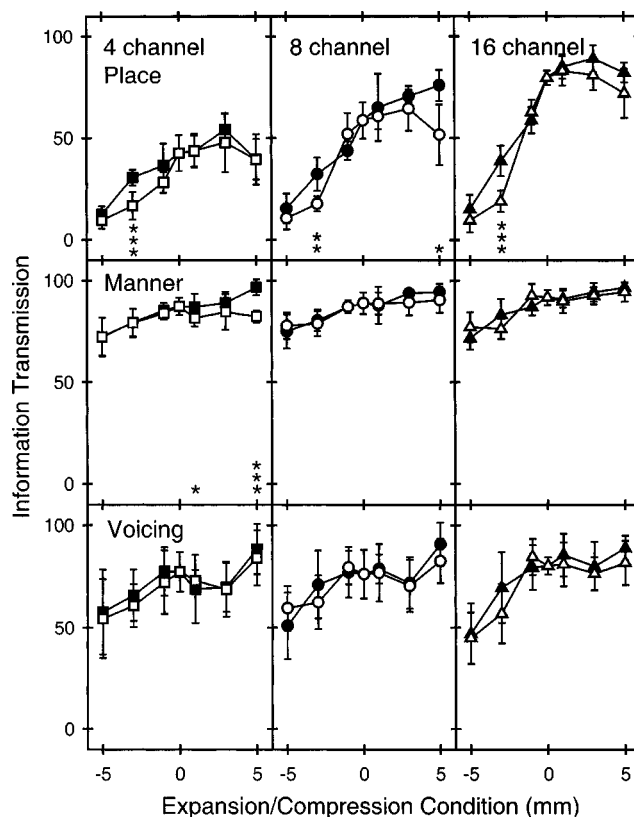


FIG. 4. Information transmission percent scores for consonant features at 20-mm simulated insertion depth as a function of frequency-place mismatch conditions. The features are grouped into production-based categories of place, manner, and voicing, in the top, middle, and bottom rows, respectively. The number of spectral bands increases from 4 to 8 to 16 in the left, middle, and right panels, respectively.

ner information received at 25-mm simulated insertion depth was only slightly affected. Yet, both manner and voicing information transmission scores dropped significantly with expansion (see Table IV for corresponding F and p values). Also, the compression/expansion mismatch scores were similar to baseline scores for manner and voicing, implying that these features are mostly affected by the bandwidth of acoustic information. The performance for both features was similar for different number of channels and for the two simulated insertion depths.

In contrast, place information was strongly affected by expansion at both simulated insertion depths, and by compression at 25-mm simulated insertion depth (Table IV). The overall pattern of the performance changed from 20-mm insertion to 25-mm insertion, and the information received increased with increasing number of channels. This pattern is very similar to the consonant recognition results of Figs. 2 and 3, and therefore it appears that the overall shape of consonant recognition performance was primarily determined by the loss of place information. This observation agrees well with previous studies, which found that manner and voicing cues are more robust to spectral manipulations (Shannon *et al.*, 1998).

C. Vowels

Vowel recognition scores for simulated 20-mm and 25-mm insertion depths are presented in Figs. 6 and 7, re-

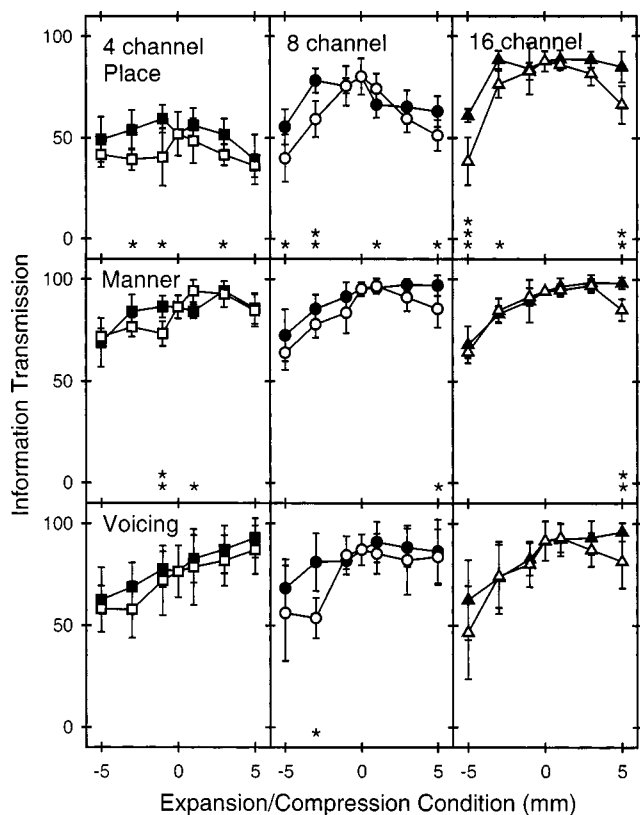


FIG. 5. Consonant feature information transmission, similar to Fig. 4, but for carrier bands simulating a 25-mm insertion depth.

spectively. Vowel recognition was much more strongly affected by frequency-place mismatch than consonant recognition; performance decreased significantly from the matched 0-mm condition with both expansion and compression mis-

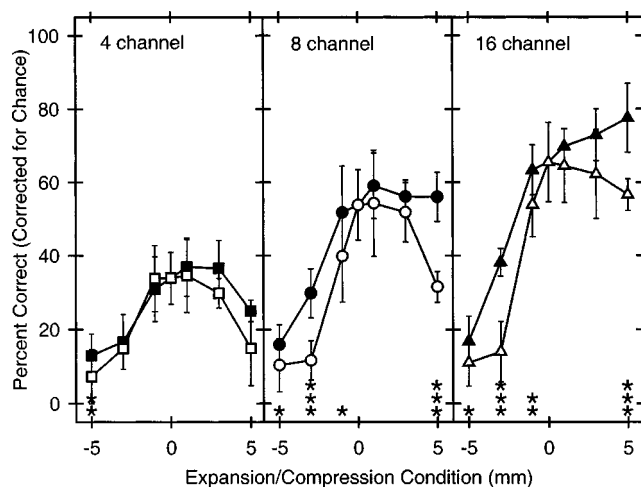


FIG. 6. Vowel recognition percent scores for carrier bands simulating a 20-mm insertion depth, as a function of compression or expansion in the frequency-place mapping. The number of spectral bands increases from 4 to 8 to 16 in the left, middle, and right panels, respectively. The percent-correct scores represent the average performance of 6 normal-hearing subjects, corrected for chance, and the error bars represent the standard deviation. Filled symbols denote the baseline condition where the carrier bands were always matched to the analysis band range. Open symbols denote the compression–expansion conditions where the carrier bands were fixed and the analysis bandwidth was varied. The stars at the bottom of the figure denote significant differences between the baseline (filled symbols) and mismatch (open symbols) conditions: one star indicates $p < 0.05$, two stars for $p < 0.01$, and three stars for $p < 0.001$.

match conditions (F and p values obtained by one-way repeated-measures ANOVA are given in Table V). Overall performance improved as more spectral bands were used but the pattern of results was similar for 4, 8, and 16 bands. For 8 and 16 channels, vowel recognition decreased as the analy-

TABLE IV. F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for consonant feature recognition at 20-mm and 25-mm simulated insertion depths.

Expansion at 20-mm insertion	Place		Manner		Voicing	
	$F(3,15)$	p	$F(3,15)$	p	$F(3,15)$	p
4 channel	63.42	<0.001	9.09	<0.01	6.90	<0.01
8 channel	63.01	<0.001	10.18	<0.001	12.97	<0.001
16 channel	347.90	<0.001	21.15	<0.001	25.77	<0.001
Compression at 20-mm insertion						
4 channel	1.55	0.24	1.51	0.25	4.47	<0.05
8 channel	1.48	0.26	0.16	0.92	3.52	<0.05
16 channel	2.49	0.10	2.43	0.11	0.62	0.62
Expansion at 25-mm insertion						
4 channel	2.19	0.13	8.98	<0.01	6.10	<0.01
8 channel	38.92	<0.001	24.90	<0.001	16.32	<0.001
16 channel	55.35	<0.001	72.88	<0.001	26.62	<0.001
Compression at 25-mm insertion						
4 channel	7.46	<0.01	3.77	<0.05	1.19	0.35
8 channel	25.38	<0.001	5.06	<0.05	0.75	0.54
16 channel	15.59	<0.001	15.43	<0.001	6.04	<0.01

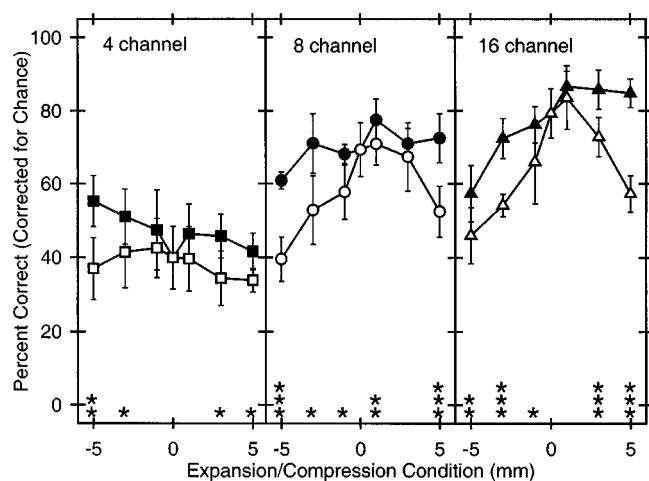


FIG. 7. Vowel recognition percent scores for noise carrier bands simulating a 25-mm electrode insertion depth, as a function of compression or expansion in the frequency-place mapping.

sis frequency range was reduced (baseline condition going from +5 mm to -5 mm), and a further drop in recognition was seen for both frequency-place compression and expansion as a result of the mismatch.

As in consonant recognition the stars at the bottom of each panel indicate significant difference between baseline (filled symbols) and mismatch (open symbols) conditions determined by paired t-test.

For 8- and 16-channel processors both frequency-place expansion of -5 mm and compression of +5 mm resulted in a 20%–30% drop in recognition compared to the matched condition (the flat performance with the 4-channel processor at 25-mm insertion depth may have been limited by the overall poor level of performance). Note that the Nucleus cochlear implant processor mentioned above typically uses a frequency-place assignment that is similar to the +5-mm compression condition, which produced a significant reduction in vowel recognition.

Although a reduction in the acoustic frequency range normally causes a drop in performance, Fig. 7 shows an improvement in vowel recognition for the baseline expansion condition for 4-channel processor at 25-mm insertion [$F(6,30)=4.13, p<0.01$]. This improvement could be due either to an increase in resolution or to a better frequency partition of the analysis bands.

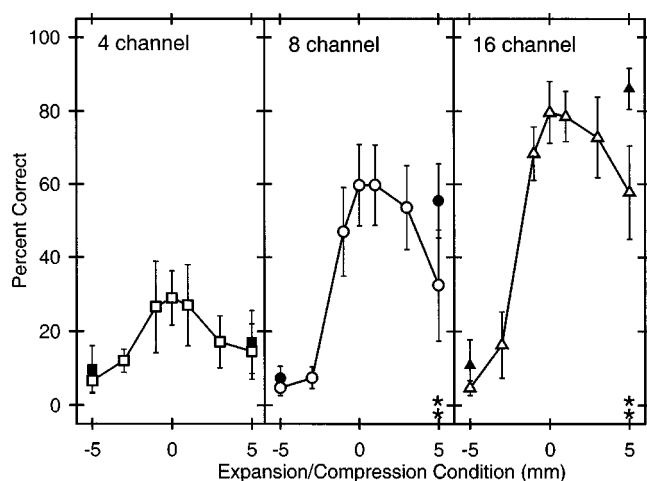


FIG. 8. Sentence recognition percent scores for noise carrier bands simulating a 20-mm electrode insertion depth, as a function of compression or expansion in the frequency-place mapping. The number of channels increases from 4 to 8 to 16 in the left, middle, and right panels, respectively. The percent-correct scores represent the average performance of 6 normal-hearing subjects, and the error bars represent the standard deviation. Filled symbols denote the baseline condition where the carrier bands were always matched to the analysis band frequency range. Open symbols denote the compression-expansion conditions where the carrier bands were fixed and the analysis bandwidth was varied.

Note that there was probably a floor effect for the expansion conditions with the simulated 20-mm insertion, where increasing the number of channels did not increase the intelligibility (Fig. 6, leftmost data point of each panel). Both -5-mm and -3-mm expansion results were close to chance level, which might be due to the loss of all low-frequency information below 1850 Hz (-3-mm condition) or below 2476 Hz (-5-mm condition).

D. Sentences

The percentage of words recognized in TIMIT sentences is presented in Figs. 8 and 9 for 20-mm and 25-mm simulated insertion depths, respectively. Due to the limited number of sentence sets available, the matched baseline performance was measured only for extreme mismatch conditions (-5-mm expansion and +5-mm compression). For all numbers of channels and both simulated insertion depths the best performance was obtained when the analysis and carrier bands were matched.

TABLE V. F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for vowel recognition at 20-mm and 25-mm simulated insertion depths.

Expansion at 20-mm insertion	$F(3,15)$	p	Compression at 20-mm insertion	$F(3,15)$	p
4 channel	40.26	<0.001	4 channel	14.28	<0.001
8 channel	78.55	<0.001	8 channel	14.86	<0.001
16 channel	235.08	<0.001	16 channel	2.05	0.15
Expansion at 25-mm insertion	$F(3,15)$	p	Compression at 25-mm insertion	$F(3,15)$	p
4 channel	1.29	0.31	4 channel	1.95	0.17
8 channel	32.67	<0.001	8 channel	22.31	<0.001
16 channel	39.54	<0.001	16 channel	32.94	<0.001

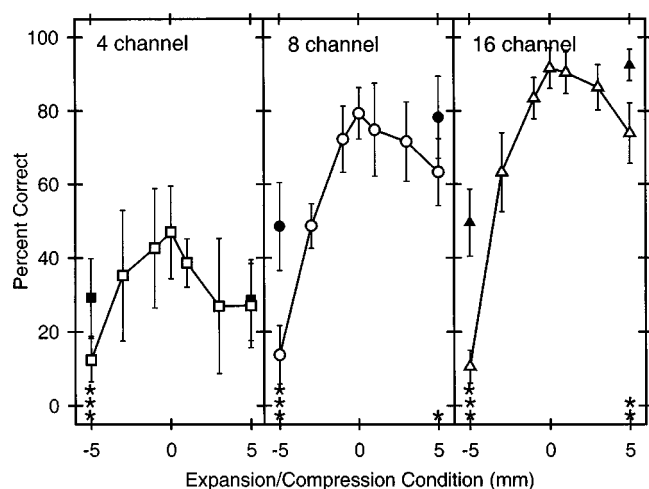


FIG. 9. Sentence recognition percent scores for noise carrier bands simulating a 25-mm electrode insertion depth, as a function of compression or expansion in the frequency-place mapping.

Note that the drop in performance for frequency-place expansion (-5 mm) was dramatic compared to the matched condition (Table VI). Performance at the -5 -mm condition drops to 5% correct for the 20-mm simulated insertion depth, and 15% for 25-mm insertion depth for all three spectral resolutions. For the 16-channel processor this drop was 75–80 percentage points. Although much of this loss was due to the information loss (filled symbols), there was an additional 20–40-point drop in performance due to the frequency-place expansion at the 25-mm simulated insertion depth.

A smaller drop in performance was observed for frequency-place compression. There was a drop of 15–25 percentage points from the 0-mm baseline condition to the $+5$ -mm baseline condition, even though the overall frequency range increased by about 2 octaves. There was an additional drop of 20–30 percentage points from the $+5$ -mm baseline to the $+5$ -mm compression condition. This $+5$ -mm compression condition is similar to the typical frequency-place assignment used in Nucleus cochlear implants.

IV. DISCUSSION

The best speech recognition performance was always observed in conditions where frequency information was

matched to its normal acoustic cochlear place. In most conditions, altering the frequency-place mapping by either compression or expansion resulted in poorer speech recognition.

A. Implications for cochlear implants

The present study observed a reduction of approximately 20 percentage points in vowel and sentence recognition when a frequency range was compressively mapped onto a cochlear range that was smaller by 2 octaves ($+5$ -mm compression condition). Even though a broader frequency range of acoustic information is presented in this condition, performance was reduced due to the distortion in the frequency-place assignment. This compressive frequency-to-place mapping is similar to the mapping used in Nucleus cochlear implant systems, in which the acoustic frequency range of 150 Hz to 10 kHz is typically mapped onto electrodes that occupy the cochlear locations that normally respond to an acoustic range of only 500–6000 Hz. This result implies that speech recognition performance in cochlear implants might be improved by as much as 20% if the frequency range for each electrode could be mapped according to the normal acoustic characteristic frequency of that cochlear location.

How can a cochlear implant speech processor be adjusted to achieve the best mapping of frequency information onto the most appropriate cochlear place, given the variability in cochlear length and electrode insertion depth across patients? In implant listeners there is uncertainty in the exact location of the electrodes and further uncertainty as to the location of the stimulated neurons. Recent advances in imaging technology allow sufficient resolution to evaluate the depth of electrode insertion and to detect the presence of any kinks or abnormalities in the electrode carrier (Ketten *et al.*, 1998). However, these imaging procedures are costly, time consuming, deliver large doses of radiation, and still may not provide all of the necessary information. For example, even knowledge of the exact cochlear location of an electrode is no guarantee that the stimulation of neurons is actually occurring at that location. The actual stimulation location can be affected by the pattern of local nerve survival or by unusual current pathways due to bone growth and fibrous blockage. In addition, the actual site of stimulation may be in the spiral ganglion, whereas Greenwood's formula holds for stimulation at the basilar membrane. These factors produce additional uncertainties regarding the appropriate frequency-place mapping in implant patients. The data presented above

TABLE VI. F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for sentence recognition at 20-mm and 25-mm simulated insertion depths.

Expansion at 20-mm insertion	$F(3,15)$	p	Compression at 20-mm insertion	$F(3,15)$	p
4 channel	18.53	<0.001	4 channel	6.87	<0.01
8 channel	66.88	<0.001	8 channel	8.76	<0.01
16 channel	256.22	<0.001	16 channel	13.91	<0.001
Expansion at 25-mm insertion	$F(3,15)$	p	Compression at 25-mm insertion	$F(3,15)$	p
4 channel	11.64	<0.001	4 channel	4.29	<0.05
8 channel	147.89	<0.001	8 channel	3.59	<0.05
16 channel	233.12	<0.001	16 channel	22.17	<0.001

control for these factors by making the measurements in normal-hearing listeners, in whom the actual stimulation locations can be controlled, at least within the constraints of the normal acoustic spread of excitation. Due to these uncertainties, the optimal frequency-place alignment in an implant patient might best be determined functionally, by adjusting the frequency range and distribution across the implanted electrode array to achieve the best performance. A simple optimizing algorithm, paired with a sensitive phonetic contrast test, could provide an efficient method for converging on an optimal frequency-place mapping for an individual patient, without the costs and risks of x rays and CT scans. The present results may help to define the inherent trade-offs between electrode depth, number of electrodes, and frequency range.

Another uncertainty comes from anatomical and geometrical issues regarding the stimulation of deeper turns of the cochlea. Electrodes in cochlear implants do not always reside between 9 and 25 mm or between 4 and 20 mm inside the round window, the two conditions simulated in this study. The latest generation of CI electrodes, such as Clarion HiFocus, Nucleus Contour, or Med-El Combi40+ offers deeper insertion, possibly up to 30 mm. Even though these specially designed electrodes make it possible to reach more apical locations inside the cochlea, it is unknown if it is possible to stimulate the spiral ganglia corresponding to low frequencies. Cell bodies of the spiral ganglia from the apical turn of the cochlea are located in the modiolus of the cochlear middle turn, and so are physically (and presumably electrically) closer to electrodes in the middle turn than to the medial wall of the cochlea in the apical turn. Studies of pitch have shown little change in pitch with electrode location for electrodes that were deeply inserted into the apical turn, suggesting that there may be a point of diminishing returns in terms of electrode insertion depth (Cohen *et al.*, 1996). And, even with the new electrode designs, the array cannot always be fully inserted due to cochlear ossification or otosclerosis. Thus, the actual location of the implanted electrode is difficult to determine accurately, and the location of the neurons actually stimulated by each electrode adds a further layer of uncertainty.

New studies are presently being conducted in patients with substantial residual hearing, where the electrode array is only shallowly inserted in an attempt to preserve any residual acoustic hearing. In these special cases it may be particularly important to assign the appropriate frequency-place mapping, because the electrically stimulated hearing must combine with residual acoustic hearing. Indeed, preliminary results (Turner and Gantz, 2001; Brill *et al.*, 2001) suggest that combined electric and acoustic hearing is best when the electrodes in the basal turn receive high-frequency information that is matched to their tonotopic location.

B. Trade-off between spectral resolution and overall bandwidth

Some of the present results indicate a trade-off between spectral resolution (number of bands) and overall bandwidth. For a given number of bands there may be an optimal bandwidth—too small a bandwidth would discard too much

important speech information, and too wide a bandwidth would increase the frequency range of each band, reducing the relative resolution. Consider the baseline conditions (filled symbols in Figs. 2–9). In these conditions the analysis bands were always matched in frequency to the carrier bands, while the number of bands was held constant. For the 16-band conditions, a larger bandwidth generally produced better performance whereas for 8 bands the performance was generally unchanged as the overall bandwidth was increased relative to the standard matched condition (0 mm). However, when there were only 4 bands of spectral resolution available, there was a complex interaction between the bandwidth and spectral resolution. In many cases, performance dropped both when the bandwidth was increased, and decreased relative to the standard matched condition. As bandwidth decreased the relative spectral resolution increased, but this was not sufficient to offset the loss of information. As bandwidth increased the additional spectral information was offset by the loss of relative spectral resolution (e.g., Figs. 6, 8, and 9: compare 4-channel 0-mm and +5-mm baseline conditions), resulting in poorer performance in spite of the larger bandwidth.

An expansion in the frequency-place mapping could theoretically improve speech recognition by spreading out the critical speech spectral region to a larger range in the cochlea. Echolocating animals have evolved such a strategy to provide better tonotopic resolution in the small frequency region of their echo signal. However, in the present study such expansion conditions mostly resulted in poorer speech recognition. The only exception was for the 4-channel processor with a 25-mm simulated insertion depth. This was likely due to an artifact of band edge placement: the 0-mm condition contained no band division in between 999 and 1843 Hz (see Table II), while the –5-mm expansion condition contained a band division at 1472 Hz, which is an important frequency for distinguishing high from low second formant frequencies. In this particular condition, the contribution from this better frequency partition might have compensated for the loss of bandwidth. With a limited number of bands, the placement of the frequency divisions appears to be more important than the overall frequency range. Alternatively, the slight improvement in performance in the expansion condition could be due to the improved resolution in this condition. The small frequency range of 1168–2864 Hz was represented across a larger cochlear region that would have normally responded to a range of 513–5860 Hz. This expansive mapping may have helped recognition by stimulating a larger neural population with information from the smaller frequency range. Whichever explanation is correct (better band partition or expanded representation) the same effects were not observed with more than 4 bands.

C. Potential effects of learning

One aspect of speech pattern recognition not addressed by the present study is the potential effect of learning. Recent work has demonstrated that NH subjects listening to simulations of cochlear implants can improve their scores on speech recognition with only a modest amount of practice (Rosen *et al.*, 1999). Rosen *et al.* used noise-band vocoders

in which the frequency-place mapping was shifted basally by as much as 6.5 mm. Listeners improved significantly in their ability to recognize phonemes and words with these shifted representations after only a few hours of training. However, their performance after this limited amount of training was still far poorer than their recognition with the unshifted speech. It is not clear if further training would allow complete recovery of performance to the unshifted levels. Fu *et al.* (2002) measured speech recognition in three cochlear implant listeners after a 3-mm apical shift in the frequency-to-electrode assignments. Initially, speech recognition was reduced dramatically. After 10 days of everyday experience with the shifted map there was a significant improvement in recognition, but then only little further improvement was observed over the next 3 months. This result suggests that there may be a limit to the amount of possible relearning. It is not clear if listeners would be able to adapt to a frequency-place compression or expansion over time. In the present experiments the emphasis was solely on speech pattern recognition with no practice or time for accommodation.

D. Implications for speech pattern recognition

The present results illustrate the limitations of central pattern recognition mechanisms for speech, which may provide insights into the critical parameters of the pattern storage and retrieval process of central nervous system. Even though it is known that the patients adapt over time and improve their understanding of speech, it is still unclear how much plasticity exists in these central nervous system mechanisms and whether the ability to accommodate to some types of alterations (e.g., frequency-place shift) might be easier than other types of alterations (e.g., nonlinear frequency-place distortion).

The pattern of results observed in the present experiment, when combined with previous results on frequency-place shifting (Fu and Shannon, 1999; Dorman *et al.*, 1997), warping (Shannon *et al.*, 1998), and frequency lowering (Braida *et al.*, 1979; Reed *et al.*, 1983) suggests that the central pattern recognition of speech is not stored in terms of an abstract pattern, but in terms of an absolute pattern. Speech recognition in healthy acoustic hearing can tolerate a small degree of distortion in this frequency-place pattern, probably to accommodate the natural range of variation in real-world listening conditions, e.g., differences in the gender of the talkers, talking speeds and styles, and different amounts of masking and interference in the listening environment. The results of the present study, combined with the results of previous studies on frequency-place distortions, suggest that speech patterns can tolerate only a relatively small amount of distortion (2–3 mm) in tonotopic space. If the peripheral representation of the pattern of speech information is shifted, warped, expanded, or compressed beyond this tolerated cochlear distance of 2–3 mm, speech recognition will be significantly reduced. (It should be noted that none of these studies gave the subjects the opportunity to adapt to the distorted mappings.)

Figure 10 presents a schematic representation of a vowel spectrum and the various types of distortion that result in a reduction in multitalker vowel recognition to approximately

Effect of frequency/place distortion on vowel recognition

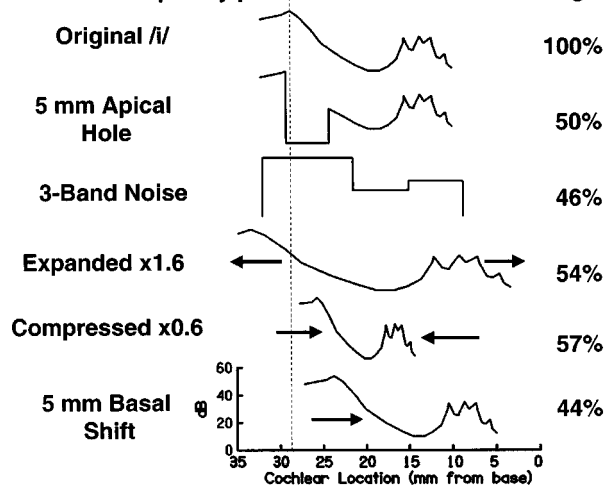


FIG. 10. A review of various frequency-place distortions on vowel recognition. The top curve is the original spectral representation of the vowel /i/ plotted in terms of cochlear distance. The second curve shows the same vowel with a 5-mm apical hole. Next is the vowel represented by a 3-band noise vocoder. The fourth curve is the frequency-place expansion with an expansion factor of 1.6 (10-mm extent expanded to 16 mm), followed by the frequency-place compression with a compression factor of 0.6 (26-mm extent compressed to 16 mm). Finally, the bottom curve shows the spectrum of an /i/ that has been shifted by 5 mm basally in the cochlea.

50% correct. These representations will be altered as they are processed by the cochlea and the central nervous system, but for simplicity they are shown here by their physical spectral representation in terms of distance along the cochlea. The top curve shows the original spectrum of the vowel /i/, presented in terms of mm along the cochlea. In this undistorted representation listeners will generally be able to identify 12 vowels at nearly 100% correct, even with multiple talkers and with the spectral resolution reduced to 16 channels. The second curve shows the same vowel in which spectral information has been removed to create a 5-mm hole in the apical spectral region, resulting in a drop to 50%-correct recognition (Shannon *et al.*, 2001). The third curve shows the vowel represented by a three-band noise vocoder, which allows 46% correct on multitalker vowel recognition (Fu *et al.*, 1998). The fourth curve shows the effect of a frequency-place expansion by a factor of 1.6, which results in 54%-correct vowel recognition with 16 bands (–3-mm condition from Fig. 7). The fifth curve shows the effect of frequency-place compression by a factor of 0.6, which results in 57%-correct recognition (+5-mm condition from Fig. 7). And, the bottom curve shows the spectrum of an /i/ that has been shifted 5 mm basally in the cochlea, resulting in 44%-correct vowel recognition (from Fu and Shannon, 1999). This comparison suggests that the central pattern recognition mechanisms are sensitive to the absolute tonotopic location of the cochlear pattern. If the frequency-place information is in the correct location, the central pattern recognition can tolerate the loss of a full octave of spectral information in the critical low-frequency region, or an extreme loss of spectral resolution—down to three bands. However, if the pattern is distorted by a frequency-place shift or compression or expansion, then speech recognition is impaired even with good spectral resolution. In terms of cochlear implants, even

if an implant patient is able to use many electrodes effectively, their performance might be limited by distortion in the frequency-place mapping.

It is possible that distortion in frequency-place mapping is responsible for at least part of the variability in performance across implant patients. If this is the case, then adjustments to the speech processor to produce a better match in frequency-place mapping may produce improvements in speech recognition.

ACKNOWLEDGMENTS

We would like to thank John Galvin for his help in the experimental setup and Qian-Jie Fu for letting us use his Tiger Speech Recognition System for sentence recognition tests. We also appreciate our subjects' efforts for the study. Funding for this research was provided in part by NIDCD Grant R01-DC-01526 and Contract N01-DC-92100.

Advanced Bionics Corporation (2001). CII Bionic Ear Programming System.

Braida, L. D., Durlach, N. I., Lippmann, R. P., Hicks, B. L., Rabinowitz, W. M., and Reed, C. M. (1979). "Hearing aids: A review of past research on linear amplification, amplitude compression, and frequency lowering," *ASHA Monogr.* **19**, 87–108.

Brill, S., Lawson, D., Wolford, R., Wilson, B., and Schatzer, R. (2001). *Speech Processors for Auditory Prostheses*, 11th Quarterly Progress Report, NIH Contract N01-DC-8-2105.

Cochlear Corporation (1995). *Technical Reference Manual*. Englewood, CO.

Cohen, L. T., Busby, P. A., Whitford, L. A., and Clark, G. M. (1996). "Cochlear implant place psychophysics. I. Pitch estimation with deeply inserted electrodes," *Audiol. Neuro-Otol.* **1**, 265–277.

Daniloff, R. G., Shiner, T. H., and Zemlin, W. R. (1968). "Intelligibility of vowels altered in duration and frequency," *J. Acoust. Soc. Am.* **44**, 700–707.

Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). "Simulating the effect of cochlear-implant electrode insertion depth on speech understanding," *J. Acoust. Soc. Am.* **102**, 2993–2996.

Echteler, S. M., Fay, R. R., and Popper, A. N. (1994). "Structure of the mammalian cochlea," in *Comparative Hearing: Mammals*, edited by R. R. Fay and A. N. Popper (Springer, New York), pp. 134–171.

Fu, Q.-J., Shannon, R. V., and Wang, X. (1998). "Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing," *J. Acoust. Soc. Am.* **104**, 3586–3596.

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., Shannon, R. V., and Galvin, J. J. (2002). "Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant," *J. Acoust. Soc. Am.* **112**, 1664–1674.

Greenwood, D. D. (1990). "A cochlear frequency-position function for several species—29 years later," *J. Acoust. Soc. Am.* **87**, 2592–2605.

Gstoettner, W., Franz, P., Hamzavi, J., Plenck, H., Baumgartner, W., and Czerny, C. (1999). "Intracochlear position of cochlear implant electrodes," *Acta Otolaryngol. (Stockh)* **119**, 229–33.

Hillenbrand, J., Getty, L., Clark, M., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am.* **97**, 3099–3111.

Ketten, D. R., Skinner, M. W., Wang, G., Vannier, M. W., Gates, G. A., and Neely, J. G. (1998). "In vivo measures of cochlear length and insertion depth of Nucleus cochlear implant electrode arrays," *Ann. Otol. Rhinol. Laryngol. Suppl.* **175**, 1–16.

Miller, G. A., and Nicely, P. E. (1955). "An analysis of perceptual confusions among some English consonants," *J. Acoust. Soc. Am.* **27**, 338–352.

Nagafuchi, M. (1976). "Intelligibility of distorted speech sounds shifted in frequency and time in normal children," *Audiology* **15**, 326–337.

National Institute of Standards and Technology (1990). *DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus CD-ROM*. Gaithersburg, MD.

Reed, C. M., Hicks, B. L., Braida, L. D., and Durlach, N. I. (1983). "Discrimination of speech processed by low-pass filtering and pitch-invariant frequency lowering," *J. Acoust. Soc. Am.* **74**, 409–419.

Robert, M. E. (1998). *CONDOR: Documentation for Identification Test Program* (House Ear Institute, Los Angeles, CA).

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.

Shannon, R. V., Galvin, J. G. III, and Baskent, D. (2001). "Holes in hearing," *J. Assoc. Res. Otolaryngol.* **3**(2), 185–199.

Tiffany, W. R., and Bennett, D. A. (1961). "Intelligibility of slow-played speech," *J. Speech Hear. Res.* **4**, 248–258.

Turner, C. W., Fabry, D. A., Barrett, S., and Horowitz, A. R. (1992). "Detection and recognition of stop consonants by normal-hearing and hearing-impaired listeners," *J. Speech Hear. Res.* **35**, 942–949.

Turner, C. W., Chi, S.-L., and Flock, S. (1999). "Limiting spectral resolution in speech for listeners with sensorineural hearing loss," *J. Speech Hear. Res.* **42**(4), 773–784.

Turner, C., and Gantz, B. (2001). "Combining acoustic and electric hearing for patients with high frequency hearing loss," *Abstracts of the 2001 Conference on Implantable Auditory Prostheses*, Asilomar, CA, p. 33.

Van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (1987). "Speech waveform envelope cues for consonant recognition," *J. Acoust. Soc. Am.* **82**, 1152–1161.

Numerical simulation of self-sustained oscillation of a voice-producing element based on Navier–Stokes equations and the finite element method

Martinus P. de Vries, Marc C. Hamburg, Harm K. Schutte, and Gijsbertus J. Verkerke
*Department of BioMedical Engineering, Division of Artificial Organs, University of Groningen,
Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands*

Arthur E. P. Veldman
*Institute of Mathematics and Computing Science, University of Groningen, Blauwborgje 3,
9747 AC Groningen, The Netherlands*

(Received 15 February 2002; revised 7 January 2003; accepted 24 January 2003)

Surgical removal of the larynx results in radically reduced production of voice and speech. To improve voice quality a voice-producing element (VPE) is developed, based on the lip principle, called after the lips of a musician while playing a brass instrument. To optimize the VPE, a numerical model is developed. In this model, the finite element method is used to describe the mechanical behavior of the VPE. The flow is described by two-dimensional incompressible Navier–Stokes equations. The interaction between VPE and airflow is modeled by placing the grid of the VPE model in the grid of the aerodynamical model, and requiring continuity of forces and velocities. By applying and increasing pressure to the numerical model, pulses comparable to glottal volume velocity waveforms are obtained. By variation of geometric parameters their influence can be determined. To validate this numerical model, an *in vitro* test with a prototype of the VPE is performed. Experimental and numerical results show an acceptable agreement. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1560163]

PACS numbers: 43.70.Bk, 43.70.Aj [AL]

I. INTRODUCTION

Surgical removal of the larynx is sometimes the treatment of last resort for patients with laryngeal cancer. The consequences of this removal are serious, implying among others the loss of the vocal folds. Since the first surgical removal of the larynx, over a hundred years ago, several methods to replace the voice source have been developed (Mahieu, 1988). Currently most patients can achieve a useful substitute voice for speaking. For this, air is brought into the esophagus by an air intake via the mouth or from the lungs via a shunt valve, surgically placed in the wall between trachea and esophagus (Fig. 1). While air escapes from the esophagus into the pharynx, the tissue at the esophagus entrance starts to vibrate, as in the case of belching (Blom *et al.*, 1988). The one-way shunt valve prevents leakage of food and fluids from the pharynx into the lungs (Mahieu, 1988).

The basic sound is (as in laryngeal speech production) acoustically converted to speech sound in the pharynx and oral and nasal cavity. However, the source sound has a low fundamental frequency F_0 (60–80 Hz) (Damsté, 1958; Snidecor and Curry, 1959; Cornut *et al.*, 1968), whereas values for F_0 in laryngeal speech are about 110 Hz for males and 210 Hz for females. Especially for a female laryngectomized person, this low F_0 is disturbing. Moreover, speech is often monotonous.

To overcome these drawbacks, a voice-producing element that produces a source sound with a higher F_0 and frequency variation during speech is under development. The underlying principle is comparable to the oscillating lips of a

musician playing a brass instrument (Sram, 1989; Adachi and Sato, 1996). In the voice-producing element, only one oscillating lip will be placed because a voice-producing element consisting of one lip is easier to manufacture and reproduce (Fig. 2). In the neutral position, the lip is slightly pressed against the opposite wall. When air pressure is applied at the inlet of the voice-producing element, the initially closed lip opens and the airflow makes the lip oscillate as a result of aerodynamic and mechanic forces.

The voice-producing element will be placed in the shunt valve between the trachea and the esophagus and consists of a single lip of silicone rubber. Silicone rubber is used for its very low elasticity that makes low F_0 possible. Since the voice-producing element is to be placed in the shunt valve several geometric requirements have to be met: it has to be placed in a lumen with an inner diameter of 5 mm and a length of 5 to 11 mm.

Other requirements concern the sound produced by the voice-producing element; F_0 , pressure, flow, and sound pressure level (SPL) have to fall within normal human ranges:

- (i) We aim at a voice-producing element with an F_0 of about 110 Hz for males and about 210 Hz for females.
- (ii) The sound pressure level at comfortable effort has to be between 70 and 80 dB, measured at a 30-cm distance from the mouth.
- (iii) The required pressure must fall within the range for normal speech, 0.4 to 3.0 kPa (Schutte, 1980).
- (iv) In normal speaking sentences of a certain length are necessary. Considering a total lung volume of 3.5–5

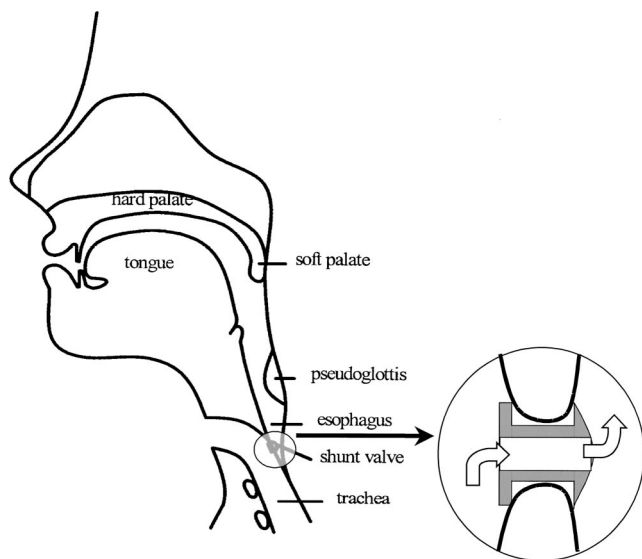


FIG. 1. Situation after laryngectomy with the location of the shunt valve. The arrows indicate the airflow direction during phonation.

L, the air flow range for driving the voice-producing element should fall in a range of 0.1–0.3 L/s as in normal phonation (Schutte, 1980).

- (v) The user of the voice-producing element must be able to vary the intensity and F_0 of the sound. Because no external control mechanisms are desired, both frequency and intensity must increase under influence of an increasing flow through the voice-producing element. While varying the applied aerodynamic forces F_0 must vary between about -10% and $+30\%$ of the values, mentioned above, and the intensity must vary between 70 and 80 dB at a distance of 30 cm (Schutte, 1992).

II. PROBLEM

To find the optimal geometry for the voice-producing element within the constraints of physiological possibilities, numerical modeling is used to avoid a trial-and-error approach. Because no numerical model on voice-producing prostheses exists, numerical models of the vocal folds are

studied. Since the human vocal folds and the new voice-producing element have much resemblance (both sound generators produce glottal waves as a result of interaction between the airflow and the oscillating lip), adapting a numerical model of the vocal folds could be a fast way to obtain a realistic model of the voice-producing element.

In numerical models of the vocal folds, both the vocal folds and the flowing air are simplified considerably, as in the models of several authors (Ishizaka and Flanagan, 1972; Titze, 1973, 1974; Pelorson *et al.*, 1994; Story and Titze, 1995). These models are often referred to as lumped-parameter models, because the properties of the vocal folds are lumped together in a small number of parameters. In most commonly used lumped-parameter models, the vocal folds are described by a number of masses, connected to each other by a number of springs and dampers. In these models, the behavior of the flowing air is approximated by the one-dimensional Bernoulli equation, implying a constant pressure distribution in a cross section. The air around the vocal folds does not behave like a one-dimensional flow, because aerodynamic pressure acts on both sides of the folds. As a consequence, a more detailed description of the aerodynamics is essential for an accurate examination of the behavior of the vocal folds. New numerical models are developed to describe the behavior of the airflow around the lip more accurately using two-dimensional Navier–Stokes equations (Alipour and Titze, 1996; Vries *et al.*, 2002). These models are a good basis for modeling the voice-producing element. Although it is possible to describe the lip by a number of lumped parameters using the numerical method described by de Vries *et al.* (1999), the lip can be represented more accurately with a finite element method (FEM) model.

The aim of this study is to determine the influence of the geometry of the lip and housing using a numerical model. For different configurations of the lip, the pressure where vibration starts, the range of self-sustained oscillation, the flow needed for voice production, and the possibilities to vary F_0 will be determined. From the results of the numerical simulations, an optimal configuration for the lip in the voice-producing element will be derived. To validate the numerical model, *in vitro* experiments are performed.

III. METHOD

A. Numerical model

The developed numerical model consists of two parts: a model of the aerodynamics based on Navier–Stokes equations and a model of the mechanics of the lips based on FEM. First the aerodynamics and mechanics are discussed separately, followed by the description of the interaction between the two models.

B. Aerodynamics

The model of the aerodynamics is a second-order version of the symmetry-preserving method of Verstappen and Veldman (1998) that is used extensively in simulating turbulent flow.

For the computation of the aerodynamic part, incompressible two-dimensional Navier–Stokes equations are

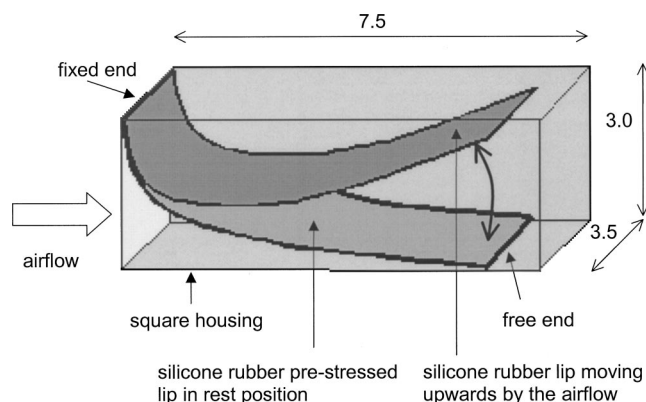


FIG. 2. Schematic representation of the voice-producing element. In rest the lip is slightly pressed against the opposite (lower) wall. The width of the lip is slightly less than the width of the square housing. Typical dimensions are given in mm.

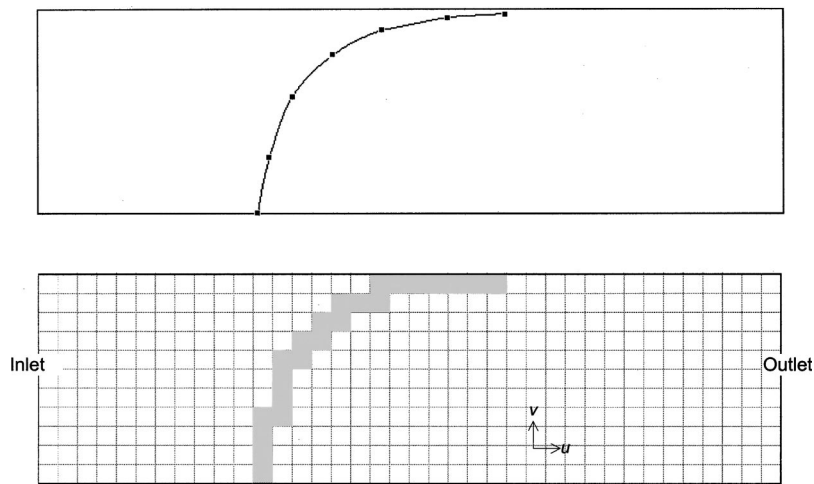


FIG. 3. Mechanical element distribution (above) and schematic representation of the grid of aerodynamic cells (white) in interaction with the mechanic cells (gray, below). Direction of velocities u and v is indicated.

used, as described by de Vries *et al.* (2002). Spatial discretization is performed in a Cartesian grid. This grid can be refined at places of particular interest. The three degrees of freedom of each cell in the grid (Fig. 3) are the horizontal velocity u , the vertical velocity v , and the pressure p .

As boundary conditions, velocities at the solid walls are $u=0$ and $v=0$. This describes an impermeable wall and a no-slip condition, which means that the fluid sticks to the wall due to its viscosity. At the inlet and outlet the pressure is prescribed, with the pressure difference driving the flow. Additionally, at the inlet and outlet the normal derivatives of the velocity components are set to zero, modeling fully developed flow. The latter condition requires the inlet and outlet to lie sufficiently far away from the interesting parts of the flow field.

The boundary conditions at the moving lip will be described below.

C. Mechanics

FEM is used to describe the mechanical behavior of the voice-producing element. Since the numerical model describing the aerodynamics is two-dimensional, the FEM model of the voice-producing element is two-dimensional as well. The thickness of the lip is small compared to its length, therefore it is allowed to approximate the lip using beam elements. In this way, the geometry in a cross-section is fixed, which would not be the fact when a three-dimensional element type, like a shell, would be chosen. The lip is divided in a number of beam elements, connected to each other in nodes. The beam element used has three degrees of freedom in every node: an axial and transverse translation, and a rotation.

The movement of the lip is restricted by the upper and lower wall of the voice-producing element. Therefore, during a cycle of the movement of the lip, contact between the lip and the upper and lower wall is assumed to collide with dissipation of all kinetic energy.

The material of the lip is chosen to be silicone rubber. This material is used for the shunt valves also. In the numerical simulations the silicone rubber lip is assumed to have a density of 1130 kg/m^3 , a Young's modulus of 8.6 MPa , and a proportional damping of 0.01 . The initially straight lip with a

length of 7 mm is bent 90° to fit into the housing that has an inner height of 3 mm . The shape of the deformed lip and the corresponding forces to keep the lip in the deformed shape are calculated with the nonlinear algorithm of the FEM program ANSYS 5.5 (SWANSON Analysis, USA). For all configurations of the lip, this calculation is performed. The forces that are needed to bend the lip are transformed into a pressure distribution along the lip and added to the pressure distribution resulting from the aerodynamic calculations. In this way, the lip is assumed to behave linearly in the final numerical model, moving around an equilibrium state that is calculated using a nonlinear method.

D. Interaction between aerodynamics and mechanics

To study the interaction between aerodynamics and mechanics, the two separate models are integrated by placing the mechanical model (calculated by FEM) in the grid of the aerodynamic model (Fig. 3). The cells of this grid can contain a volume of air or a part of the geometry. The interaction between aerodynamics and mechanics is obtained by exchanging information in the common grid in every time step.

The velocity of the air is equalized to the velocity of the surface of the lip; in this way a continuous velocity field is obtained.

The Navier–Stokes equations compute a pressure field. From this pressure field, forces acting on the lip are calculated by integrating the pressure distribution in the cells adjacent to every beam element and sum the resulting forces in every node between the beam elements. In this way, continuity of the force field is achieved. These forces are used to compute the movement of the lip using FEM. The new positions and velocities of the lips form the input to the next time step of the aerodynamics.

E. Transient numerical simulation

To achieve self-sustained oscillation, a pressure is prescribed at the inlet of the voice-producing element. In the simulations, pressure is increased from 0 up to 3 kPa . During the pressure rise, the lip starts to oscillate at a certain pressure, defined as vibration threshold pressure (Fig. 4), analogous to the phonation threshold pressure defined by Titze

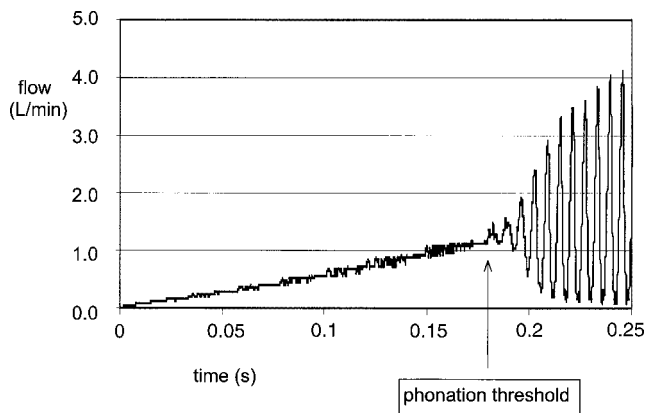


FIG. 4. Determination of vibration threshold pressure: while increasing the tracheal pressure, the moment of the start of self-sustained oscillation is established. The corresponding value of the tracheal pressure at that time is called the vibration threshold pressure for that configuration.

(1992). From this point on to higher pressures, the relation between tracheal pressure, mean flow, and F_0 are investigated, up to the pressure where no regular glottal waves are produced.

Several configurations of the lip are implemented, see Table I. A basic configuration is chosen and relative to this configuration, three parameters are varied: the thickness of the lip, the length of the lip, and the height of the housing that the lip is placed in. From this parameter study, an indication of the desired value of the different parameters will be obtained.

F. Numerical verification

The behavior of the mechanical part of the model is accurate when the behavior of the lip by the linear beam theory in static and dynamic behavior is valid. Because the quotient of thickness and length of the beam is very low, the lip can be considered to behave like the beam theory.

The numerical verification of the complete model, including the interaction between the aerodynamic model and the mechanic model, is done by varying the density of the grid of the aerodynamic model from 80×25 to 150×50 and the number of elements in the mechanic model from 3 to 9. When finer grids and more elements do not result in different results, the grid size and number of elements is considered to be optimal.

TABLE I. Dimensions of the simulated configurations of the voice-producing element. Dimensions that differ from the basic configuration are indicated in bold.

	Thickness lip (mm)	Length lip (mm)	Height element (mm)
Basic configuration	0.25	7	3
Long lip configuration	0.25	9	3
Short lip configuration	0.25	5	3
Higher opening configuration	0.25	7	4
Tapered configuration	0.25 (bottom)– 0.125 (free tip)	7	3
Thin lip configuration	0.125	7	3

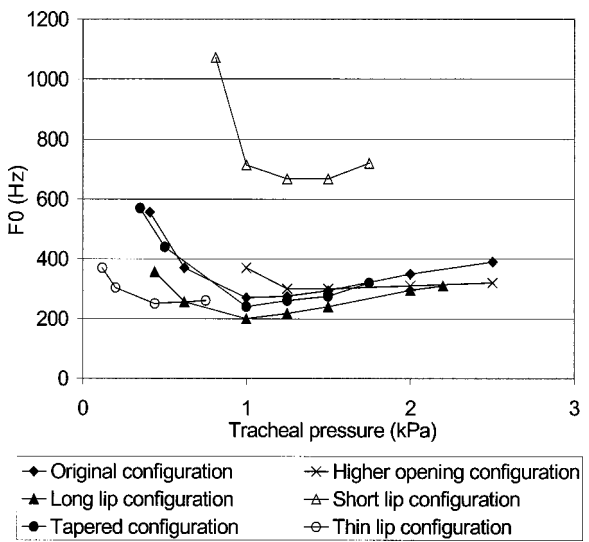


FIG. 5. F_0 as a function of tracheal pressure, as calculated by the simulation model. Curves start at pressures, where oscillation starts, and ends at pressures where oscillations become nonperiodical.

IV. EXPERIMENTS

To validate the numerical model of the lip principle and to obtain insight into the behavior of the lip principle, *in vitro* experiments are performed with a prototype of the voice-producing element in a test setup as described by Van der Plaats *et al.* (submitted). The flow is slowly increased by changing the pressure relief valve of a pressurized air cylinder from 0 to a flow that in this setup relates to a pressure of 3 kPa. The vibration threshold pressure F_0 as a function of tracheal pressure and mean flow rate is registered.

Both pressure and flow values of the *in vitro* experiments mentioned above are decreased by 40% to let the dip of the *in vitro* curve coincide with the dip of the simulated curve. This necessity can be explained by leakage along the lip and leakage caused by torsion of the lip. This rather high amount of leakage is caused by the provisional way of manufacturing and corresponds to other experiences with valve systems (Mihaylov *et al.*, 2000).

The corrected results of the *in vitro* experiments are compared to the values determined with the numerical model. Also a comparison with values obtained in laryngeal phonation is made to examine the correspondence between the voice produced by the voice-producing element and voice produced by the vocal folds.

V. RESULTS

A basic grid of 100 cells in the flow direction and 30 cells in the transverse direction appeared to be optimal because refining the grid did not change the results significantly, it only increased calculation time. For the same reason, a representation of the lip by six elements appeared to be the best.

In Fig. 5, the relation between F_0 and the simulated tracheal pressure is shown. At low tracheal pressure, the lip is in rest and no sound is made. At very high tracheal pressure, the lip is pressed against the opposite wall and no periodic sound is produced. Between these extremes, the lip vibrates

TABLE II. Characteristics of the relation between F_0 and tracheal pressure and between F_0 and mean flow though the voice-producing element, resulting from the numerical simulation model.

	Vibration threshold pressure (kPa)	Minimum F_0 (Hz)	Pressure at minimum F_0 (kPa)	Mean flow at minimum F_0 (L/s)	Increase of F_0 with mean flow (1/L)
Basic configuration	0.41	270	1.00	0.10	822
Long lip configuration	0.44	200	1.00	0.13	688
Short lip configuration	0.81	667	1.30	0.08	3533
Higher opening configuration	1.00	300	1.50	0.26	87
Tapered configuration	0.35	230	1.00	0.10	1067
Thin lip configuration	0.12	250	0.40	0.05	256

periodically. From Fig. 5 it can be seen that F_0 depends on the tracheal pressure, but in a different way for each configuration. All configurations show an initial decrease and slow increase of F_0 with increasing pressure. It is clear that the vibration threshold pressure (Fig. 4) differs for the different configurations, as is presented in Table II. It can be seen that the thickness and length of the lip and the height of the opening in the housing is important for the determination of the vibration threshold pressure: glottal volume velocity waveforms are produced at 0.12 kPa for the thin lip, whereas 0.41 kPa is needed for the basic configuration. The short lip starts oscillating at a pressure of 0.81 kPa, whereas the basic and long lip start oscillating at pressures of 0.41 and 0.44 kPa, respectively. The configuration with the highest opening has the highest vibration threshold pressure.

The range over which self-sustained oscillation occurs depends on the configuration: the thin lip has a phonation range of only 0.6 kPa, whereas the basic configuration has a range of more than 2 kPa.

The relation between F_0 and the mean flow is shown in Fig. 6. The mean flow through the simulated voice-producing element is almost linearly related to the tracheal pressure, therefore a minimum in the F_0 as a function of the mean flow

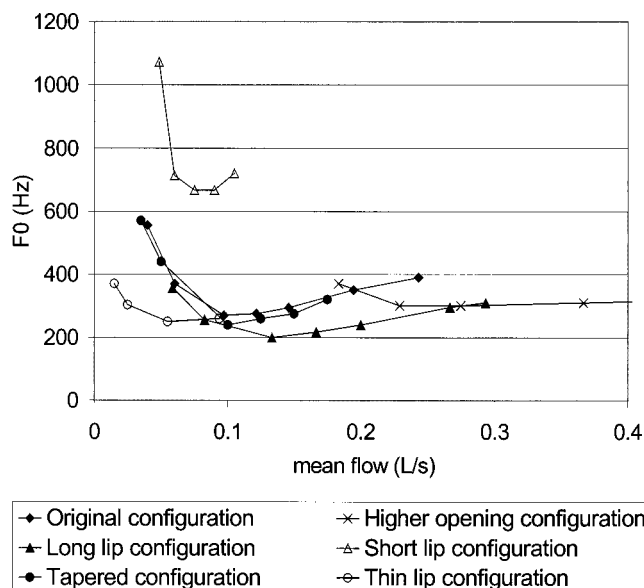


FIG. 6. F_0 as a function of mean flow, as calculated by the simulation model. Curves start at flows, where oscillation starts and ends at flows where oscillations become nonperiodical.

is also present. It must be noted that the higher opening configuration results in a considerable shift towards higher flow values.

Figures 4, 6, and 7 show that the relation between F_0 and tracheal pressure and between F_0 and mean flow all show a U-shape of the graph. A rise of tracheal pressure and flow results in a decrease of F_0 for all configurations until a

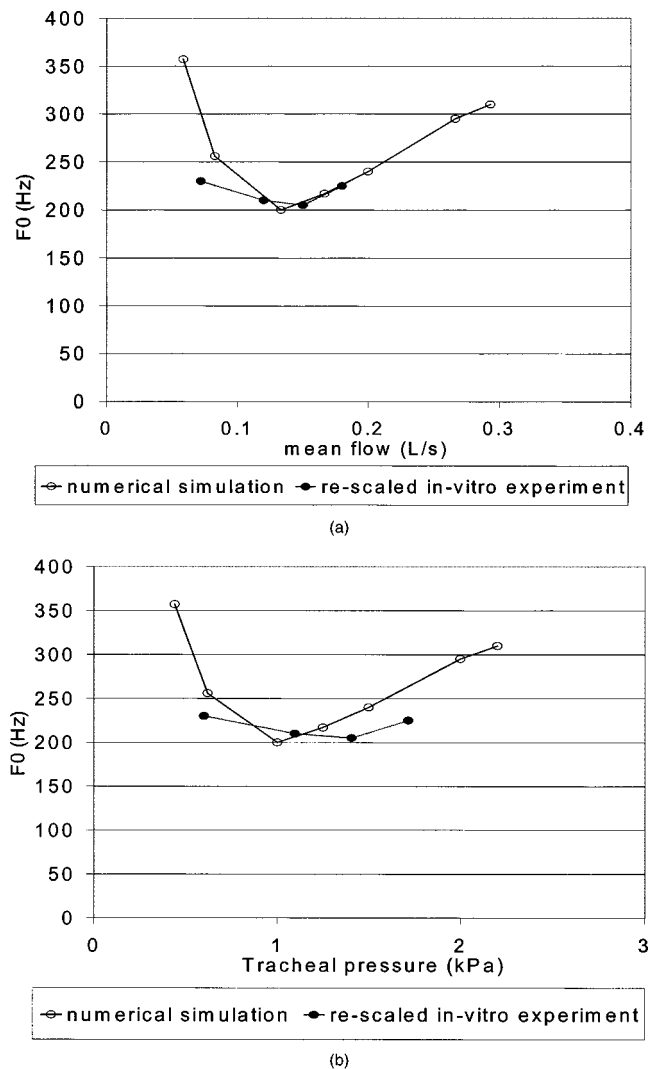


FIG. 7. Relation between F_0 and mean flow (a) and between F_0 and mean pressure (b) of the long lip configuration, resulting from the numerical simulation model and from the *in vitro* experiments. *In vitro* results are re-scaled by 0.6 to account for leakage.

minimum F_0 is reached. The minimum F_0 and the corresponding tracheal pressure and mean flow are presented in Table II. After that point F_0 increases slightly with increasing pressure and flow. In Table II, the rate of increase is represented by the slope of the line connecting the two most extreme F_0 values. The tracheal pressure at which the minimum F_0 occurs appears to be related with the vibration threshold pressure: a low vibration threshold pressure corresponds to a low pressure at which the minimum F_0 occurs.

The relations between F_0 , tracheal pressure, and mean flow of the prototype used in the *in vitro* experiments are shown in Fig. 7, together with the simulation results of the corresponding long lip configuration. In this way, the difference and correspondence between the results of the *in vitro* experiments and of the numerical simulations can be used to judge the validity of the numerical model.

From Fig. 7(a) it can be seen that the two curves of the relation between F_0 and mean flow, resulting from the *in vitro* experiments and from the simulations both have a falling and raising part, so show the same trend. The rising parts of both graphs even coincide almost completely. The two curves of the relation between F_0 and tracheal pressure, resulting from the *in vitro* experiments and from the simulations [Fig. 7(b)], correspond less, but are still acceptable. Both in Figs. 7(a) and (b), it can be seen that the range in mean flow, respectively tracheal pressure, as is simulated is larger than the range found during the *in vitro* experiments.

VI. DISCUSSION

The values of tracheal pressure needed for the production of glottal waves with the voice-producing element should fall in a range comparable to laryngeal voice production. Therefore, these values are compared to those measured in laryngeal phonation (Schutte, 1980). A mean value for the tracheal pressure of 0.44 kPa and a mean airflow rate of about 0.2 L/s were measured. Considering the results of the simulations, it appeared that four of the six configurations oscillate at that pressure. However, the ideal relation between F_0 and pressure is not yet obtained. At low pressures, an increasing pressure causes a decreasing F_0 . The normal intonation pattern is characterized by an increase of F_0 with increasing pressure. Patients who are able to put more effort in speech than the comfortable value of about 1 kPa (Schutte, 1980), by producing pressures up to 3 kPa, can reach that intonation pattern. The other two configurations (short lip and higher opening) start to oscillate only at higher pressures.

In the ideal flow range of 0.1–0.3 L/s, three of the six configurations oscillate according to the requirements: increasing mean flow causes an increasing F_0 . The higher opening configuration requires too much mean flow, and speech duration will be shortened too much; the thin lip and short lip configuration require a nonphysiological low mean flow.

The long lip configuration fulfills frequency requirements for female patients best. The lowest frequency obtained is 200 Hz. The required female frequency increase of

80 Hz per 0.2 L/s = 400 Hz/L/s is realized by the basic, long lip, short lip, and tapered configuration.

The assumption of the flow to be incompressible is valid in our model because in our application, the airflow velocity did not exceed 50 m/s, which is below Mach 0.2, at which compressible effects can be considered negligible. The 2D approach sometimes leads to unrealistic results. The pressure and flow range at which oscillation occurs is larger in the simulations than in the *in vitro* experiments: *In vitro* the vibration threshold pressure appeared to be slightly higher; also, oscillation above the maximum phonation tracheal pressure and mean flow appears to be nonperiodic and not useful for voice production, whereas during simulation much higher pressures and flows still produce periodic oscillations. Most probably, 3D effects like leakage along the lip and torsion of the lip (absent in the 2D numerical model) cause the limited range found during the *in vitro* experiments.

In both Figs. 7(a) and (b), the curves resulting from the simulation and from the *in vitro* experiments show a large similarity, thus providing the validity of the numerical model.

As a continuation, the numerical model has been used to determine the optimal male and female configuration. The resulting prototypes have been tested on aero-acoustics (Van der Torn *et al.*, 2001).

VII. CONCLUSIONS

The numerical model appears to be a valid tool to study the mechanical and aerodynamical behavior of a voice-producing element.

Lip length, lip thickness, and height of the housing influence the performance of the voice-producing element. From the simulated versions, the best female voice-producing element is the “long lip configuration.”

The numerical model can be used for further optimization studies of the voice-producing element, for instance by combining several successful configurations like the long and tapered one. A lower opening configuration could lead to a shift towards a lower flow range.

ACKNOWLEDGMENTS

This research is supported by the Technology Foundation STW, applied science division of NWO and the technology program of the Ministry of Economic Affairs.

- Adachi, S., and Sato, M. (1996). “Trumpet sound simulation using two-dimensional lip vibration model,” *J. Acoust. Soc. Am.* **2**, 1200–1209.
- Alipour, F., and Titze, I. R. (1996). “Combined simulation of two-dimensional airflow and vocal fold vibration,” in *Vocal Fold Physiology: Controlling Complexity and Chaos*, edited by P. J. Davis and N. H. Fletcher (Singular, San Diego), pp. 17–29.
- Blom, E. D., Singer, M. I., and Hamaker, R. C. (1988). “Total laryngectomy with voice preservation,” in *The Larynx: A Multidisciplinary Approach*, edited by M. P. Fried (Little, Brown, Boston), pp. 517–531.
- Cornut, G., Laganier, C., and Agnes, M. J. (1968). “Contribution à l’étude acoustique de la voix oesophagienne,” *J. Fr. d’Oto-Rhino-Laryngologie* **XVII**, 269–276.
- Damsté, P. H. (1958). “Oesophageal speech after laryngectomy (slokdarmspraak),” Ph.D. thesis, University of Groningen.
- Ishizaka, K., and Flanagan, J. L. (1972). “Synthesis of voiced sounds from a two-mass model of the vocal cords,” *Bell Syst. Tech. J.* **51**, 1233–1268.

- Mahieu, H. F. (1988). "Voice and Speech Rehabilitation following Laryngectomy," Ph.D. thesis, University of Groningen.
- Mihaylov, D., Verkerke, G. J., Plaats, A. v. d., Loon, J.-P. v., Hummel, J. M., and Rakhorst, G. (2000). "Design and test of the PUCA-2, a single valved left ventricular assist device," *Int. J. Artif. Organs* **23**, 462–472.
- Pelorson, X., Hirschberg, A., van Hassel, R. R., Wijnands, A. P. J., and Auregan, Y. (1994). "Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model," *J. Acoust. Soc. Am.* **96**, 3416–3431.
- Plaats, A. v. d., Vries, M. P. d., Schutte, H. K., and Verkerke, G. J. (submitted). "In-vitro measurements of a voice producing element under physiologic acoustic conditions," *J. Acoust. Soc. Am.*
- Schutte, H. K. (1980). "The Efficiency of Voice Production," Ph.D. thesis, University of Groningen.
- Schutte, H. K. (1992). "Integrated Aerodynamic Measurements," *J. Voice* **6**, 127–134.
- Snidecor, J., and Curry, E. T. (1959). "Temporal and pitch aspects of superior oesophageal speech," *Ann. Otol. Rhinol. Laryngol. (St. Louis)* **68**, 623–635.
- Sram, F. (1989). "Laute und Silben der Hilfsartikulation beim Blasen von Blechblasinstrumenten," in *Das Instrumentalspiel* (Verlag Doblinger, Prag), pp. 137–146.
- Story, B. H., and Titze, I. R. (1995). "Voice simulation with a body cover model of the vocal folds," *J. Acoust. Soc. Am.* **97**, 1249–1260.
- Titze, I. R. (1973). "The human vocal cords: a mathematical model. Part I," *Phonetica* **28**, 129–170.
- Titze, I. R. (1974). "The human vocal cords: a mathematical model: Part 2," *Phonetica* **29**, 1–21.
- Titze, I. R. (1992). "Phonation threshold pressure: A missing link in glottal aerodynamics," *J. Acoust. Soc. Am.* **91**, 2926–2935.
- Van der Torn, M. (2001). "Aero-acoustics of silicone rubber lip reeds for alternative voice production in laryngectomees," *J. Acoust. Soc. Am.* **110**, 2548–2559.
- Verstappen, R. W. C. P., and Veldman, A. E. P. (1998). "Spectro-consistent discretization: a challenge to RANS and LES," *J. Eng. Math.* **34**, 163–179.
- Vries, M. P. d., Schutte, H. K., and Verkerke, G. J. (1999). "Determination of parameters for lumped parameters models of the vocal folds using a finite element method approach," *J. Acoust. Soc. Am.* **106**, 3620–3628.
- Vries, M. P. d., Schutte, H. K., Veldman, A. E. P., and Verkerke, G. J. (2002). "Glottal flow through a two-mass model: Comparison of Navier–Stokes solutions with simplified models," *J. Acoust. Soc. Am.* **111**, 1847–1853.

Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing

Judy R. Dubno,^{a)} Amy R. Horwitz, and Jayne B. Ahlstrom

Department of Otolaryngology-Head and Neck Surgery, Medical University of South Carolina,
135 Rutledge Avenue, P.O. Box 250550, Charleston, South Carolina 29425

(Received 25 June 2002; accepted for publication 3 January 2003)

In a previous study [Dubno *et al.*, J. Acoust. Soc. Am. **111**, 2897–2907 (2002)], older subjects benefitted less than younger subjects from momentary improvements in signal-to-noise ratio when listening to speech in interrupted maskers. It has been hypothesized that the benefit derived from interrupted maskers may be related to recovery from forward masking, i.e., the recovery of a response to a suprathreshold signal from prior stimulation by a masker. The effect of interrupted maskers on speech recognition may be well suited to test hypotheses regarding recovery from prior stimulation, given that both involve the perception of signals following a masker. Here, younger and older adults with normal but not identical audiograms listened to nonsense syllables at moderate and high levels in a speech-shaped noise that was modulated by a 2-, 10-, 25-, or 50-Hz square wave. An additional low-level noise was always present that was shaped to produce equivalent masked thresholds for all subjects. To assess recovery from forward masking, forward-masked thresholds were measured at 0.5 and 4.0 kHz as a function of the delay between the speech-shaped masker and the signal. Speech recognition in interrupted noise was poorer for older than younger subjects. Small but consistent age-related differences were observed in the decrease in score with interrupted noise relative to the score without interrupted noise. Forward-masked thresholds of older subjects were higher than those of younger subjects, but there were no age-related differences in the amount of forward masking or in simultaneous masking. Negative correlations were observed between speech-recognition scores in interrupted noise and forward-masked thresholds. That is, the benefit derived from momentary improvements in speech audibility in an interrupted noise decreased as forward-masked thresholds increased. Stronger correlations with forward masking were observed for the higher frequency signal, for higher noise interruption rates, and when the signal-to-noise ratio was poor. Comparisons of speech-recognition scores at moderate and high levels for younger and older subjects were not consistent with the hypothesis of an age-related difference in the contribution of low-spontaneous-rate fibers to speech recognition in interrupted noise. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1555611]

PACS numbers: 43.71.Lz, 43.66.Dc, 43.66.Sr [CWT]

I. INTRODUCTION

For younger subjects with normal hearing, the ability to understand speech in an interrupted masker varies with the masker interruption rate, the type of speech material, and the relative level of the speech and masker (e.g., Miller and Licklider, 1950; Howard-Jones and Rosen, 1993; Gustafsson and Arlinger, 1994). When listening to speech in interrupted or modulated noise, audible speech information during the momentary improvements in signal-to-noise ratio that occur provides substantial improvements in speech recognition relative to performance in steady-state noise. However, speech recognition of subjects with hearing loss, even those with relatively mild loss, improves less or not at all in modulated or interrupted noise relative to performance in steady-state noise (Festen and Plomp, 1990; Takahashi and Bacon, 1992; Eisenberg *et al.*, 1995; Stuart and Phillips, 1996; Bacon *et al.*, 1998). That is, hearing-impaired subjects benefit less from improvements in signal-to-noise ratio that occur in

modulated or interrupted noises. This result is often explained in two ways. First, elevated thresholds limit the improvement in signal-to-noise ratio that may be realized when the masker level momentarily decreases. Second, the benefit derived from modulated or interrupted maskers may be related to the recovery from forward masking, i.e., the recovery of a response to a signal from prior stimulation by a masker (Festen and Plomp, 1990; Bronkhorst and Plomp, 1992; Festen, 1993; Gustafsson and Arlinger, 1994). Although it is possible that the deficit for subjects with hearing loss is due simply to limited audibility, in some cases this explanation does not entirely account for differences in speech recognition in modulated maskers between subjects with normal and impaired hearing (e.g., Eisenberg *et al.*, 1995; Bacon *et al.*, 1998).

Few studies have focused on age-related differences in benefit from interrupted noise, independent of hearing loss, and the results are mixed. Older subjects with normal hearing or only mild, high-frequency hearing loss benefitted less from modulated maskers than younger subjects with normal hearing (Takahashi and Bacon, 1992; Stuart and Phillips,

^{a)}Electronic mail: dubnojr@muscd.edu

1996; Dubno and Ahlstrom, 1997). However, Souza and Turner (1994) found no differences in speech recognition in multitalker babble and steady-state noise modulated by multitalker babble between younger and older subjects with equivalent mild-to-moderate sloping hearing loss.

Differences between younger and older subjects may be explained by the same two factors mentioned previously, namely reduced audibility (due to slightly elevated thresholds) and age-related changes in forward masking. With regard to reduced audibility, results of Dubno and Ahlstrom (1997) suggested that older subjects' slightly higher pure-tone thresholds in quiet can result in substantially elevated word-recognition thresholds under certain conditions. With regard to forward masking, although there is considerable evidence of age-related changes in various measures of temporal resolution (e.g., Abel *et al.*, 1990; Peters and Hall, 1994; Fitzgibbons and Gordon-Salant, 1994; Phillips *et al.*, 1994; Strouse *et al.*, 1998), little is known about age-related changes in recovery from forward masking. Walton *et al.* (1999) reported "prolonged" recovery from forward masking for older subjects with normal hearing measured using the auditory brainstem response. Studies of gap detection, a closely-related measure, suggest that older subjects have larger gap detection thresholds (Snell, 1997). However, other studies find age-related differences in gap detection only when the gap is placed near the onset or offset of a noise (He *et al.*, 1999). Thus, it remains unclear whether age-related changes in recovery from forward masking contribute to older subjects' reduced benefit from interrupted maskers.

To explore these two factors, an experiment was conducted whereby pure-tone thresholds and speech recognition of younger and older subjects with normal hearing were measured in a speech-shaped steady-state noise at two levels and in the same noise modulated by a 10-Hz square wave (Dubno *et al.*, 2002). Forward-masked thresholds were also measured at several signal frequencies in two masker levels. An additional low-level broadband noise was always present which was shaped to produce equivalent masked thresholds for all subjects. Forward-masked thresholds were higher and benefit of interrupted maskers was smaller for older than younger subjects, particularly at a higher masker level. Moreover, significant correlations were observed between speech recognition in interrupted noise and forward-masked thresholds, especially at higher frequencies.

Based on these initial results, the current study was undertaken to explore further the relationship between age-related changes in recovery from forward masking and speech recognition in interrupted maskers. The effect of interrupted maskers on speech recognition may be well suited to test hypotheses regarding recovery from prior stimulation, given that both involve the perception of signals following a masker. Further, recovery from forward masking may be related to subgroups of auditory neurons with different thresholds. For tonal signals, a two-segment recovery function measured in animals has been attributed to low-spontaneous-rate (SR) fibers recovering more slowly than high-SR fibers from prior stimulation (Salvi *et al.*, 1986; Relkin and Doucet, 1991; Relkin *et al.*, 1995). At shorter delays between the masker and signal, a steeper segment of the recovery func-

tion reflects the faster recovery of the high-SR fibers; at longer delays, a shallower segment reflects the slower recovery of the low-SR fibers. Results consistent with this notion were observed in human subjects by measuring the amplitude of the forward-masked compound action potential (CAP) in response to a 4.0-kHz probe (Murnane *et al.*, 1998). In a psychophysical study in human subjects by Zeng *et al.* (1991), intensity discrimination following a high-level masker was degraded at moderate levels compared to lower and higher levels (the "mid-level hump"). The results were attributed to a gap in intensity coding at moderate levels where high-SR fibers were saturated and low-SR fibers had not recovered from the forward masker.

Of potential importance to age-related effects was a finding by Schmiedt *et al.* (1996) demonstrating that at high frequencies (>6.0 kHz), forward-masked CAP recovery functions were steeper in aged gerbils than in young gerbils. These authors attributed this result to a selective loss (or inactivity) of low-SR fibers with high characteristic frequencies, possibly related to age-related reductions of the endocochlear potential (EP; Schulte and Schmiedt, 1992) which may differentially affect low-SR fibers (Sewell, 1984). This finding may have significance for several aspects of auditory function of older subjects. In addition to recovering more slowly relative to high-SR fibers, low-SR fibers have higher thresholds, larger dynamic ranges, smaller effective response areas (narrower bandwidths), and are better able to preserve timing information and amplitude modulation (see review in Schmiedt *et al.*, 1996). These characteristics may provide low-SR fibers with an increased resistance to the effects of masking.

Here, speech recognition in the presence of interrupted maskers was regarded as analogous to recovery of a suprathreshold signal from prior stimulation by a masker. Younger and older adults with normal but not identical audiograms listened to nonsense syllables at moderate and high levels in a noise that was modulated by a 2-, 10-, 25-, or 50-Hz square wave. This resulted in interrupted noises with a range of interruption rates. An additional low-level noise was always present that was shaped to produce masked thresholds of 20–30 dB HL across frequency to equate audibility across subjects despite differences in quiet thresholds. To assess recovery from forward masking, forward-masked thresholds were measured at 0.5 and 4.0 kHz as a function of the delay between the speech-shaped masker and the signal.

II. METHODS

A. Subjects

There were two subject groups: (1) ten younger subjects (mean age: 25.1 years; range: 21–30) and (2) ten older subjects (mean age: 65.7 years; range: 61–72). All subjects had "normal hearing" defined as thresholds ≤ 20 dB HL (ANSI, 1996) at octave frequencies from 0.25 to 4.0 kHz. Immitance measures were within normal limits. Subjects did not have experience with the psychophysical tasks used in this study and thus received approximately 2 h of practice with the

various tasks. Data collection was completed in 5–7 2-h sessions. Subjects were paid an hourly rate for their participation.

B. Apparatus and stimuli

1. Tonal signals

Tonal signals were digitally generated (TDT DA3-4) pure tones sampled at 50 kHz and low-pass filtered at 12 kHz (TDT FT6). Signals were either 350 ms, including 10-ms raised-cosine rise/fall ramps, or 20 ms, including 5-ms raised-cosine rise/fall ramps. Thresholds for the 350-ms signals were measured at 16 one-third-octave frequencies ranging from 0.2 to 6.3 kHz in quiet and in noise. Thresholds for the 20-ms signals were measured at 0.5 and 4.0 kHz in quiet, in simultaneous masking, and in forward masking. Signal level was varied adaptively. For forward masking, the signal was presented at 10, 20, 50, or 100 ms following masker offset (at the 0-voltage points). For simultaneous masking, the onset of the 20-ms signal was 90 ms after the onset of the 200-ms masker. Quiet thresholds for frequencies from 8.0 to 18.0 kHz were measured with a Demlar audiometer and Koss headphones.

2. Speech

Speech signals were 66 consonant-vowel and 63 vowel-consonant syllables formed by combining 22 or 21 English consonants with /a,i,u/ spoken by one male and one female talker (a total of 258 syllables). The stimulus set was divided into 12 subsets to provide for a reasonable listening interval. A detailed description of the speech stimuli can be found in Dubno and Schaefer (1992, 1995). Each syllable was presented two times in random order; thus, each speech-recognition score was the average of 516 responses. Digital speech waveforms (without a carrier phrase) were output at a sampling rate of 32.89 kHz, mixed with various noises, and delivered through one of a pair of TDH-49 earphones mounted in supra-aural cushions.

3. Noises

For both forward and simultaneous masking, the masker was a 200-ms speech-shaped noise (with 5-ms raised-cosine rise/fall). Steady-state broadband noise was digitally generated at a sampling rate of 28 kHz and then spectrally shaped in one-third-octave intervals to match the long-term spectrum of the speech. This speech-shaped noise was then modulated by a 2-, 10-, 25-, or 50-Hz square wave to create interrupted noises with varying interruption rates (Cool Edit Pro™ Version 1.2, Syntrillium Software Corp.). The overall level of the interrupted noise was 80 dB SPL. The level was set such that during the period when the interrupted noise was “on,” it was identical to the original steady-state noise. Accordingly, the overall level of each interrupted noise was lower than the steady-state noise; no adjustments in level were made to compensate for this difference. The timing of each syllable presentation was random and independent of the noise so that the location of the consonant portion of the syllable with respect to the “on” or “off” periods of the noise was also random. For speech-recognition measures,

noise was output through a 16-bit digital-to-analog converter (TDT DA3-4), low-pass filtered at 12 kHz (TDT FT6), and recorded onto digital audio tape (DAT) for playback.

We assumed that quiet thresholds for older subjects would be slightly higher than those of younger subjects; in addition, it was likely that there would be small differences in thresholds among subjects within each group. As a result, differences in signal audibility could arise during the “off” periods of the interrupted noise. To minimize these differences, a “threshold matching noise” (TMN) was present during speech-recognition measurements. This low-level broadband noise was generated and its spectrum adjusted at one-third-octave intervals (Cool Edit Pro™) to produce equivalent masked thresholds for all subjects. Band levels of the TMN were set to shift thresholds to 20–25 dB HL from 0.2 to 3.15 kHz, 30 dB HL from 4.0 to 5.0 kHz, and 40 dB HL at 6.3 kHz (thresholds that were slightly higher than the highest thresholds of any subject). TMN was output through a 16-bit digital-to-analog converter (TDT PD1), low-pass filtered at 12 kHz (TDT FT6), and recorded onto DAT for playback. Spectral characteristics of all noises were verified on an acoustic coupler and a signal analyzer (Stanford Research SR780). The amplitudes of the signals and maskers were controlled individually using programmable and manual attenuators (TDT PA4). The signals were added to the masker (TDT SM3), low-pass filtered (TDT PF1) at 7.0 kHz, and delivered to the earphone.

C. Procedures

For each subject, thresholds were measured in the following order: (1) thresholds in quiet for 350-ms pure tones; (2) masked thresholds for 350-ms tones in the low-level TMN; (3) masked thresholds for 350-ms pure tones in a speech-shaped steady-state noise at 80 dB SPL; (4) thresholds in quiet for 20-ms signals at 0.5 and 4.0 kHz; and (5) thresholds for the 20-ms signals in forward and simultaneous masking. Masked thresholds in TMN and in speech-shaped steady-state noise were used for computing articulation index (AI) values.

Thresholds were obtained using a single-interval (yes–no) maximum-likelihood psychophysical procedure, similar to that described by Green (1993) and discussed in detail in Leek *et al.* (2000). Each threshold was determined from 24 trials, four of which were catch trials. Listen and vote periods were displayed on the screen of a computer monitor. Subjects responded by clicking one of two mouse buttons corresponding to the responses “yes, I heard the tone” and “no, I did not hear the tone.” Thresholds at frequencies from 8.0 to 18.0 kHz were measured using a modified method of limits psychophysical procedure.

Following threshold measures, recognition of consonant-vowel and vowel-consonant syllables was measured in TMN alone and in each of the four interrupted noises, with speech presented at either 65 or 80 dB SPL. For the 2-Hz rate, scores were obtained for speech at 65 dB SPL only. Subjects were instructed to respond by clicking the mouse on the appropriate alphabetic symbol displayed on the computer monitor. All consonants contained in the subset were available alternatives; correct answer feedback was not

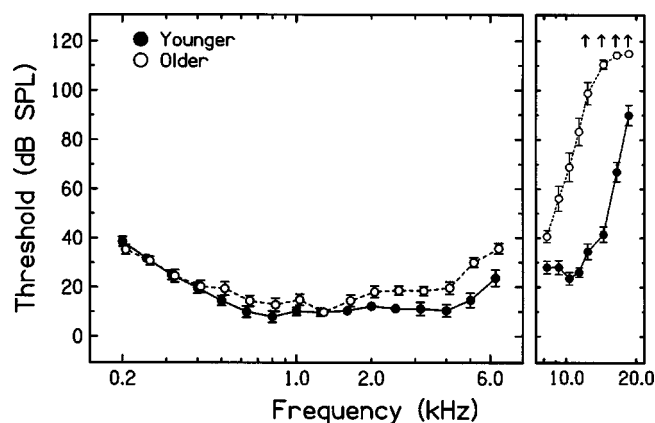


FIG. 1. Mean thresholds for 350-ms tones measured in quiet for younger subjects (filled) and older subjects (open). Error bars indicate ± 1 standard error of the mean (SE). Thresholds for frequencies in the audiometric range are in the left panel and thresholds for extended high frequencies are in the right panel. If subjects did not respond at the maximum intensity presented (110 dB SPL), thresholds were assigned a value of 115 dB SPL and the mean values designated with an arrow (\uparrow).

provided. Speech and noise levels, and noise interruption rates, were selected based on results of Zeng *et al.* (1991), which demonstrated that intense forward maskers with specific delays increased the size of the intensity difference limen at moderate signal levels. In addition to scores in TMN and in interrupted noise, the “penalty” of interrupted noise was computed, that is, the difference between scores with and without the interrupted noise. Differences in scores due to age, speech level, and interruption rate were assessed by repeated-measures ANOVA (effects were considered significant when $p < 0.05$).

III. RESULTS AND DISCUSSION

A. Quiet thresholds

Mean thresholds (in dB SPL) for younger and older subjects are shown in Fig. 1. Thresholds for frequencies

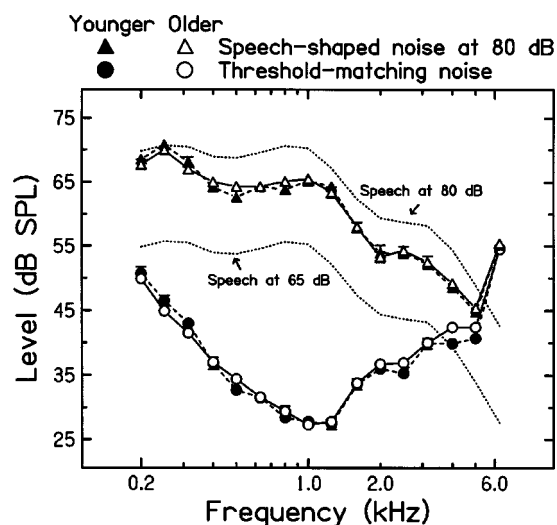


FIG. 2. Mean thresholds (± 1 SE) for 350-ms tones for younger and older subjects in the speech-shaped noise at 80 dB SPL (triangles) and in the threshold-matching noise (circles). As a reference, long-term average spectra for nonsense syllables presented at 65 and 80 dB SPL are shown by the dotted lines.

in the audiometric range are in the left panel; thresholds for extended high frequencies are in the right panel. From 0.5 to 6.3 kHz, thresholds for the older subjects were higher generally than thresholds for the younger subjects, although all subjects met the criterion for “normal hearing.” Over this frequency range, thresholds averaged 6.7 dB higher for older than younger subjects; from 2.0 to 6.3 kHz, thresholds averaged 9.5 dB higher for older than younger subjects. Much larger threshold elevations for older subjects were observed at frequencies > 8.0 kHz, suggesting substantial age-related changes to the base of the cochlea.

B. Masked thresholds

1. Thresholds in steady-state noise (350 ms)

Figure 2 includes thresholds for 350-ms tones for younger and older subjects in the speech-shaped steady-state noise at 80 dB SPL (triangles) and in the threshold-matching noise (circles). Also shown for comparison are the long-term average spectra of the nonsense syllables at two levels (65 and 80 dB). To obtain a single value to compare masked thresholds between groups, weighted average thresholds in the steady-state noise and TMN were computed, using weights from the frequency importance function for the speech stimuli used in this experiment (Dirks *et al.*,

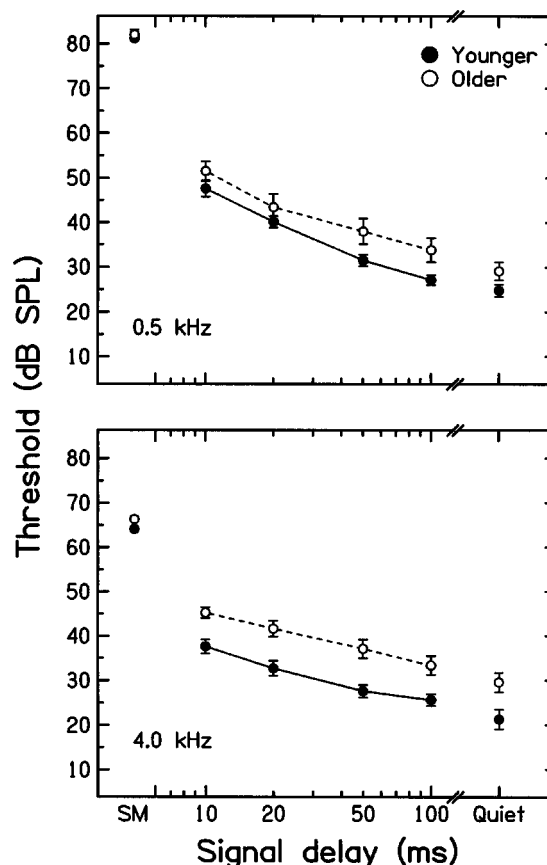


FIG. 3. Mean thresholds (± 1 SE) for 20-ms tones for younger and older subjects at 0.5 kHz (top) and 4.0 kHz (bottom), measured in forward masking as a function of masker-signal delay, in simultaneous masking (SM, left-most data points), and in quiet (right-most data points).

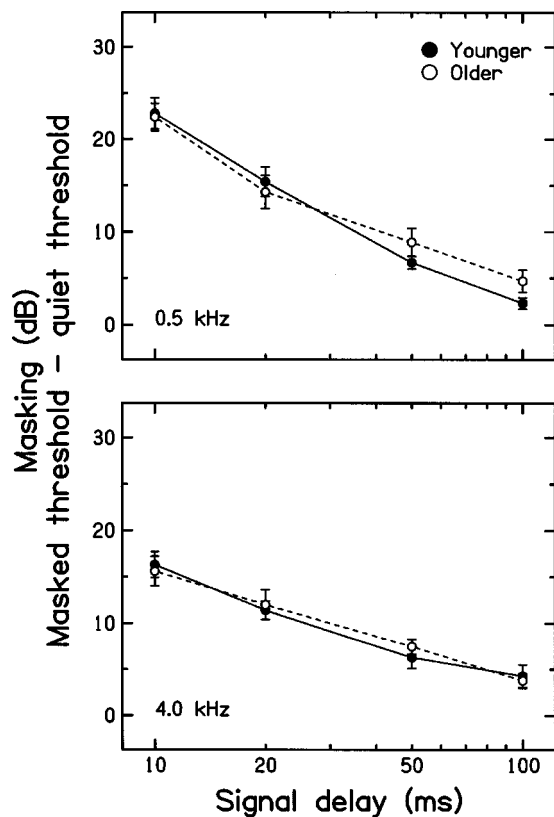


FIG. 4. Mean forward masking (± 1 SE) for younger and older subjects at 0.5 kHz (top) and 4.0 kHz (bottom), as a function of masker-signal delay. Masking was computed as the difference between thresholds in forward masking and in quiet.

1990a,b). For both noises, differences in weighted-average thresholds between younger and older subjects were less than 1 dB.

2. Thresholds in forward and simultaneous masking (20 ms)

Figure 3 plots mean thresholds for 20-ms tones at 0.5 kHz (top) and 4.0 kHz (bottom) for younger and older subjects. The four connected data points are thresholds measured in forward masking as a function of masker-signal delay. The masker was the speech-shaped noise at 80 dB SPL. The left-most data points are thresholds in simultaneous masking. The right-most data points are thresholds in quiet. In forward masking and in quiet, thresholds were higher at both signal frequencies for older than younger subjects. Averaged across frequency and delay, forward-masked thresholds were significantly higher by 6.8 dB for older than younger subjects. Similarly, thresholds in quiet averaged 6.3 dB higher for older than younger subjects. There were no significant age-related differences in simultaneous-masked thresholds.

To determine if age-related differences in forward masking were attributable to threshold differences in quiet, amount of forward masking for each subject was computed, i.e., the difference between each threshold in forward masking and in quiet. Figure 4 shows amount of forward masking at 0.5 kHz (top) and 4.0 kHz (bottom), as a function of masker-signal delay. Across frequency and delay, age-related

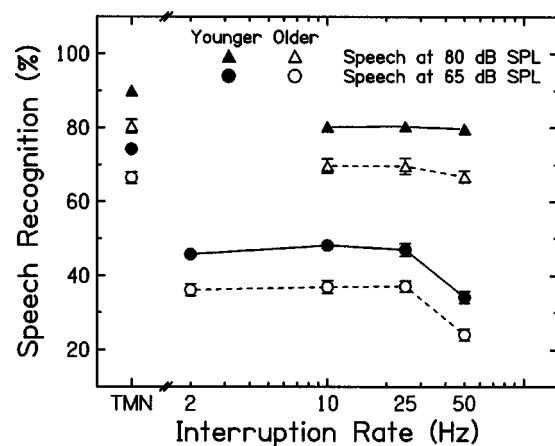


FIG. 5. Mean speech-recognition scores (± 1 SE) for younger and older subjects in interrupted noise as a function of interruption rate and in threshold-matching noise (TMN, left-most data point) for speech at 65 dB SPL (circles) and 80 dB SPL (triangles).

differences in forward masking averaged less than 0.5 dB. Given that there were no significant age-related differences in forward masking, differences observed in forward-masked thresholds in Fig. 3 were likely attributable to age-related differences in quiet thresholds. Also consistent with this conclusion is the finding that for older subjects no significant correlations were observed between age and forward-masked thresholds.

In a previous study of speech recognition in interrupted noise (Dubno *et al.*, 2002), forward-masked and simultaneous-masked thresholds for 20-ms signals were significantly higher for older than younger subjects by 3–4 dB at 0.5, 1.0, 2.0, and 4.0 kHz. Quiet thresholds and simultaneous-masked thresholds for longer signals were also elevated for older subjects but only by 1–2 dB. In that study, it was not possible to determine amount of forward masking because quiet thresholds for 20-ms signals were not available. We concluded that age-related differences in masked thresholds were more likely to be observed for shorter than longer signals. Here, no age-related differences were observed in simultaneous-masked thresholds for short signals and age-related differences in forward masking for short signals were attributed to differences in quiet thresholds. Although the signals were identical for the two studies, masker spectra and levels were different and, in the previous study, thresholds for only one very short masker-signal delay were measured. Further clarification is needed regarding possible age-related effects on psychophysical measures of forward masking.

C. Speech recognition

1. Threshold-matching noise

Figure 5 plots speech-recognition scores for younger and older subjects in TMN and interrupted noise. Consider first the left-most data points, which are speech-recognition scores in TMN alone. As expected, due simply to increased audibility, scores are higher for speech at 80 dB than for speech at 65 dB. However, averaged across speech level, scores were significantly higher by 8.6% for younger sub-

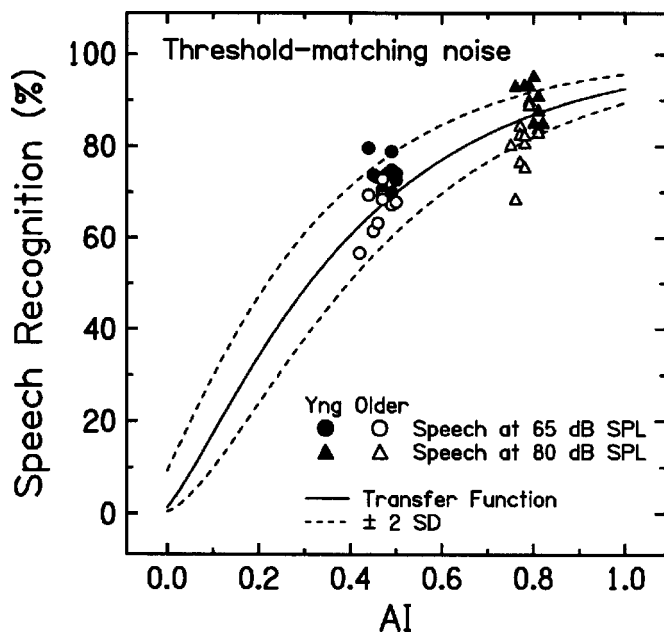


FIG. 6. Speech-recognition scores for individual younger subjects (filled) and older subjects (open) in threshold-matching noise plotted against articulation index (AI) values computed using each subject's thresholds in the threshold-matching noise. Scores were obtained for speech at 65 dB SPL (circles) and for speech at 80 dB SPL (triangles). The solid line is the AI-recognition transfer function for the nonsense syllables and the dashed lines encompass the 95% confidence interval.

jects than for older subjects ($F[1,18]=21.07$, $p=0.0002$). This was an unexpected result, given that subjects had essentially equal audibility and most of the variance in speech-recognition scores may be attributed to differences in audibility. To determine if poorer scores in TMN for older subjects could be attributed to their slightly higher thresholds in TMN, AI values were computed and scores were predicted for each speech level, using procedures similar to ANSI (1969), but using each subject's thresholds in the TMN.

Figure 6 shows scores for younger and older subjects at two speech levels in TMN plotted against AI. The solid line is the normal transfer function relating the AI to consonant recognition established for the speech stimuli used in this experiment (Dirks *et al.*, 1990a,b); the dashed lines encompass the 95% confidence interval. At the lower speech level (circles), scores for nearly all subjects were within the 95% confidence interval and the distribution of scores for younger and older subjects overlapped. However, for equivalent AI values, scores for older subjects fell at the lower end of the distribution of scores for younger subjects. Whereas scores for all younger subjects fell above the transfer function and half the scores for older subjects fell above the transfer function and half fell below. Similar distributions of scores for younger and older subjects were seen for the higher speech level (triangles), but scores for three older subjects were poorer than the 95% confidence interval of their predicted scores. Two of these subjects also had the lowest scores for the lower speech level. These patterns suggested that threshold differences among subjects in the TMN may account for only a small portion of the lower scores observed for older subjects in that masker.

To quantify the contribution of masked thresholds to

age-related differences in scores in TMN, differences were computed between individual observed scores and scores predicted by the AI (represented by the transfer function in Fig. 6). Here, the difference between the observed and predicted score was used as an estimate of speech recognition that was independent of threshold elevation. If poorer scores of older subjects were entirely attributed to reduced audibility related to their thresholds in TMN, differences between observed and predicted scores would be equivalent for younger and older subjects. When effects of TMN thresholds were accounted for, the average difference in scores between younger and older subjects decreased from 8.6% to 7.6%; this group difference remained statistically significant ($F[1,18]=16.84$, $p=0.0007$), with no significant interaction with level ($F[1,18]=1.96$, $p=0.179$). Thus, age-related differences in speech recognition in TMN could not be explained by audibility.

To explore further age-related differences in speech recognition in TMN, percent-correct scores for each consonant were computed from confusion matrices and differences in consonant scores between younger and older subjects were determined. Among the consonants with the largest age-related differences were the three consonants /dʒ/, /ʒ/, and /ð/. It was noted that these consonants do not have simple orthographic representations. Subjects received training with all syllables and correct responses were consistently observed during practice trials. Nevertheless, it is possible that some of the age-related difference could be attributed to difficulty associating these three sounds with their appropriate symbols. With these three consonants removed, the age-related difference in score in TMN was reduced to 5.0% but the difference remained statistically significant ($F[1,17]=13.19$, $p=0.0021$).¹

2. Interrupted noise

a. Effect of speech level. Returning to Fig. 5, the remaining data points are speech-recognition scores in interrupted noise plotted as a function of interruption rate. The higher scores (triangles) are for speech presented at 80 dB SPL and lower scores (circles) are for speech presented at 65 dB SPL. For younger and older subjects, speech recognition declined significantly in interrupted noise relative to scores in TMN, but scores with speech at 65 dB declined more than scores with speech at 80 dB. That is, interrupted noise was a significantly more effective masker of lower level speech than of higher level speech. During the periods when the interrupted noise was off, speech audibility was poorer at the lower speech level (65 dB) than at the higher speech level (80 dB, refer to Fig. 2). Then, during the periods when the noise was on, speech audibility (i.e., signal-to-noise ratio) was much poorer with speech at 65 dB than 80 dB. As a result, relative to scores in TMN alone, scores declined more in interrupted noise with speech at 65 dB than with speech at 80 dB.

Another way to describe these results is by the improvement in speech-recognition scores as speech level increased from 65 to 80 dB. Without interrupted noise, scores increased by ~16% for younger subjects and ~14% for older subjects. These amounts were slightly less than the ~19%

improvement predicted by the AI on the basis of increased speech audibility. In 10- and 25-Hz interrupted noise, scores increased by 32%–33% for younger and older subjects as speech level increased from 65 to 80 dB, which was similar to the 34% improvement predicted by the AI on the basis of increased speech audibility in interrupted noise.² In 50-Hz interrupted noise, however, scores increased 45.4% and 42.6% for younger and older subjects, respectively. The greater improvement for the 50-Hz rate was attributed to poorer scores in lower level speech.

b. Effect of interruption rate. With speech presented at 65 and 80 dB SPL, scores in interrupted noise remained relatively constant as interruption rate increased (Fig. 5). However, for speech at 65 dB, scores for the 50-Hz interruption rate were significantly poorer than scores for the slower rates, for all subjects. For speech at 80 dB, scores for the 50-Hz rate were significantly poorer than those for the slower rates, but for older subjects only. Results for younger subjects are comparable to those in Miller and Licklider (1950, Fig. 8). Their results showed that interrupted maskers were least effective at 10 Hz and, at poor signal-to-noise ratios (comparable to the 65-dB speech condition), scores declined at higher rates. As suggested by Miller and Licklider, at slower rates (in this case, 2, 10, and 25 Hz), listeners benefitted from multiple “glimpses” of each syllable. Syllable durations in the current study ranged from 480 to 851 ms (for consonant-vowel syllables) and from 453 to 950 ms (for vowel-consonant syllables), with consonant durations shorter. Given “off” periods of the interrupted noise ranging from 250 to 20 ms, many silent periods generally occurred during each syllable presentation. In addition, with longer “off” periods, recovery from forward masking was more complete. For the 50-Hz rate, however, the 10-ms “off” period may have been too short to provide sufficient audible information for consonant recognition, and higher amounts of forward masking would persist throughout the “off” period.

c. Effect of age. In the interrupted noise, averaging across speech level and interruption rate, scores were significantly higher by 10.7% for younger than older subjects, an even larger difference than in TMN. In addition, scores for older subjects for the 50-Hz rate were significantly poorer than those for the slower rates, even at the higher speech level. Poorer scores for the fast interruption rate with relatively good speech audibility suggest that older subjects may be less able to extract small bits of speech information or may be disadvantaged by higher forward-masked thresholds.

To determine if age-related differences in scores in interrupted noise were attributable to score differences in TMN alone, the “penalty” of interrupted noise was computed for each subject, i.e., the score in threshold-matching noise minus the score in interrupted noise. Figure 7 plots mean observed penalty as a function of interruption rate for speech at 65 and 80 dB SPL. The right-most pairs of bars are mean penalty predicted from AI values for threshold-matching and interrupted noise; no age-related differences in penalty were predicted. As noted earlier, scores in interrupted noise declined more for 65-dB speech than for 80-dB speech, and more for the 50-Hz rate than for slower rates. Although

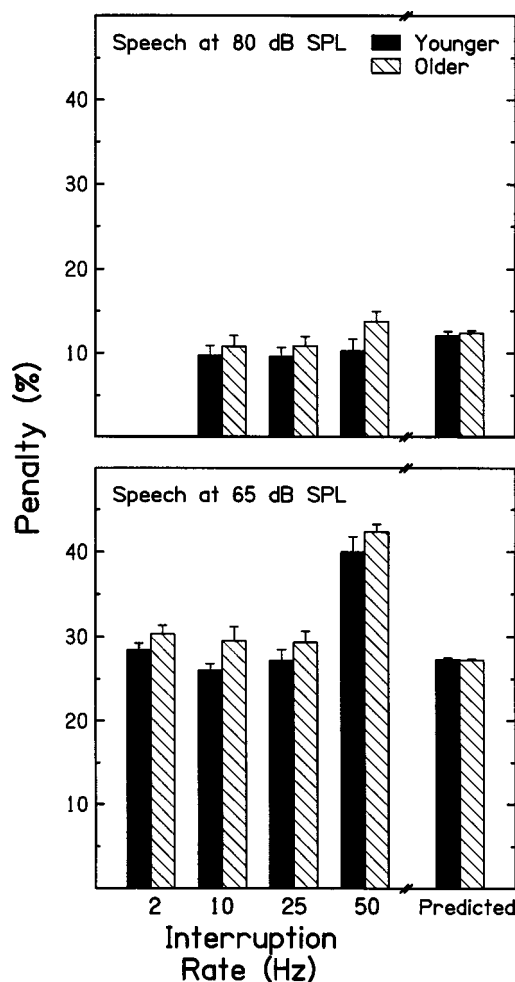


FIG. 7. Mean observed penalty (+1 SE) as a function of interruption rate for speech at 80 dB SPL (top) and 65 dB SPL (bottom). The right-most pairs of bars are mean penalty predicted from AI values for threshold-matching and interrupted noise. Penalty was defined as the score in threshold-matching noise minus the score in interrupted noise.

equivalent penalty was predicted for younger and older subjects, scores for older subjects consistently declined more (i.e., larger penalty) in interrupted noise than scores for younger subjects. These differences did not reach statistical significance ($F[1,18]=3.90$, $p=0.064$). We explored these findings further by repeating the statistical analysis without including results from three older subjects whose scores at the higher speech level in TMN were poorer than predicted by the AI (see Fig. 6). Age-related differences in scores in TMN and in interrupted noise remained statistically significant. Moreover, age-related differences in penalty now also reached statistical significance ($F[1,15]=5.84$, $p=0.02$). A similar analysis was conducted with scores from all subjects computed after removing results for /dʒ/, /ʒ/, and /ð/, as described earlier. Age-related differences in scores in interrupted noise remained statistically significant, and statistically significant age-related differences in penalty were also observed ($F[1,16]=11.3$, $p=0.004$).

Figure 8 shows observed scores for younger and older subjects in interrupted noise plotted against AI. As was the case for TMN (Fig. 6), AI values computed for younger and older subjects were nearly equivalent (consistent with equal

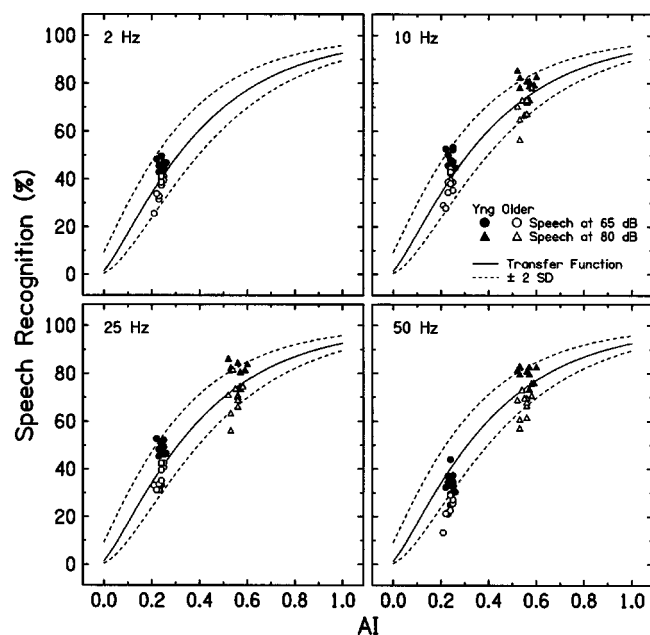


FIG. 8. Speech-recognition scores for individual younger subjects (filled) and older subjects (open) in interrupted noise plotted against articulation index (AI) values computed by taking the mean of the AI for threshold-matching noise and the AI for speech-shaped steady-state noise at 80 dB SPL. Scores for a 2-, 10-, 25-, and 50-Hz interrupted noise are plotted in separate panels. Scores were obtained for speech at 65 dB SPL (circles) and for speech at 80 dB SPL (triangles). The solid line in each panel is the AI-recognition transfer function for the nonsense syllables and the dashed lines encompass the 95% confidence interval.

predicted penalty). However, for equivalent AI values, scores for older subjects fell at the lower end of the distribution of scores for younger subjects. Except for lower level speech in 50-Hz noise, observed scores for younger subjects were better than predicted whereas those for older subjects were poorer than predicted (although most scores were within the 95% confidence interval). These patterns suggested that threshold differences among subjects in TMN or speech-shaped steady-state noise could not account for the lower scores observed for older subjects in interrupted noise.

As described above for TMN, a comparison of observed-predicted differences for younger and older subjects provided an estimate of the age-related difference in scores in interrupted noise that remained after accounting for differences in masked thresholds among subjects. When effects of masked thresholds were accounted for, the average difference in scores in interrupted noise between younger and older subjects decreased only slightly, from 10.7% to 9.8%, a difference that remained statistically significant. Thus, consistent with the pattern of results in Fig. 8, age-related differences in speech recognition in interrupted noise could not be explained by audibility.

In our previous study (Dubno *et al.*, 2002), a significant age-related deficit for speech recognition in interrupted noise was observed. In that study, benefit of interrupted noise was assessed by comparing scores in steady-state speech-shaped noise and the same noise interrupted at a 10-Hz rate. Reduced benefit of interrupted noise for older subjects was only partially attributed to their poorer scores in the steady-state masker. In the current study, rather than measure improve-

ment in speech recognition with interrupted noise relative to steady-state noise, we measured the penalty of interrupted noise on speech recognition, that is, the decline in speech recognition with and without interrupted noise. Direct comparisons between the two studies are not straightforward because of differences in speech and noise levels. In the current study, due to speech and noise level requirements and to avoid floor effects, it was not possible to obtain scores in the steady-state condition for both speech levels. As an alternative, scores in steady-state noise were predicted from AI values computed using masked thresholds for each subject measured in steady-state noise (see Fig. 2). For speech at 80 dB and noise at 80 dB, the predicted score in steady-state noise was 51.3% for both younger and older subjects.³ Thus, given scores in 10-Hz interrupted noise in the current study and assuming equal scores in steady-state noise for younger and older subjects, benefit was estimated to be 29.0% for younger subjects and 18.4% for older subjects. This is comparable to benefit for 10-Hz interrupted noise in our previous study with similar speech and noise levels (27.9% for younger subjects, 22.2% for older subjects). Indeed, if scores in steady-state noise for older subjects were assumed to be somewhat poorer than for younger subjects (as was observed previously), differences between the two studies would be even smaller. This suggests that the conclusions of the two studies are generally similar, that is, that older subjects benefit less than younger subjects from momentary improvements in signal-to-noise ratio in interrupted maskers.

d. Relationship between forward masking and speech recognition in interrupted noise. Recall that in this experiment speech recognition in interrupted noise was regarded as analogous to recovery of the response to a suprathreshold signal from prior stimulation by a masker. Accordingly, linear regression analysis was used to assess the relationship between scores in interrupted noise and forward-masked thresholds. Regression analyses included forward-masked thresholds at 0.5 and 4.0 kHz measured in each of the four masker-signal delays, slopes and y-intercepts of the forward-masked threshold functions at 0.5 and 4.0 kHz fitted by linear functions, and parameter values of the recovery functions fitted by the equations given in Jesteadt *et al.* (1982). There were no significant correlations with slopes or with any of the recovery-function parameters from Jesteadt *et al.* (1982). Correlations with forward-masked thresholds at the four masker-signal delays were comparable to those with the y-intercepts of the linear functions; for clarity, results described in the following sections are those using the forward-masked thresholds.

Negative correlations were observed between scores in interrupted noise and forward-masked thresholds at both 0.5 and 4.0 kHz, suggesting that speech recognition in interrupted noise decreased as forward-masked thresholds increased. To illustrate these relationships, Figs. 9 and 10 show scores for 65-dB speech in 25- and 50-Hz interrupted noise, respectively, plotted against forward-masked thresholds at 4.0 kHz. Data for the four masker-signal delays are presented in different panels; results for younger and older subjects are indicated by filled and open symbols, respectively. Linear regression functions were fitted to the pooled data.

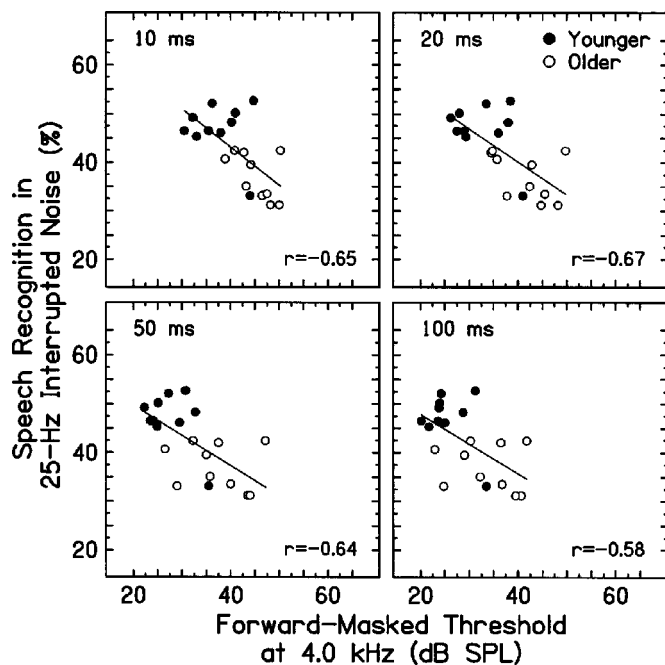


FIG. 9. Recognition scores for speech at 65 dB SPL in 25-Hz interrupted noise plotted against forward-masked thresholds at 4.0 kHz. Data for the four masker-signal delays are presented in different panels; results for younger and older subjects are indicated by filled and open symbols, respectively. Linear regression functions were fitted to the pooled data. Correlation coefficients are given in each panel ($p < 0.01$ for each masker-signal delay).

Generally, correlation coefficients were stronger and more reached statistical significance for the lower than the higher level speech and for the higher than the lower frequency signal. The dependence on speech level means that speech recognition in interrupted noise may be more dependent on recovery from forward masking when speech audi-

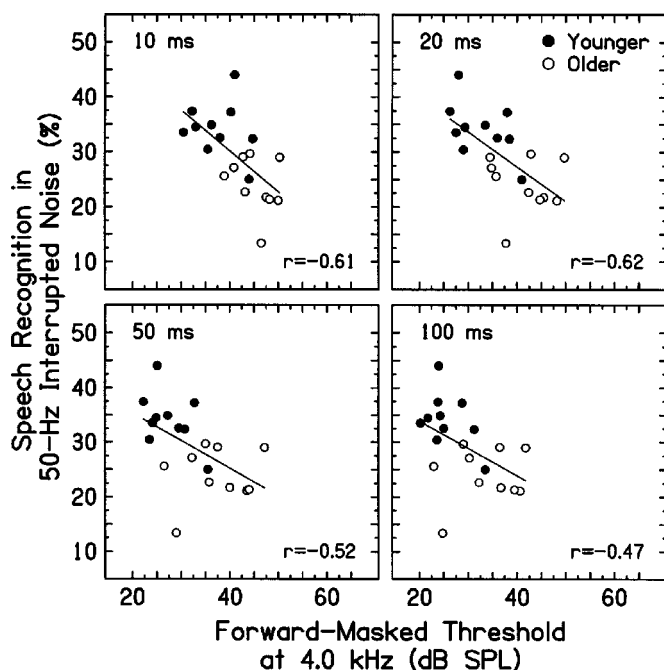


FIG. 10. Same as Fig. 9, but for 50-Hz interrupted noise. Correlation coefficients are given in each panel ($p < 0.01$ for 10- and 20-ms masker-signal delays; $p < 0.05$ for 50 and 100 ms).

bility (i.e., signal-to-noise ratio) is relatively poor; a similar dependence on noise level was observed in our previous study. The frequency dependence observed in the association between scores in interrupted noise and forward masking was also observed in our previous study and may relate to the relative importance of lower versus higher frequencies to nonsense syllable recognition, the shape of the noise spectrum, or differences in the temporal properties of higher and lower frequency auditory nerve fibers (see Dubno *et al.*, 2002 for additional discussion). A systematic pattern was also observed for the various masker-signal delays. For forward masking at 0.5 kHz, correlations with scores in interrupted noise were strongest for the longest delays; for 4.0 kHz, correlations were strongest for the shortest delays. This pattern could not be attributed to threshold differences between younger and older subjects as a function of masker-signal delay or to differences in ranges of scores or thresholds across conditions.

Finally, in the lower level speech, stronger correlations between scores in interrupted noise and forward-masked thresholds were observed for the two faster noise interruption rates than for the slower rates. As noted earlier, with 65-dB speech, scores declined significantly as the noise interruption rate increased to 50 Hz, perhaps because the 10-ms “off” period provided less audible information or because there was not sufficient time for recovery from forward masking by the noise burst. Accordingly, the shorter “off” times in the 25- and 50-Hz interrupted noises may have been more sensitive to small differences in forward-masked thresholds and the subsequent effects on speech audibility.

Assuming an age-related decline in activity of low-SR fibers (Schmiedt *et al.*, 1996), we hypothesized that older subjects would benefit less from the contribution of these fibers than younger subjects. Recall that previous psychophysical data suggested that the encoding ability of low-SR fibers was diminished for mid-level signals compared to higher level signals (Zeng *et al.*, 1991). Given these findings, we further hypothesized that younger subjects would derive more benefit from low-SR fibers at the higher speech level than at the lower speech level. Thus, we predicted less improvement in speech recognition for older than younger subjects as speech level increased from 65 to 80 dB. This was not observed (although a nonsignificant trend was seen in the 50-Hz interrupted noise).

The absence of the hypothesized level effect may be attributed to several factors. First, the moderate (65 dB) and high (80 dB) speech levels and the noise level (80 dB) used in this experiment may not have been best suited to reveal the contributions of low-SR fibers. Ideally, based on results of Zeng *et al.* (1991), the “moderate” speech level would be lower and the noise level would be higher. However, the choice of speech and noise levels was restricted by the need to avoid floor and ceiling effects in speech-recognition measures. Second, given the slow recovery of the low-SR fibers, with the interrupted noise on continuously, it is possible that these fibers would not be completely recovered between each syllable presentation. Third, according to Schmiedt *et al.* (1996), the inactivity of low-SR fibers may relate to a chronically depressed EP in aged gerbils. Given the strong relation-

ship between threshold shift and loss of EP (Sewell, 1984), it is possible that these older subjects with exceptionally good hearing (at least at frequencies <6.0 kHz) maintained an EP and distribution of low-SR fibers more like younger subjects. Fourth, the change in recovery functions shown by Schmiedt *et al.* (1996) in aging gerbils was observed only at high frequencies (>6.0 kHz). A recent study of forward-masked intensity discrimination by He *et al.* (2002) found larger age-related changes at 1.0 kHz than at 6.0 kHz, which was consistent with age-related declines in intensity discrimination at lower frequencies observed previously (He *et al.*, 1998) but not with an age-related loss of low-SR fibers at higher frequencies. Thus, the functional effects of a loss of low-SR fibers at high frequencies in older human subjects, and the extent of any such loss, are not yet known. Nevertheless, subjects in the current experiment and in Dubno *et al.* (2002) had substantial threshold elevations at frequencies >6.0 kHz (see Fig. 1, right panel). Given the specialized characteristics of low-SR fibers, especially as they pertain to resistance to masking and preservation of timing and amplitude modulation, a selective loss of these fibers, even if only at high frequencies, may underlie age-related difficulties in interrupted maskers. That is, functional changes in the base of the cochlea may alter the use of lower frequency cues for preserving timing information that may be important for understanding speech in time-varying noise (Joris *et al.*, 1994; Horwitz *et al.*, 2002). Finally, elevated forward-masked thresholds and reduced benefit to speech recognition from interrupted noise could also result from age-related changes in the auditory central nervous system, which may make the contribution of low-SR fiber response properties more difficult to detect.

IV. SUMMARY AND CONCLUSIONS

Younger and older adults with normal hearing listened to nonsense syllables at moderate and high levels in a speech-shaped noise that was modulated by a 2-, 10-, 25-, or 50-Hz square wave. An additional low-level "threshold-matching noise" was always present which was shaped to produce equivalent masked thresholds for all subjects. To assess recovery from forward masking, forward-masked thresholds at 0.5 and 4.0 kHz were measured as a function of the delay between the speech-shaped masker and the signal. Results may be summarized as follows:

- (1) Forward-masked thresholds of older subjects were higher than those of younger subjects. However, there were no age-related differences in the amount of forward masking or in simultaneous masking.
- (2) Relative to scores without interrupted noise, speech recognition in interrupted noise declined more for lower level speech than for higher level speech. For lower level speech, scores for all subjects remained constant across lower interruption rates and declined for the 50-Hz interruption rate. Scores for older subjects declined in the 50-Hz interrupted noise for the higher speech level as well.
- (3) Speech recognition measured in threshold-matching noise alone was poorer for older than younger subjects.

Similarly, speech recognition in interrupted noise was poorer for older than younger subjects. These age-related differences could not be attributed to differences in audibility among subjects. Small but consistent age-related differences were observed in the decrease in score with interrupted noise, although equivalent decreases in scores were predicted for younger and older subjects. Thus, older subjects benefit less than younger subjects from momentary improvements in speech audibility in interrupted maskers.

- (4) Negative correlations were observed between speech-recognition scores in interrupted noise and forward-masked thresholds at 0.5 and 4.0 kHz. Thus, the benefit derived from momentary improvements in speech audibility in an interrupted noise decreased as forward-masked thresholds increased. Stronger associations with forward masking were observed for the higher frequency signal, for higher noise interruption rates, and when the signal-to-noise ratio was poor.
- (5) Only a small age-related difference was observed in the increase in speech-recognition score as speech level increased from a moderate to high level. This was not consistent with the hypothesis of an age-related difference in the contribution of low-spontaneous-rate fibers to speech recognition in interrupted noise.

ACKNOWLEDGMENTS

This work was supported (in part) by Grant Nos. P50 DC00422 and R01 DC00184 from NIH/NIDCD, and the MUSC General Clinical Research Center (M01 RR 01070). The authors thank Chris Ahlstrom for computer and signal-processing support, Fu-Shing Lee for advice on data analysis, and John H. Mills and Richard A. Schmiedt for editorial comments.

¹Due to an error in retrieving confusion matrices, percent correct scores for individual consonants could not be obtained for all conditions for one older subject. Therefore, results for this subject were not included in this latter analysis.

²AI values for interrupted noise were computed by taking the mean of the AI for threshold-matching noise and the AI for speech-shaped steady-state noise at 80 dB SPL. AI values for threshold-matching noise and steady-state noise were computed using each subject's thresholds in those noises. See Dubno *et al.* (2002) for additional details on computing AI values for interrupted noise.

³For speech at 65 dB and noise at 80 dB, the predicted score in steady-state noise was 0%. The minimum speech level for measurable scores was 68 dB. As such, estimates of benefit of interrupted noise for comparison to the previous study were not possible.

Abel, S. M., Krever, E. M., and Alberti, P. W. (1990). "Auditory detection, discrimination, and speech processing in ageing, noise-sensitive and hearing-impaired listeners," *Scand. Audiol.* **19**, 43–54.

ANSI (1969). ANSI S3.5-1969, "Methods for the calculation of the articulation index" (American National Standards Institute, New York).

ANSI (1996). ANSI S3.6-1996, "Specifications for audiometers" (American National Standards Institute, New York).

Bacon, S. P., Opie, J. M., and Montoya, D. Y. (1998). "The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds," *J. Speech Lang. Hear. Res.* **41**, 549–563.

Bronkhorst, A. W., and Plomp, R. (1992). "Effect of multiple speech like maskers on binaural speech recognition in normal and impaired hearing," *J. Acoust. Soc. Am.* **92**, 3132–3139.

- Dirks, D. D., Dubno, J. R., Ahlstrom, J. B., and Schaefer, A. B. (1990a). "Articulation index importance and transfer functions for several speech materials," *Asha* **32**, 91.
- Dirks, D. D., Dubno, J. R., and Bell, T. S. (1990b). "Problems in the prediction of speech recognition among normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am. Suppl. 1* **88**, S32.
- Dubno, J. R., and Ahlstrom, J. B. (1997). "Additivity of multiple maskers of speech," in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Hillsdale, NJ), pp. 253–272.
- Dubno, J. R., and Schaefer, A. B. (1992). "Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked-normal hearing listeners," *J. Acoust. Soc. Am.* **91**, 2110–2121.
- Dubno, J. R., and Schaefer, A. B. (1995). "Frequency selectivity and consonant recognition for hearing-impaired and normal-hearing listeners with equivalent masked thresholds," *J. Acoust. Soc. Am.* **97**, 1165–1174.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2002). "Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing," *J. Acoust. Soc. Am.* **111**, 2897–2907.
- Eisenberg, L. S., Dirks, D. D., and Bell, T. S. (1995). "Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing," *J. Speech Hear. Res.* **38**, 222–233.
- Festen, J. M. (1993). "Contributions of comodulation masking release and temporal resolution to the speech-recognition threshold masked by an interfering voice," *J. Acoust. Soc. Am.* **94**, 1295–1300.
- Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *J. Acoust. Soc. Am.* **88**, 1725–1736.
- Fitzgibbons, P. J., and Gordon-Salant, S. (1994). "Age effects on measures of auditory duration discrimination," *J. Speech Hear. Res.* **37**, 662–670.
- Green, D. M. (1993). "A maximum-likelihood method for estimating thresholds in a yes–no task," *J. Acoust. Soc. Am.* **93**, 2096–2105.
- Gustafsson, H. A., and Arlinger, S. A. (1994). "Masking of speech by amplitude-modulated noise," *J. Acoust. Soc. Am.* **95**, 518–529.
- He, N.-j., Dubno, J. R., and Mills, J. H. (1998). "Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects," *J. Acoust. Soc. Am.* **103**, 553–565.
- He, N.-j., Dubno, J. R., and Mills, J. H. (2002). "Forward-masked intensity discrimination measured from young and aged subjects with normal hearing," *J. Acoust. Soc. Am.* submitted.
- He, N.-j., Horwitz, A. R., Dubno, J. R., and Mills, J. H. (1999). "Psychometric functions for gap detection in noise measured from young and aged subjects," *J. Acoust. Soc. Am.* **106**, 966–978.
- Horwitz, A. R., Dubno, J. R., and Ahlstrom, J. B. (2002). "Recognition of low-pass-filtered consonants in noise with normal and impaired high-frequency hearing," *J. Acoust. Soc. Am.* **111**, 409–416.
- Howard-Jones, P. A., and Rosen, S. (1993). "The perception of speech in fluctuating noise," *Acustica* **78**, 258–272.
- Jesteadt, W., Bacon, S. P., and Lehman, J. R. (1982). "Forward masking as a function of frequency, masker level, and signal delay," *J. Acoust. Soc. Am.* **71**, 950–962.
- Joris, P. X., Smith, P. H., and Yin, T. C. T. (1994). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail," *J. Neurophysiol.* **71**, 1037–1051.
- Leek, M. R., Dubno, J. R., He, N.-j., and Ahlstrom, J. B. (2000). "Experience with a yes–no single-interval maximum-likelihood procedure," *J. Acoust. Soc. Am.* **107**, 2674–2684.
- Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.* **22**, 167–173.
- Murmane, O. D., Prieve, B. A., and Relkin, E. M. (1998). "Recovery of the human compound action potential following prior stimulation," *Hear. Res.* **124**, 182–189.
- Peters, R. W., and Hall, J. W. (1994). "Comodulation masking release for elderly listeners with relatively normal audiograms," *J. Acoust. Soc. Am.* **96**, 2674–2682.
- Phillips, S. L., Gordon-Salant, S., Fitzgibbons, P. J., and Yeni-Komshian, G. H. (1994). "Auditory duration discrimination in young and elderly listeners with normal hearing," *J. Am. Acad. Audiol.* **5**, 210–215.
- Relkin, E. M., and Doucet, J. R. (1991). "Recovery from prior stimulation. I: Relationship to spontaneous firing rates of primary auditory neurons," *Hear. Res.* **55**, 215–222.
- Relkin, E. M., Doucet, J. R., and Sterns, A. (1995). "Recovery of the compound action potential following prior stimulation: Evidence for a slow component that reflects recovery of low spontaneous-rate auditory neurons," *Hear. Res.* **83**, 183–189.
- Salvi, R. J., Saunders, S. S., Ahroon, W. A., Shivapuja, B. G., and Arehole, S. (1986). "Psychophysical and physiological aspects of auditory temporal processing in listeners with noise-induced sensorineural hearing loss," in *Basic and Applied Aspects of Noise-Induced Hearing Loss*, edited by R. J. Salvi, D. Henderson, and R. P. Hamernik (Plenum, New York), pp. 179–194.
- Schmiedt, R. A., Mills, J. H., and Boettcher, F. A. (1996). "Age-related loss of activity of auditory-nerve fibers," *J. Neurophysiol.* **76**, 2799–2803.
- Schulte, B. A., and Schmiedt, R. A. (1992). "Lateral wall Na, K-ATPase and endocochlear potentials decline with age in quiet-reared gerbils," *Hear. Res.* **61**, 35–46.
- Sewell, W. F. (1984). "The relation between the endocochlear potential and spontaneous activity in auditory nerve fibers of the cat," *J. Physiol. (London)* **347**, 685–696.
- Snell, K. B. (1997). "Age-related changes in temporal gap detection," *J. Acoust. Soc. Am.* **101**, 2214–2220.
- Souza, P. E., and Turner, C. W. (1994). "Masking of speech in young and elderly listeners with hearing loss," *J. Speech Hear. Res.* **37**, 665–661.
- Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (1998). "Temporal processing in the aging auditory system," *J. Acoust. Soc. Am.* **104**, 2385–2399.
- Stuart, A., and Phillips, D. P. (1996). "Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing and presbycusic listeners," *Ear Hear.* **17**, 478–489.
- Takahashi, G. A., and Bacon, S. P. (1992). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," *J. Speech Hear. Res.* **35**, 1410–1421.
- Walton, J., Orlando, M., and Burkard, R. (1999). "Auditory brainstem response forward-masking recovery functions in older humans with normal hearing," *Hear. Res.* **127**, 86–94.
- Zeng, F. G., Turner, C. W., and Relkin, E. M. (1991). "Recovery from prior stimulation II: Effects upon intensity discrimination," *Hear. Res.* **55**, 223–230.

Segmental intelligibility of four currently used text-to-speech synthesis methods

Horabail S. Venkatagiri^{a)}

Department of Psychology, 210 Pearson Hall, Iowa State University, Ames, Iowa 50011

(Received 31 March 2002; accepted for publication 7 January 2003)

The study investigated the segmental intelligibility of four currently available text-to-speech (TTS) products under 0-dB and 5-dB signal-to-noise ratios. The products were IBM ViaVoice™ version 5.1, which uses formant coding, Festival version 1.4.2, a diphone-based LPC TTS product, AT&T Next-Gen™, a half-phone-based TTS product that uses harmonic-plus-noise method for synthesis, and FlexVoice™2, a hybrid TTS product that combines concatenative and formant coding techniques. Overall, concatenative techniques were more intelligible than formant or hybrid techniques, with formant coding slightly better at modeling vowels and concatenative techniques marginally better at synthesizing consonants. No TTS product was better at resisting noise interference than others, although all were more intelligible at 5 dB than at 0-dB SNR. The better TTS products in this study were, on the average, 22% less intelligible and had about 3 times more phoneme errors than human voice under comparable listening conditions. The hybrid TTS technology of FlexVoice had the lowest intelligibility and highest error rates. There were discernible patterns of errors for stops, fricatives, and nasals. Unrestricted TTS output—e-mail messages, news reports, and so on—under high noise conditions prevalent in automobiles, airports, etc. will likely challenge the listeners. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1558356]

PACS numbers: 43.72.Ja, 43.72.Dv, 43.71.Gv [DOS]

I. INTRODUCTION

Text-to-speech (TTS) systems that generate speech from electronically stored text have emerged from the laboratories to find a niche in serious, mission-critical applications in industry, commerce, education, and consumer products. It is anticipated that, in the future, we will be able to listen to a large portion of the information we presently receive in text form (Sproat, Ostendorf, and Hunt, 1999). The settings in which this might occur include challenging listening environments such as automobiles, airports, and offices. These real-world applications require that TTS systems produce speech that is both highly intelligible and natural sounding.

More than a decade ago, Pisoni and his colleagues (Green, Logan, and Pisoni, 1986; Logan, Greene, and Pisoni, 1989; Pisoni, Nusbaum, and Green, 1985) reported on the segmental (phonemic) intelligibility of a number of then-commercially available TTS products. They found a wide range of intelligibility with segmental error rates ranging from 12.92% to 73.97% for different TTS products (Logan, Greene, and Pisoni, 1989). Clearly, the error rates were too high in these products for them to be of use in most real-life applications. One would expect that, over the period of a decade and a half, there would be significant improvement in their performance. A goal of the present study is to assess the segmental intelligibility of selected currently available TTS products to evaluate their suitability for use in voice response systems, call centers, web content delivery, e-mail reading, and other important real-life applications.

The conversion of text to speech typically takes place in three stages (Klatt, 1987). In the first stage, the text is trans-

formed into a phonetic code, which is a complicated task in languages such as English where spelling is notoriously unreliable as a guide to pronunciation. Next, values for prosodic elements—pitch, intonation, rate, and pausing—are added to the phonetic representation. In both phonetic and prosodic processing stages, the TTS system utilizes not only the phonetic and phonological knowledge of the language but also a limited analysis of the semantic, syntactic, and morphological structure of the text in an attempt to generate speech that accurately represents the meaning and intent of the text.

The final stage of text to speech conversion, speech synthesis, involves conversion of the phonetic code into speech waveform. There are broadly two different methods used for speech synthesis. In formant TTS synthesis, the phonemes represented in the phonetic code are synthesized by retrieving acoustic data for each phoneme and selected allophones from a database. The database consists of a variety of relevant acoustic parameters for each phoneme such as the center frequencies, bandwidths, and amplitudes of formants, frication, aspiration, and nasal resonance (Allen, Hunnicutt, and Klatt, 1987).

In concatenative TTS synthesis, parameterized units of natural speech are joined together to produce speech. The units of natural speech used most commonly are the diphones. A diphone extends from the center of one phoneme to the center of the next phoneme and, therefore, preserves the crucial features of interphoneme coarticulation. Linear predictive coding (LPC) is frequently used as the back-end synthesis method for concatenative TTS synthesis (Atal and Hanauer, 1971; O'Shaughnessy, 1987). The input parameters for LPC synthesizers are typically the predictor coefficients, derived from LPC analysis of units of natural speech pro-

^{a)}Electronic mail: giri@iastate.edu

duced by a speaker. Although LPC-based TTS synthesis, a frequency-domain approach, is popular, there are also waveform-based concatenative TTS synthesis methods including the time-domain pitch-synchronous overlap-add (TD-PSOLA) (Moulines and Charpentier, 1990), MBE resynthesis overlap add (MBROLA) (Dutoit and Leich, 1993), and the harmonic-plus-noise model (HNM) (Stylianou, 2000). These time-domain concatenative TTS methods afford greater control over prosodic variables such as pitch, duration, and loudness of segments than LPC synthesis.

A second goal of the present study is to determine whether different speech synthesis techniques as implemented in selected commercially available TTS products differ in intelligibility. Current speech synthesizers typically retrieve the phonetic form of the word from a pronunciation dictionary and use pronunciation rules only for uncommon words and proper names. The present study uses a set of common monosyllabic words with predictable stress and intonation to minimize differences in linguistic preprocessing (i.e., conversion of text into phonetic code) among the four products investigated so that the differences in intelligibility will be largely attributable to differences in synthesis techniques. To date, there does not appear to be a comprehensive controlled study comparing the segmental intelligibility of formant, LPC, waveform, and formant-LPC hybrid synthesis techniques.

Finally, TTS is being considered for use in noisy environments such as for mobile computing (e.g., web content delivery and e-mail reading in automobiles) and for providing information through kiosks and announcements in airports and shopping malls. The presence of competing acoustic signals in such environments interferes with the perception of speech. It has been shown that TTS suffers significant intelligibility loss in noisy listening conditions (Fucci *et al.*, 1995; Koul and Allen, 1993; Pisoni and Koen, 1982). However, it is not known whether some synthesis techniques are better than others in resisting noise interference. Therefore, a third goal of the present study is to assess the segmental intelligibility of selected TTS products at adverse signal-to-noise ratios.

II. METHOD

The four TTS products included in this study use four different methods for generating speech from text. AT & T Next-Gen TTS™ (<http://www.research.att.com/projects/tts/>) employs a half-phone-based synthesis typically using the HNM method. Festival version 1.4.2 (<http://cslu.cse.ogi.edu/research/tts.htm>), developed at the University of Edinburgh, U.K., is an open source, freely available diphone-based LPC synthesis product. FlexVoice™ 2 is a hybrid TTS product that combines the features of concatenative and formant synthesis methods (<http://www.flexvoice.com/index.html>). IBM ViaVoice™ version 5.1 (http://www-4.ibm.com/software/speech/dev/ttsdk_windows.html), uses allophone-based formant coding. In each case, the version of the product available during October–November, 2001 on their respective web sites was used in this study.

A. Stimulus materials

The stimulus materials were taken from the Modified Rhyme Test (House *et al.*, 1965), which consists of 300 monosyllabic English words arranged in six lists, A, B, C, D, E, and F, of 50 words each. Each list is further divided into two 25-words sublists, e.g., A1 and A2, each sublist systematically varying the final and initial consonant sounds, respectively. This study utilized lists A, B, C, and D after randomizing the words in each sublist to eliminate the original semialphabetical order.

All 200 words were synthesized using each of the four TTS products at a sampling rate of 16 kHz and 16-bits quantization and stored in the *wav* file format on a computer hard disk. An adult male voice speaking in the American accent was chosen for all four voices. Festival and IBM ViaVoice words were synthesized from software downloaded and installed on a local computer. AT & T Next-Gen and FlexVoice words were synthesized directly on their web sites using their interactive demo facilities. Each word was preceded by the carrier phrase “The word is _____” in the same voice. A sound file was created for each 25-word sublist with an interstimulus interval of 5 seconds using the COOLETIT 2000 (Syntrillium Software Corporation, Phoenix, AZ) sound-editing software. A stimulus was defined as the carrier phrase followed by an MRT word. The volume levels of the 32 sound files (8 sublists×4 TTS products) were equated to a mean of between −35.5 and −35 dB power using the COOLETIT program. Apart from this overall control over the power of each sound file, the intensity of individual words and sounds varied depending on the values used by the synthesis software programs.

B. Participants

The listeners in this study were 32 people, 12 males and 20 females, with a mean age of 24 years (range: 17–61 years). All were native speakers of English and none had taken part in any previous research involving TTS. The participants passed a hearing screening at 30 dB HL (ANSI, 1969) at frequencies 500, 1000, and 2000 Hz bilaterally, administered in a laboratory room. Each participant listened to 8 tapes of 25 words each, 4 tapes at a signal-to-noise ratio (SNR) of 5 dB and the remaining 4 tapes at a SNR of 0 dB. The four tapes in each of the two SNR conditions were in the four different TTS voices—AT&T Next-Gen, FlexVoice, IBM ViaVoice, and Festival. The word lists, SNR conditions, and TTS products were systematically counterbalanced so that an equal number of listeners experienced each condition in a comparable manner to minimize the effects of differences in linguistic content and the order of presentation.

C. Stimulus presentation

The noise used in the present study was the multitalker speech babble obtained from Auditec (St. Louis, MO). This noise consists of 20 young adults reading different passages simultaneously. The resulting recording is mostly unintelligible and the noise level is nearly uniform over time. Speech babble somewhat approximates the noise encountered in real-life situations and has been shown to interfere with

speech intelligibility more than spectrally matched noise (Kalikow, Stevens, and Elliott, 1977; Lewis *et al.*, 1988).

The stimulus phrases were played off of one channel of the stereo setup of COOLETIT 2000 installed on a Gateway Solo 5150 laptop computer (Gateway, San Diego, CA). The babble noise was played on the second channel. The output from the computer was routed to a GSI 16 audiometer (Grason-Stadler, Littleton, MA) and from there to a pair of TDH 50-P earphones (Telephonics Corporation, Farmingdale, NY). The peak output level of the stimulus words at each earphone was 65 dB(A) (± 2 dB) measured using a 6-cc coupler on a Quest model 215 sound-level meter (Metrosonics, Oconomowoc, WI). The peak noise level in each earphone was 60 and 65 dB(A) (± 1 dB) for the 5-dB and 0-dB SNR conditions, respectively.

Participants took part in the listening experiment individually, seated in a double-walled sound-treated room (Industrial Acoustics Company, Bronx, NY), wearing the earphones. They were given a response form with numbered blank lines arranged in columns and requested to write down the last word they heard in each phrase. The listeners were encouraged to guess if they were not sure of what they heard and to print words for improved legibility.

D. Human voice

As part of another study, a 22-year-old male college student with American (Midwestern) English accent recorded words from lists E1 and E2 of MRT. House *et al.* (1965) have demonstrated that the six lists of MRT words produce statistically equivalent results. Therefore, it is possible to make valid comparisons between results obtained from lists A–D for TTS as described above with results obtained from list E for human voice. The student spoke into a head-mounted Telex USB microphone model H-531 (Telex Communications, Minneapolis, MN) positioned approximately 2 cm from and slightly below the mouth. The stimulus tapes of human voice were prepared in exactly the same manner as the tapes of TTS voices described above.

A group of 32 participants who were not part of the group that listened to the TTS products listened to the two lists at 0-dB and 5-dB SNR, respectively, in conditions comparable to those for the TTS products. This group consisted of 15 males and 17 females with a mean age of 20 years (range: 17–42 years). All participants were native speakers of English and passed a hearing screening at 30 dB HL (ANSI, 1969) for frequencies 500, 1000, and 2000 Hz. The two word lists and the order of presentation of the two lists were counterbalanced across participants.

III. RESULTS

Errors were tallied from the response sheets. Alternate spelling (e.g., *scene* for *seen*) and phonetic spelling (*duk* for *duck*) were accepted as correct. The number of words without errors, and the number of errors in the initial, medial, and final position of words were tallied for each TTS product and for the human voice in each of the two SNR conditions.

A. Whole words correct

Table I shows the number of words without any errors for each of the four TTS products. The number of complete words correctly identified by listeners is a useful measure of the relative intelligibility of TTS products, because the presence of any error in a word would potentially alter the meaning or render the word meaningless. A two-factor (4 TTS products \times 2 SNR conditions) repeated-measures analysis of variance (ANOVA) model (Cohen, 1996) was used to determine whether different TTS products and SNR conditions differed significantly for whole words correct. All statistical analyses in this paper were carried out using the SPSS (release 10.0.5) statistical package (SPSS, Inc., Chicago, IL). The results of the ANOVA showed that the TTS products ($F [1.69, 52.30] = 61.79$, Greenhouse–Geisser¹ $p < 0.0001$) and SNR conditions ($F [1, 31] = 39.31$, Greenhouse–Geisser $p < 0.0001$) differed significantly for whole words correct. The TTS products \times SNR conditions interaction, however, was nonsignificant ($F [3, 93] = 2.0$, $p > 0.1$), indicating that TTS products and SNR conditions exerted independent influence on the number of whole words correct. Although signal-to-noise ratio was an important factor in the intelligibility of TTS products, SNR did not have a differential effect based on the TTS synthesis algorithms. The 0-dB SNR ($M = 52.19$) condition across TTS products was less intelligible than the 5-dB SNR condition ($M = 60.97$) regardless of TTS synthesis algorithms.

The *post hoc* multiple comparisons among the four TTS products were made using the Bonferroni test (Cohen, 1996) to control for type I statistical error. With six possible comparisons, the Bonferroni $\alpha = 0.05/6 = 0.008$. For this and other Bonferroni tests in this paper, separate error terms based on each pair rather than $\sqrt{MS_{\text{error}}}$ from repeated measures ANOVA were used. The t-tests for correlated samples for TTS products revealed that the AT&T NextGen ($M = 66.25$) and Festival ($M = 68.19$) were statistically equivalent in intelligibility as measured on whole words correct ($t [31] = -0.922$, 2-tailed $p > 0.3$). IBM ViaVoice ($M = 58.81$) was significantly less intelligible than AT&T NextGen ($t [31] = -3.52$, 2-tailed $p = 0.001$) and Festival ($t [31] = -3.47$, 2-tailed $p = 0.002$).

IBM ViaVoice ($t [31] = 17.58$, 2-tailed $p < 0.0001$), AT&T NextGen ($t [31] = 13.57$, 2-tailed $p < 0.0001$), and Festival ($t [31] = -11.49.58$, 2-tailed $p < 0.0001$) were significantly more intelligible than FlexVoice ($M = 33.00$). TTS synthesis algorithms that use a natural speech database (AT&T NextGen and Festival) were slightly more intelligible than those that use an acoustic database (IBM ViaVoice). However, the hybrid synthesis technology that combined the elements of formant and concatenative synthesis methods as implemented in FlexVoice was significantly less intelligible than either concatenative or formant synthesis technique alone. The two different units of speech used in the natural speech databases, diphones (Festival) and half-phones (AT&T NextGen), and the difference in synthesis methods—LPC (Festival) and harmonic-plus-noise (AT&T NextGen)—did not make a difference in intelligibility as measured on whole words correct.

TABLE I. Percent whole words correct and percent phoneme errors for human and TTS products under 0 and 5 dB SNR conditions.

		Human		AT&T NextGen		Festival		FlexVoice		IBM ViaVoice	
Whole words correct	Mean		86.81		66.25		68.19		33.00		58.81
	s.d.		5.10		8.01		11.16		9.72		8.08
	Range		78.0–96.0		48.0–84.0		40.0–92.0		12.0–56.0		42.0–72.0
Total errors	Mean		5.29		16.85		15.27		38.50		17.06
	s.d.		2.36		5.30		5.44		6.40		3.42
	Range		1.3–10.0		6.0–28.7		3.3–30.0		26.0–56.7		10.7–23.3
Consonant errors	Mean		5.59		17.78		16.69		43.66		22.00
	s.d.		2.51		6.07		6.94		7.83		4.51
	Range		1.0–10.0		6.0–29.0		3.0–36.0		29.0–62.0		14.0–31.0
Vowel errors	Mean		4.63		15.00		12.44		28.19		7.19
	s.d.		3.07		6.14		5.07		9.86		3.66
	Range		0.0–12.0		4.0–28.0		4.0–26.0		12.0–50.0		0.0–14.0
Signal-to-noise ratio		0 dB	5 dB	0 dB	5 dB	0 dB	5 dB	0 dB	5 dB	0 dB	5 dB
Whole words correct	Mean	83.50	90.62	60.00	70.50	63.25	71.50	33.13	36.75	54.38	66.38
	s.d.	7.85	5.62	9.48	8.78	17.14	13.62	14.59	13.30	11.75	11.62
	Range	64.0–92.0	76.0–100	36.0–76.0	44.0–88.0	24.0–96.0	40.0–88.0	8.0–84.0	16.0–84.0	28.0–88.0	40.0–84.0
Total errors	Mean	6.85	3.72	19.50	14.21	17.88	12.67	39.54	37.46	20.13	14.0
	s.d.	3.68	2.52	6.43	5.57	8.82	5.77	8.66	6.46	5.10	5.16
	Range	2.7–16.4	0.0–10.7	5.3–33.3	5.3–29.3	1.3–37.3	5.3–26.7	24.0–61.3	24.0–52.0	8.0–29.3	5.3–25.3
Initial sound (consonants)	Mean	3.63	2.01	15.88	13.38	16.13	10.13	40.25	40.63	20.88	11.00
	s.d.	3.57	2.28	6.85	8.10	11.24	8.68	11.27	9.91	10.55	6.01
	Range	0.0–12.0	0.0–8.0	4.0–32.0	0.0–32.0	0.0–52.0	0.0–32.0	16.0–60.0	24.0–64.0	4.0–44.0	0.0–24.0
Middle sound (vowels)	Mean	5.66	3.64	17.38	12.63	14.13	10.75	29.25	27.13	9.00	5.38
	s.d.	4.22	4.23	7.84	6.51	9.14	5.32	12.25	9.57	6.26	4.01
	Range	0.0–16.7	0.0–12.0	4.0–36.0	0.0–28.0	0.0–40.0	0.0–24.0	0.0–52.0	8.0–48.0	0.0–20.0	0.0–16.0
Final sound (consonants)	Mean	11.30	5.50	25.25	16.63	23.38	17.13	49.13	44.63	30.50	25.63
	s.d.	7.26	4.40	10.18	8.01	10.42	8.51	11.95	9.59	12.64	14.15
	Range	0.0–28.0	0.0–16.0	4.0–48.0	4.0–32.0	4.0–56.0	4.0–44.0	28.0–76.0	32.0–64.0	4.0–64.0	8.0–60.0

B. Phoneme errors

Table I shows the percent of errors (relative to the total possible errors) for each TTS voice. Figure 1 graphically shows the percentage of phoneme errors in initial, medial, and final positions of words under two SNR conditions. In MRT, each word consists of an initial consonant, a middle

vowel, and a final consonant or consonant cluster.² The percent-error measure, therefore, not only provides an overall estimation of the difficulty listeners experienced with the four TTS products but also sheds light on how well consonants and vowels and the consonants in the initial and final positions are modeled in the four different TTS synthesis

TABLE II. Percentages of single consonant substitution and omission errors in four TTS products and the human voice in word-initial and word- and utterance-final positions. Total occurrence of each phoneme (correct identification + single phoneme substitution/omission errors for 32 participants) is shown in parentheses. /ng/ (*sing*); /ch/ (*chew*); /th/ (*think*).

	AT&T		Festival		FlexVoice		IBM		Human	
	<i>Initial</i>	<i>Final</i>	<i>Initial</i>	<i>Final</i>	<i>Initial</i>	<i>Final</i>	<i>Initial</i>	<i>Final</i>	<i>Initial</i>	<i>Final</i>
/b/	5.52 (163)	68.97 (29)	18.40 (163)	25.00 (32)	61.84 (152)	67.86 (28)	32.72 (162)	42.86 (28)	2.08 (192)	0.00 (32)
/d/	20.65 (92)	28.41 (88)	16.67 (102)	13.98 (93)	26.92 (104)	49.45 (91)	2.94 (102)	46.59 (88)	3.13 (128)	3.19 (94)
/g/	5.41 (37)	19.61 (51)	0.00 (40)	6.25 (64)	21.05 (38)	60.78 (51)	15.00 (40)	21.88 (64)	0.00 (64)	3.13 (64)
/p/	20.79 (202)	19.33 (119)	18.93 (206)	17.83 (129)	94.57 (184)	86.27 (102)	4.33 (208)	29.66 (118)	1.79 (224)	15.63 (128)
/t/	16.90 (142)	1.69 (177)	4.23 (142)	15.38 (169)	56.52 (92)	22.98 (161)	4.17 (144)	21.64 (171)	0.78 (128)	10.63 (160)
/k/	23.78 (143)	4.92 (183)	2.01 (149)	20.74 (188)	54.41 (136)	22.40 (183)	5.96 (151)	34.41 (186)	0.00 (128)	1.25 (160)
/s/	9.95 (221)	3.33 (60)	18.30 (235)	17.24 (58)	0.00 (236)	19.61 (51)	6.25 (240)	0.00 (59)	0.00 (224)	0.00 (64)
/f/	4.44 (90)	4.17 (24)	26.51 (83)	20.83 (24)	15.12 (86)	68.75 (16)	54.17 (96)	56.52 (23)	1.04 (96)	0.00 (32)
/h/	5.49 (91)	(0)	16.67 (96)	(0)	26.19 (84)	(0)	25.00 (96)	(0)	12.50 (96)	(0)
/w/	0.00 (50)	(0)	0.00 (48)	(0)	22.92 (48)	(0)	45.65 (46)	(0)	1.56 (64)	(0)
/r/	0.00 (95)	0.00 (16)	1.92 (104)	35.71 (14)	4.30 (93)	26.67 (15)	17.02 (94)	6.25 (16)	0.00 (63)	(0)
/l/	1.43 (70)	1.78 (169)	2.82 (71)	11.98 (167)	8.33 (72)	16.23 (154)	1.37 (73)	0.58 (173)	0.00 (64)	0.00 (192)
/m/	6.25 (64)	32.86 (70)	0.00 (63)	22.54 (71)	14.06 (64)	65.63 (64)	25.76 (66)	49.28 (69)	3.13 (64)	11.46 (96)
/n/	0.00 (32)	24.85 (169)	0.00 (32)	6.38 (188)	6.06 (33)	32.16 (171)	6.25 (32)	27.42 (186)	(0)	16.96 (224)
/ng/	(0)	52.00 (50)	(0)	62.96 (54)	(0)	55.17 (58)	(0)	54.24 (59)	(0)	0.00 (32)
/v/	(0)	43.48 (23)	(0)	39.13 (23)	(0)	80.00 (15)	(0)	23.08 (26)	(0)	0.00 (32)
/ch/	(0)	0.00 (16)	(0)	0.00 (16)	(0)	31.25 (16)	(0)	12.50 (16)	(0)	(0)
/th/	(0)	37.50 (16)	(0)	31.25 (16)	(0)	33.33 (15)	(0)	46.67 (15)	(0)	25.00 (32)
/z/	(0)	40.91 (22)	(0)	0.00 (24)	(0)	19.05 (21)	(0)	26.09 (23)	(0)	0.00 (32)

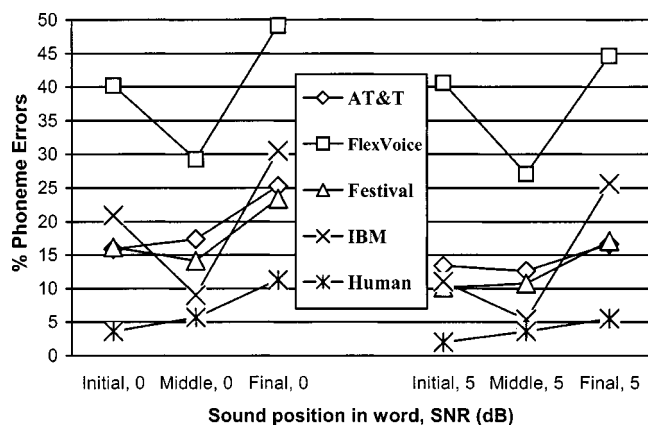


FIG. 1. Mean percent errors in initial, middle, and final positions of words at 0-dB and 5-dB signal-to-noise ratios for the four TTS products and the human voice. Initial and final sounds are consonants and the middle sounds are vowels.

algorithms. Table II shows the percentage of single-phoneme substitution and omission errors for selected consonants in the initial and final positions of words. A three-factor repeated-measures ANOVA (4 TTS products \times 2 SNR conditions \times 3 sound positions in words) was performed with percent error as the dependent variable. The results of the ANOVA along with *post hoc* Bonferroni tests to further explore the data are presented below.

The TTS products produced significantly different error rates ($F [2.44, 75.57] = 168.38$, Greenhouse–Geisser $p < 0.0001$). The *post hoc* Bonferroni test for six possible comparisons among the four TTS products required an $\alpha = 0.05/6 = 0.008$. Results showed that AT&T NextGen ($M = 16.85$) did not significantly differ from IBM ViaVoice ($M = 17.06$, $t [31] = 0.196$, 2-tailed $p = 0.85$) and Festival ($M = 15.27$, $t [31] = 1.32$, 2-tailed $p = 0.20$) and IBM ViaVoice was equivalent to Festival ($t [31] = 1.32$, 2-tailed $p = 0.10$) in error rates. However, FlexVoice ($M = 38.50$) had significantly higher error rate than AT&T NextGen ($t [31] = -18.53$, 2-tailed $p < 0.0001$), Festival ($t [31] = -17.77$, 2-tailed $p < 0.0001$), and IBM ViaVoice ($t [31] = -21.202$, 2-tailed $p < 0.0001$). The hybrid technology used in FlexVoice produced a significantly greater number of phoneme errors under adverse SNRs than either concatenative or formant synthesis alone. The two concatenative techniques and the one formant technique did not differ in phoneme error rates.

The 0-dB SNR condition ($M = 24.26$) had significantly higher percentage of errors than the 5-dB SNR condition ($M = 19.58$, $F [1, 31] = 34.96$, $p < 0.0001$). The ANOVA for percent errors also revealed an SNR \times sound positions interaction ($F [2, 62] = 3.46$, $p < 0.04$). *Post hoc* t-tests comparing sound positions across 0- and 5-dB SNR conditions with Bonferroni α set at $0.05/6 = 0.008$ were performed to explore the nature of interaction between SNR and sound positions. The listeners experienced equivalent difficulty (percent errors) at 0 dB, initial sound position ($M = 23.28$) and 5 dB, final position ($M = 26.00$, $t [31] = -2.30$, 2-tailed $p = 0.03$), and at 0 dB, middle sound position ($M = 17.44$) and 5 dB, initial position ($M = 18.78$, $t [31] = -1.33$, 2-tailed $p = 0.19$). Listeners, however, experienced more difficulty in 5

dB, final position than at 0 dB, middle position ($t [31] = 7.09$, 2-tailed $p < 0.0001$), 0 dB, initial position than 5 dB, middle position ($M = 13.97$, $t [31] = 13.43$, 2-tailed $p < 0.0001$), 0-dB final position ($M = 32.06$) than 5-dB initial position ($t [31] = 11.56$, 2-tailed $p < 0.0001$), and 0 dB, final position than 5 dB, middle position ($t [31] = 16.26$, 2-tailed $p < 0.0001$) conditions.

The preceding analysis shows that noise affects intelligibility to a greater degree in word-final position than in word-initial position and in word-initial position than in word-middle position. It is, however, interesting to note that noise affected all TTS products equally as indicated by a lack of interaction between TTS products and SNR conditions ($F [3, 93] = 1.65$, $p > 0.1$). Because a similar nonsignificant interaction between TTS products and SNR was found earlier for whole words correct, there is a strong possibility that all four TTS synthesis algorithms suffer equally in the presence of noise.

The three sound positions within MRT words, initial ($M = 21.03$), middle ($M = 15.70$), and final ($M = 29.03$), differed significantly in percent errors ($F [2, 62] = 202.91$, $p < 0.0001$). *Post hoc* Bonferroni tests (with α set at $0.05/3 = 0.02$) showed that there were significantly more errors in the final sounds of words ($M = 29.03\%$) than in the initial ($M = 21.03\%$, $t [31] = -10.84$, 2-tailed $p < 0.0001$) and middle ($M = 15.70\%$, $t [31] = -19.23$, 2-tailed $p < 0.0001$) sounds, and the initial sounds accounted for significantly more errors than the middle sounds ($t [31] = 9.63$, 2-tailed $p < 0.0001$). Since, in MRT words, the middle sounds are vowels, the TTS products were better at modeling vowels than consonants. In addition, overall, listeners recognized initial consonants of TTS words better than the final consonants. It should be noted that, in the present study, the final consonants of MRT words were also in the utterance-final position.

In order to determine if the differential error rates for sound positions in words were present in all TTS products under investigation, 21 paired t-tests comparing the three sound positions within each TTS product as well as analogous sound positions across the four TTS products were performed with Bonferroni α set at $0.05/21 = 0.002$. AT&T NextGen initial sounds ($M = 14.62$) did not differ significantly from AT&T NextGen middle sounds ($M = 15.00\%$, $t [31] = -0.35$, 2-tailed $p > 0.7$). Similarly, Festival initial sounds ($M = 13.13$) did not differ significantly from Festival middle sounds ($M = 12.44\%$, $t [31] = 0.51$, 2-tailed $p > 0.6$). The two concatenative synthesis methods modeled the initial consonants and middle vowels equally well. In contrast, IBM ViaVoice initial sounds ($M = 15.93$) had significantly worse recognition than its middle sounds ($M = 7.19$, $t [31] = 7.17$, 2-tailed $p < 0.0001$). FlexVoice initial sounds ($M = 40.43$), likewise, had significantly more errors than its middle sounds ($M = 28.18$, $t [31] = 5.59$, 2-tailed $p < 0.0001$). Formant and hybrid TTS synthesis techniques modeled initial consonants of words more poorly than the middle vowels.

Final consonants of AT&T NextGen ($M = 20.94$), FlexVoice ($M = 46.88$), Festival ($M = 20.25$), and IBM ViaVoice ($M = 28.06$) had more errors than their initial consonants.

The differences between initial and final sound positions in words for AT&T NextGen ($t[31] = -5.31$, 2-tailed $p < 0.0001$), FlexVoice ($t[31] = 6.15$, 2-tailed $p < 0.0001$), Festival ($t[31] = 6.27$, 2-tailed $p < 0.0001$), and IBM ViaVoice ($t[31] = 4.24$, 2-tailed $p < 0.0001$) were all significant. In all TTS products, the initial consonant sounds of words were more intelligible than the final consonant sounds.

The differences between the middle vowels and final consonants for AT&T NextGen ($t[31] = -4.10$, 2-tailed $p < 0.0001$), FlexVoice ($t[31] = -8.51$, 2-tailed $p < 0.0001$), Festival ($t[31] = -5.54$, 2-tailed $p < 0.0001$), and IBM ViaVoice ($t[31] = -9.99$, 2-tailed $p < 0.0001$) were also significant, indicating that middle vowels were more intelligible than final consonants in all TTS products.

The differences between initial sounds of AT&T NextGen and Festival ($t[31] = 0.91$, 2-tailed $p > 0.3$), AT&T NextGen and IBM ViaVoice ($t[31] = 1.05$, 2-tailed $p > 0.3$), and IBM ViaVoice and Festival ($t[31] = 1.05$, 2-tailed $p > 0.5$) did not reach statistical significance. The FlexVoice initial consonants, however, had significantly more errors than AT&T NextGen ($t[31] = -13.00$, 2-tailed $p < 0.0001$), Festival ($t[31] = 12.36$, 2-tailed $p < 0.0001$), and IBM ViaVoice ($t[31] = -10.26$, 2-tailed $p < 0.0001$). Concatenative and “straight” parametric synthesis techniques modeled initial consonants equally well, whereas the technique that combined the two as implemented in FlexVoice was more prone to errors.

The difference between AT&T NextGen middle and Festival middle sounds for percent errors was nonsignificant ($t[31] = 1.78$, 2-tailed $p > 0.8$). The vowel errors (middle sounds) for AT&T NextGen ($t[31] = -7.24$, 2-tailed $p < 0.0001$) and Festival ($t[31] = -4.77$, 2-tailed $p < 0.0001$) were significantly higher than for IBM ViaVoice. This suggests that the parametric synthesis technique is better at synthesizing vowel sounds than concatenative techniques. Consistent with other results presented above, the FlexVoice vowels had significantly more errors than AT&T NextGen ($t[31] = -7.20$, 2-tailed $p < 0.0001$), Festival ($t[31] = 8.06$, 2-tailed $p < 0.0001$), and IBM ViaVoice ($t[31] = -13.11$, 2-tailed $p < 0.0001$). In this study, FlexVoice’s hybrid technology performed more poorly on vowel sounds than either of the two techniques alone.

The differences between AT&T NextGen and Festival final consonants ($t[31] = 0.51$, 2-tailed $p > 0.6$), AT&T NextGen and IBM ViaVoice final consonants ($t[31] = 2.34$, 2-tailed $p > 0.02$) and Festival and IBM ViaVoice final consonants ($t[31] = 2.91$, 2-tailed $p = 0.007$) did not reach statistical significance. Unlike for vowels, the concatenative and formant synthesis techniques did not differ in how well they modeled either initial or final consonants. However, both techniques produced more errors for final consonants than initial consonants. The FlexVoice final consonants, however, had significantly more errors than AT&T NextGen ($t[31] = -10.43$, 2-tailed $p < 0.0001$), Festival ($t[31] = 13.25$, 2-tailed $p < 0.0001$), and IBM ViaVoice ($t[31] = -14.13$, 2-tailed $p < 0.0001$). Taken together, the findings for FlexVoice indicate that its synthesis technology produces significantly more errors in each of the three sound positions than the other three TTS products investigated.

The ANOVA for percent errors showed significant interaction among TTS products \times SNR \times sound positions ($F[6,186] = 2.49$, $p < 0.03$). To explore the nature of this three-way interaction, 33 correlated t-tests comparing analogous pairs of conditions at each SNR for the four TTS products were performed with Bonferroni α set at $0.05/33 = 0.002$. The means for these conditions are shown in Table I. Consistent with previously presented data, FlexVoice was found to be significantly less intelligible than the other three TTS products across all comparisons (2-tailed $p < 0.002$). IBM ViaVoice vowels were significantly more intelligible than AT&T NextGen vowels under both 0-dB and 5-dB SNRs, but were significantly more intelligible than Festival vowels only under 5-dB SNR condition (2-tailed $p < 0.002$). AT&T NextGen and Festival did not differ in vowel errors. Although there is no ready explanation for the inconsistency of finding more vowel errors for Festival compared to IBM at 5-dB SNR but not at 0-dB SNR, overall, it appears that formant TTS is superior in modeling vowels than concatenative TTS. AT&T NextGen, Festival, and IBM ViaVoice did not differ significantly in the number of initial or final consonant errors. The formant synthesis (IBM ViaVoice), LPC-based (Festival), and waveform-based (AT&T NextGen) concatenative TTS techniques were comparable in the intelligibility of initial and final consonants of words.

C. Human voice

How do the current TTS products compare with the human voice under similar listening conditions? Table I shows percent whole words correct (no errors in these words) and percent phoneme errors for the human voice and the four TTS products. First, to evaluate listeners’ performance on human voice, a two-factor (2 SNR conditions \times 3 sound positions in words) repeated measures ANOVA was performed. The dependent variable was percent errors relative to total possible errors in each condition. The results showed that the human voice was significantly less intelligible at 0-dB SNR ($M = 6.85$) than at 5-dB SNR ($M = 3.72$, $F[1,31] = 18.14$, $p < 0.0001$), and some or all of the three sound positions in words were also significantly different ($F[2,62] = 40.65$, $p < 0.0001$). The interaction between SNR and sound positions was significant ($F[1.38,42.85] = 4.17$, Greenhouse–Geisser $p < 0.04$).

Post hoc comparisons of three unique pairs of sound positions with Bonferroni α set at $0.05/3 = 0.02$ showed that the initial consonants ($M = 2.81$) were significantly more intelligible than middle vowels ($M = 4.63$, $t[31] = -3.39$, 2-tailed $p = 0.002$) and final consonants ($M = 8.38$, $t[31] = -8.37$, 2-tailed $p = 0.0001$), and the middle vowels had significantly fewer errors than final consonants ($t[31] = -5.54$, 2-tailed $p = 0.0001$). The difference between initial and final consonants for human voice is in agreement with that found for TTS products. Consonants in utterance-final positions are difficult to recognize in both human speech and TTS. It may be recalled, however, that both AT&T NextGen and Festival did not differ in error rates between initial consonants and middle vowels, and IBM ViaVoice and FlexVoice had more errors in initial consonants than in

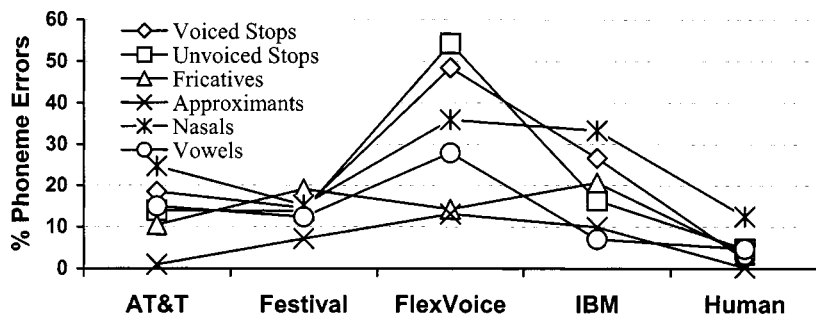


FIG. 2. Pattern of phoneme errors in four TTS products and the human voice. Only single phoneme substitution errors are included. Voiced stops: b, d, g; Unvoiced stops: p, t, k; Fricatives: s, f, h, v, th (think), z; Approximants: w, r, l; Nasals: m, n, ng (sing).

middle vowels. For human speech, however, middle vowels were less intelligible than initial consonants.

To explore the nature of interaction between SNR and sound positions in human voice, *post hoc* multiple comparisons were made with Bonferroni tests ($\alpha=0.05/9=0.006$) among the three sound positions within each SNR as well as analogous sound positions across the two SNR conditions. The 0-dB, initial ($M=3.63$)—0 dB, middle ($M=5.66$, $t[31]=-3.01$, 2-tailed $p=0.002$) pair, 0 dB, initial—0 dB, final ($M=11.30$, $t[31]=-5.43$, 2-tailed $p<0.0001$) pair, and 0 dB, middle—0 dB, final ($t[31]=-4.53$, 2-tailed $p<0.0001$) pair parallel the significant differences found for the three sound positions earlier. In contrast, only the difference between 5 dB, initial ($M=2.01$)—5 dB, final ($M=5.50$, $t[31]=-4.91$, 2-tailed $p<0.0001$) pair reached significance. The 5-dB, initial—5 dB, middle ($M=3.66$, $t[31]=-2.21$, 2-tailed $p>0.03$) conditions and the 5-dB, middle—5 dB, final ($t[31]=-1.73$, 2-tailed $p>0.09$) conditions were not significantly different. The differences between the two analogous sound positions across SNR, 0 dB, initial—5 dB, initial ($t[31]=2.34$, 2-tailed $p>0.02$) and 0

dB, middle—5 dB, middle ($t[31]=1.97$, 2-tailed $p>0.06$) pairs, were not statistically significant. The 0-dB final—5-dB final ($t[31]=3.56$, 2-tailed $p=0.001$) conditions were significantly different. Taken together, these results suggest that the differences in error rates among the three positions were, in part, a function of SNR. When noise was as strong as the signal, significantly more errors occurred on final consonants than on middle vowels, and on middle vowels than on initial consonants of human voice. However, at a more favorable SNR of 5 dB, the effect of noise on differences in consonant and vowel errors of human voice tended to disappear except in utterance-final position.

A t-test for related samples was performed comparing the two SNR conditions of human voice for percentage of whole words correct. Consistent with results for percent errors for human voice and the results obtained for TTS products reported above, the human voice was significantly more intelligible at 5-dB SNR ($M=90.62$) than at 0-dB SNR ($M=83.50$, $t[31]=-4.19$, 2-tailed $p<0.0001$).

Because the data for human voice were collected from a

TABLE III. Patterns of single phoneme substitution errors for selected consonants as a percentage of the total number of errors for each consonant rounded off to the nearest whole number. Only substitutions that accounted for a minimum of 10% of the errors for a given phoneme are shown. /ng/ (sing); /th/ (think).

	AT&T NextGen	Festival	FlexVoice	IBM ViaVoice	Human voice
/b/	/p/, 52; /f/, 14	/p/, 16; /f/, 34; /v/, 16	/f/, (59); /v/, (19)	/d/, 35; /p/, 28; /f/, 14	/f/, 25; /r/, 25; /m/, 25
/d/	/g/, 11; /t/, 59; /k/, 14	/g/, 17; /t/, 13	/g/, 10; /t/, 26; /n/, 10	/t/, 100	/b/, 71
/g/	/k/, 75; /ng/, 17	/k/, 100	/d/, 10; /k/, 56	/d/, 15; /k/, 45; /h/, 15	/b/, 50; /v/, 50
/p/	/b/, 69	/b/, 53; /t/, 13; /f/, 10; /v/, 13	/b/, 16; /f/, 53	/b/, 20; /d/, 18; /t/, 45	/k/, 63
/t/	/d/, 63; /h/, 15; /th/, 22	/d/, 59	/d/, 54	/d/, 67; /h/, 12	/p/, 33; /k/, 44
/k/	/g/, 72; /t/, 14	/d/, 14; /g/, 50; /t/, 12	/d/, 12; /g/, 57; /t/, 16	/d/, 10; /g/, 18; /p/, 30; /t/, 16	/p/, 100
/s/	/h/, 17; /z/, 42	/z/, 74	/z/, 100	/h/, 13; /z/, 13	No errors
/f/	/b/, 20; /p/, 40; /t/, 20; /s/, 29; /h/, 20	/b/, 41; /p/, 19; /v/, 26	/th/, 38	/b/, 42; /h/, 22	/k/, 100
/h/	/n/, 20	/w/, 13; /n/, 25	/f/, 59; /th/, 14	/k/, 83	No errors
/v/	/t/, 50; /f/, 30	/p/, 11; /r/, 11; /l/, 33; /m/, 22; /n/, 11	/d/, 17; /t/, 58; /z/, 25	/d/, 33; /f/, 17; /z/, 17	No errors
/th/	/t/, 100	/t/, 40; /v/, 20	/t/, 20; /s/, 40; /f/, 20; /v/, 20	/d/, 29; /t/, 71	/p/, 13; /t/, 25; /k/, 50
/z/	/s/, 100	No errors	/s/, 100	/s/, 100	No errors
/w/	No errors	No errors	/r/, 73; /m/, 18	/r/, 14; /l/, 57; /m/, 29	/r/, 100
/r/	No errors	/b/, 29; /p/, 14; /k/, 14; /l/, 14	/b/, 13; /v/, 13; /w/, 25	/l/, 24; /m/, 47	No errors
/l/	/w/, 25; /n/, 25	/v/, 14	/d/, 10; /n/, 23	No errors	No errors
/m/	/n/, 44; /ng/, 41	/n/, 81	/n/, 27; /ng/, 20	/b/, 18; /d/, 16; /n/, 55	/n/, 77
/n/	/d/, 10; /t/, 31; /m/, 17	/b/, 17; /t/, 25; /m/, 50	/d/, 14; /m/, 21	/d/, 58; /m/, 21	/d/, 11; /m/, 82
/ng/	/g/, 15; /t/, 15; /k/, 12; /n/, 46	/n/, 74	/l/, 13; /m/, 31; /n/, 47	/n/, 72	/g/, 13

different group of 32 listeners from those for TTS products, a series of four t-tests for independent samples comparing the number of whole words correct in human voice with each of the TTS voices was performed. The results revealed that a significantly higher number of whole words were correct for the human voice ($M = 86.81$) than for AT&T NextGen ($M = 66.25$, $t(52.59) = 12.25$, 2-tailed $p < 0.0001$), Festival ($M = 68.19$, $t(43.42) = 8.59$, 2-tailed $p < 0.0001$), FlexVoice ($M = 33.00$, $t(46.88) = 27.73$, 2-tailed $p < 0.0001$), and IBM ViaVoice ($M = 58.81$, $t(52.34) = 16.58$, 2-tailed $p < 0.0001$). In spite of advances in synthesis technology, TTS poses a significant challenge to listeners compared to natural speech.

The human voice was also compared to each of the TTS products separately with an independent samples t-test using percent errors as dependent variable. The human voice ($M = 5.29$) had significantly fewer phoneme errors than AT&T NextGen ($M = 16.85$, $t(62) = -11.28$, 2-tailed $p < 0.0001$), Festival ($M = 15.27$, $t(62) = -9.53$, 2-tailed $p < 0.0001$), FlexVoice ($M = 38.50$, $t(62) = -27.56$, 2-tailed $p < 0.0001$), and IBM ViaVoice ($M = 17.06$, $t(62) = -16.03$, 2-tailed $p < 0.0001$), supporting the findings presented above for whole words correct.

D. Error patterns

Figure 2 and Table III present data on error rates for selected groups of consonants and selected consonants, respectively. Only single phoneme substitution and omission errors are included in this analysis, excluding those errors that involved substitution of more than one sound (e.g., the final consonant of *wed* heard as *west*) and addition of sounds. Consonant clusters in words (e.g., *ld* in *sold*) are also excluded from this analysis. The error data reported in the previous sections and in Table I include all errors.

Figure 2 shows that there were a relatively large number of errors on nasal consonants in all TTS products and in human voice ranging from 12.5% for human voice to 35.9% for FlexVoice. Somewhat surprisingly, formant coding method used in IBM ViaVoice produced more nasal consonant errors than the LPC coding of Festival. Typically, a separate filter is used in formant synthesizers to simulate the zeros of nasal resonance, whereas the all-pole LPC synthesis does not adequately account for antiresonances of nasal sounds. Nasal consonants as a group also have the largest number of errors among consonants in AT&T NextGen.

Table III shows the pattern of substitution errors for nasal and other consonants. Although a large number of nasal errors involved confusing one nasal sound for another nasal sound, for IBM ViaVoice, 58% of /n/ errors involved /d/ substitution and 33% of /m/ errors involved either /b/ or /d/ substitution. Similarly, listeners frequently mistook AT&T NextGen /n/ for /t/ (31%)³ and /ng/ for either /g/ (15%) or /t/ (15%). In Festival, /n/ was confused for /b/ (17%) and for /t/ (25%). FlexVoice /m/ and /n/ errors were widely scattered among many sounds with just 47% and 49%, respectively, being another nasal sound. All human voice errors on /m/ and /n/ were the result of mutual confusion and there was a single /g/ for /ng/ substitution. Although nasal sound errors in TTS products involve multiple feature confusion—manner (oral/nasal), place (labial, alveolar and velar), and voicing

(voiced/unvoiced), attempts to better represent nasal resonances may improve the intelligibility of nasal sounds in both formant and concatenative synthesis methods.

For stop sounds, the TTS voices had a relatively large number of errors, ranging from 14.57% to 48.49% for voiced stops and 13.84% to 54.31% for unvoiced stops as compared to the human voice (2.26% and 4.74%, respectively). When listening to TTS voices, manner confusion is less common (typically, a stop is substituted for another stop) although both place of production and voicing confusions occur within stops (see Table III). Voicing errors were the most common in TTS products mistaking voiced stops for unvoiced stops and vice versa. Human stop consonants accounted for a relatively small number of errors and the errors, with the exception of those associated with the /p/ sound, were in place of production rather than voicing. Because place and manner of production errors were less frequent than voicing errors in TTS stops, the intelligibility of these sounds may improve with better modeling of voicing feature.

Fricative errors among TTS products ranged from 10.48% for AT&T NextGen to 20.71% for IBM ViaVoice as compared to 3.45% for the human voice. Among the TTS products, FlexVoice had the second smallest number of errors on fricatives, while in all other groups of sounds, it had the largest number of errors. In all TTS products, /z/ was frequently confused for /s/ and vice versa, indicating that voicing feature was not adequately modeled in these two sounds. For /f/, /v/, and /th/, the confusion frequently revolved around the manner of production and to a lesser degree, voicing confusion. These fricatives were frequently mistaken for a stop sound, oftentimes with the added confusion of voicing.

Approximants had the fewest number of errors among the five groups of consonants (see Fig. 2). AT&T NextGen had no errors on /w/ and /r/ sounds and had four errors on the /l/ sound. Festival had no errors on the /w/ sound, while human voice had no errors on the /r/ and /l/ sounds. There is no discernible pattern in the small number of errors found in approximants.

Vowel errors for /i/, /ɪ/, /e/, /ɛ/, /æ/, /u/, /o/, /ɑ/, /ʌ/ and /aɪ/ were tabulated. The formant coding technique used in IBM ViaVoice produced the smallest number of vowel errors at a mean of 7.06% with two of the ten vowels, /o/ and /aɪ/, registering zero errors. Festival had the second lowest vowel errors among the TTS products ($M = 12.4\%$) with 25% of /æ/, 20% of /ɑ/, and 22% of /aɪ/ being mistaken for another sound. The overall vowel error rate in AT&T NextGen was 15% with 30% of /ɪ/, 21% of /ɛ/, and 23% of /ɑ/ confused for another sound. FlexVoice vowel errors averaged 27.9% with 19% or more errors in all but two (/i/ and /ʌ/) of the ten vowels. The mean vowel errors in human voice stood at 4.72% with only the diphthong /aɪ/ mistaken for another vowel sound 34% of the times.

IV. DISCUSSION AND CONCLUSIONS

The results of this study are not directly comparable to those of Pisoni and his colleagues (Green, Logan, and Pisoni, 1986; Logan, Greene, and Pisoni, 1989; Pisoni, Nusbaum,

and Green, 1985) because the present study tested different synthesis products and under adverse SNR conditions. However, the patterns of performance of TTS products in these studies are somewhat similar in that there were significantly more errors with TTS voices than with the human voice, and that the final consonant errors were the most numerous with both TTS and human voices. It is noteworthy that the three better TTS products in the present study—AT&T NextGen, Festival, and IBM ViaVoice with total error rates ranging from 15.27% to 17.06%—did slightly worse in noise than the best TTS products of the past [DECtalk 1.8, Paul with 12.92% errors in Logan, Greene, and Pisoni (1989) study done in quiet] although granted that the present results were obtained under challenging listening conditions. This study does not support the industry claim (included in their promotional literature and web sites) that the current TTS products are adequate for high noise environments typically prevalent in automobiles, airports, and in many offices and classrooms. Listening to unrestricted TTS output—e-mail messages, web content, news reports, lectures, and so on—under these adverse listening conditions will likely require a disproportionate amount of cognitive resources of the listener impairing other ongoing activities.

It must, however, be pointed out that, in addition to the adverse SNRs, the listening materials in the present study consisting of monosyllabic words contained very little linguistic and communicative contextual cues that would normally be present in real-life listening conditions. In actual listening environments, the TTS products investigated may be expected to perform better due to top-down processing relying on the listener's knowledge of the topic, participants, and previous utterances. However, it remains to be seen if rich language and contextual clues can partially bridge the gap in intelligibility between natural speech and TTS.

The present study found evidence that formant coding may be slightly superior to concatenative techniques in synthesizing vowels. The study also found that the three major synthesis techniques, LPC-based diphone and waveform-based concatenative synthesis techniques and allophone-based formant coding technique, do not differ in error rates for initial and final consonants. Although not statistically significant, the two concatenative techniques had fewer consonant errors than the formant technique. In addition, the two concatenative techniques produced significantly more whole words correct than the formant technique. These findings suggest that the consonants of concatenative techniques may be slightly more intelligible than the formant technique (see Fig. 1).

In the present study, no TTS algorithm emerged as a clear better choice for synthesis of segments. Although the current efforts in TTS research appear to be largely aimed at improving naturalness of speech, the results of the present study indicate that much additional work is needed in the area of synthesis of phonemes. The three better TTS products in the present study (AT&T NextGen, Festival, and IBM ViaVoice) were, on the average, 22% less intelligible on the measure of whole words correct and had 3 times more segmental errors than the human voice under similar listening conditions (see Table I).

SNR is an important factor in the intelligibility of TTS output. A surprising finding of the present study is that noise affected all TTS products equally. No TTS technique was better suited to resist noise interference than others. In addition, the TTS output was more susceptible to noise interference than human voice. TTS output was consistently more intelligible at 5-dB SNR than at 0-dB SNR in all three positions of words. For human voice, the 0-dB SNR was more detrimental to intelligibility than 5-dB SNR for the utterance-final consonant sounds only.

The hybrid speech technology that combines the elements of concatenative and formant synthesis techniques as implemented in FlexVoice did not fare well in this study.⁴ The details of this technique are proprietary and are not available. The concept is attractive because it has the potential to combine the computational efficiency (during synthesis) and the choice of a range of different voices obtainable in formant coding with short and less-expensive developmental cycle and naturalness of speech found in concatenative synthesis. However, FlexVoice may be illustrative of the difficulty faced in combining these two techniques.

Error patterns are discernible for four of the five groups of sounds—stops, fricatives, nasals, and vowels. Intelligibility of TTS products would probably improve with better modeling of nasal resonances in nasal sounds, voicing in both stops and fricatives, and by enhancing frication feature in fricatives. In addition, concatenative and hybrid synthesis methods would benefit from better modeling of vowels.

A. Future directions

Although the four synthesis products investigated in the present study differ in linguistic preprocessing in addition to the synthesis techniques, it is unlikely that the pronunciation dictionaries from which the phonetic forms of common monosyllabic words used in this study were drawn would differ significantly among the four TTS products. The differences in performance are, therefore, mostly attributable to units of speech selected for synthesis (phonemes, half-phones, or diphones) and the method of generating speech waveform (formant, LPC, or harmonic-plus-noise waveform coding). Therefore, a straightforward explanation for the poorer intelligibility of TTS monosyllables may be that there are important, as yet not completely understood, differences in the spectral and temporal composition of synthesized words as compared to natural speech. In spite of advances in acoustic modeling of speech, especially as they relate to coarticulation, the TTS products investigated in the present study may be deficient in accurately modeling acoustic properties known to be important for perception of speech. A comparative acoustic analysis of TTS and natural speech words may shed light on this possibility.

There is, however, another intriguing possibility to account for large differences between TTS and human speech. Several different lines of evidence suggest that human listeners process speech differently from nonspeech auditory signals (see Jusczyk and Luce, 2002 for a review). For instance, listeners are capable of processing 400 words per minute (Orr, Friedman, and Williams, 1965), whereas they are unable to assign correct temporal sequence to more than four

different nonspeech sounds per second (Warren, 1974). It has also been shown that listeners process sine wave-synthesized speech differently depending on whether they are told that it consisted of a series of beeps and tones or synthesized speech (Remez *et al.*, 1981). These and related findings have led to the formulation of theories that emphasize the importance of underlying articulatory gestures over surface acoustic features in speech perception (Fowler, 1994; Liberman and Mattingly, 1989). It is possible that listeners may initially use a nonspeech auditory processing mode to analyze TTS output because the output is perceived as originating from a nonhuman source, switching to the speech mode only in later stages resulting in poor speech perception. While segmental and suprasegmental cues for identification of units of speech may be mostly encoded in the lower frequency regions of speech, the “humanness” cues that put the auditory system in the speech mode may be partly located in the higher frequency regions, especially in the higher formants. Additional research is needed to explore these and other explanations for the poor intelligibility of current TTS products.

ACKNOWLEDGMENT

Kelly Schmitt ably served as a research assistant for this study. The study could not have been completed without her diligent efforts.

¹In this paper, Greenhouse–Geisser *p* is used when the data failed the Mauchly sphericity test (Cohen, 1996).

²All lists have two words without final consonants and list D has one word without the initial consonant.

³All percentages referring to phoneme confusion are computed from total single phoneme substitution errors on that phoneme. Thus, in this instance, 31% of single phoneme substitution errors on the /n/ sound were attributable to /t/ sound.

⁴Even as this study was underway using FlexVoice 2, a new version, FlexVoice 3, was released.

Allen, J., Hunnicutt, S., and Klatt, D. (1987). *From Text To Speech, The MITTALK System* (Cambridge University Press, New York).

ANSI (1969). ANSI S3.6-1969, “American National Standards Specification for Audiometers” (American National Standards Institute, New York).

Atal, B. S., and Hanauer, S. L. (1971). “Speech analysis and synthesis by linear prediction of the speech wave,” *J. Acoust. Soc. Am.* **50**, 637–655.

Cohen, B. H. (1996). *Explaining Psychological Statistics* (Brooks/Cole, Pacific Grove, CA).

Dutoit, T., and Leich, H. (1993). “MBR-PSOLA: Text-to-speech synthesis based on an MBE re-synthesis of the segments database,” *Speech Commun.* **13**, 435–440.

Fowler, C. A. (1994). “Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation,” *Percept. Psychophys.* **55**, 597–610.

Fucci, D., Reynolds, M. E., Bettagere, R., and Gonzales, M. D. (1995). “Synthetic speech intelligibility under several experimental conditions,” *Augment. Alt. Commun.* **11**, 113–117.

Green, B. G., Logan, J. S., and Pisoni, D. B. (1986). “Perception of synthetic speech produced automatically by rule: Intelligibility of eight text-to-speech systems,” *Behav. Res. Methods Instrum. Comput.* **18**, 100–107.

House, A. S., Williams, C. E., Hecker, M. H. L., and Kryter, K. D. (1965). “Articulation-testing methods: Consonantal differentiation with a closed-response set,” *J. Acoust. Soc. Am.* **37**, 158–166.

Juszyk, P. W., and Luce, P. A. (2002). “Speech perception,” in H. Pashler & S. Yantis, *Stevens' Handbook of Experimental Psychology*, 3rd ed., edited by H. Pashler and S. Yantis (Wiley, New York), Vol. 1, pp. 493–536.

Kalikow, D. N., Stevens, K. N., and Elliott, L. L. (1977). “Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability,” *J. Acoust. Soc. Am.* **61**, 1337–1351.

Klatt, D. H. (1987). “Review of text-to-speech conversion of English,” *J. Acoust. Soc. Am.* **82**, 737–793.

Koul, R. K., and Allen, G. D. (1993). “Segmental intelligibility and speech interference thresholds of high-quality synthetic speech in the presence of noise,” *J. Speech Hear. Res.* **36**, 790–798.

Lewis, H., Benignus, V. A., Muller, K. E., Malott, C. M., and Barton, C. N. (1988). “Babble and random-noise masking of speech in high and low context cue conditions,” *J. Speech Hear. Res.* **31**, 108–114.

Liberman, A. M., and Mattingly, I. G. (1989). “A specialization for speech perception,” *Science* **243**, 489–494.

Logan, J. S., Greene, B. G., and Pisoni, D. B. (1989). “Segmental intelligibility of synthetic speech produced by rule,” *J. Acoust. Soc. Am.* **86**, 566–581.

Moulines, E., and Charpentier, F. (1990). “Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones,” *Speech Commun.* **9**, 453–467.

Orr, D. B., Friedman, H. L., and Williams, J. C. C. (1965). “Trainability of listening comprehension of speech discourse,” *J. Educ. Psychol.* **56**, 148–156.

O’Shaughnessy, D. (1987). *Speech Communication: Human and Machine* (Addison-Wesley, Reading, MA).

Pisoni, D. B., and Koen, E. (1982). “Some comparisons of intelligibility of synthetic and natural speech at different speech-to-noise ratios,” *J. Acoust. Soc. Am. Suppl.* **1** **71**, S94.

Pisoni, D. B., Nusbaum, H. C., and Greene, B. G. (1985). “Perception of synthetic speech generated by rule,” *Proc. IEEE* **73**, 1665–1676.

Remez, R. E., Rubin, P. E., Pisoni, D. B., and Carrell, T. D. (1981). “Speech perception without traditional cues,” *Science* **212**, 947–950.

Sproat, R. M., Ostendorf, M., and Hunt, A., eds., (1999). “The need for increased speech synthesis research” (a report of the 1998 NSF workshop for discussing research priorities and evaluation strategies in speech synthesis). (Available at <http://cslu.cse.ogi.edu/publications/>).

Stylianou, Y. (2000). “On the Implementation of the Harmonic Plus Noise Model for Concatenative Speech Synthesis,” Paper presented at ICASSP 2000, Istanbul, Turkey (available at http://www.research.att.com/projects/tts/papers/2000_ICASSP/fastHNM.pdf).

Warren, R. M. (1974). “Auditory temporal discrimination by trained listeners,” *Cogn. Psychol.* **6**, 237–256.

Modal analysis of a violin octet

George Bissinger

East Carolina University, Greenville, North Carolina 27858

(Received 1 February 2001; revised 21 December 2002; accepted 23 December 2002)

Experimental modal analysis of a complete Hutchins–Schelleng violin octet, combined with cavity mode analysis and room-averaged acoustic analysis, gives a highly detailed characterization of the dynamics for this historic group of instruments. All the “signature” modes in the open string pitch region—cavity modes A0 (“main air”) and A1 (lowest longitudinal), C-bout “rhomboid,” the first corpus bending modes $B1^-$ and $B1^+$ (comprising the “main wood”)—were observed across the octet. A0 was always the lowest dominant radiator, below all corpus modes. A1 contributed significant acoustic output only for larger instruments, but was the dominant contributor for the large bass in the “main wood” region. Acoustic results indicate *either* $B1^-$ or $B1^+$ can be the major radiator. Damping results indicate that B1 modes overall radiate approximately 28% of their vibrational energy. “Doublet” B1 modes from substructure couplings were observed for three instruments. “A0-B0” coupling was not significant for the largest instruments. The original flat-plate-based scaling of the “main wood” resonance was generally successful across the octet, although that for the “main air” was not. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1555614]

PACS numbers: 43.75.De [NHF]

I. INTRODUCTION

The Hutchins–Schelleng violin octet occupies an honored place in the field of musical acoustics because it is the very first example of the use of physical scaling laws to design a plucked or bowed string instrument.¹ Designing a bowed string musical instrument from scratch is a veritable “briar-patch” of difficulty—once inside it is impossible to move in any direction without serious consequences. As Schelleng remarked in 1963, “The process of tuning a violin includes more than adjusting string tension; the luthier must first ‘tune’ its pattern of resonances with respect to those of string and air.”² Deciding the pitch range of the open strings places you in the middle of the briar-patch; getting out requires navigation through the tangled mess of air, corpus (wood), and substructure modes, and their mechanical and vibro-acoustic couplings. The stringent ergonomic and practical limitations for a successful musical instrument—strength and size of the player (or their fingers), strength of materials, transportability, etc., all somewhat removed from desired pitch and resonance properties—create an envelope for mechanical design that is constrained ultimately to a positive *acoustic* judgment by the musician and listener.

Schelleng used the violin, generally considered the most successful of the bowed string instruments, as the design standard for the octet. To simplify scaling Schelleng chose two major low resonances, the “main air” (hereinafter called A0) and “main wood” resonances, originally identified from Saunders loudness curves.³ Assuming “similarity” of instrument shape he derived physical scaling laws based on cavity and flat plate theory and applied them to situate “main air” and “main wood” frequencies at prescribed pitches relative to the open strings. Schelleng did not investigate the shape of the instrument, primarily because no substantive analysis

was possible at that time. This first modal analysis of an octet experimentally evaluates the success of Schelleng’s scaling (and hence whether its underlying basic physical assumptions are appropriate), as well as other vibrational and acoustical aspects of the octet that are pertinent to the dynamic normal mode properties, and therefore the sound of the instruments.

The experimental modal analysis presented here—including expected and complicating vibro-acoustic coupling, cavity and room-averaged acoustic analysis—covered a complete violin octet. These instruments ranged over three and a half octaves (pitch of lowest string), starting at the high end with a treble (C. M. Hutchins, sus 123), soprano (sus 301), mezzo (sus 153, replacing traditional violin), alto (sus 137), tenor (sus 151), baritone (sus 198), small bass (sus 172), and ending with the large bass (sus 178). (The four instruments were all “shallow-ribbed.”) Rather than an exhaustive analysis of each instrument our goal is to emphasize normal mode similarities and trends across the octet for low-lying corpus “signature” modes, and to examine the success of scaling with principal emphasis on the “main wood.” A companion octet cavity mode analysis, newly incorporating wall compliance corrections and coupling between A0 and A1 modes and new scaling equations for the four lowest cavity modes, has already been published.⁴

II. EXPERIMENT

Measurements on the violin octet consisted of three major parts: (1) experimental modal analysis to characterize mechanical vibrations in terms of the instruments’ normal modes, (2) cavity mode analysis for reliable frequencies for the two lowest cavity modes A0 and A1, and (3) averaged room radiation measurements to determine the strongest radiating modes. On the larger instruments with endpins (alto to large bass) all measurements were made with these in the

⁴Electronic mail: bissinger@mail.ecu.edu

withdrawn position; all chin rests were removed from the smaller instruments and all strings were undamped, at pitch based on $A=440$ Hz.

A. Experimental modal analysis

All experimental modal analysis measurements on the octet were made under approximate “free-free” support conditions to ensure low support fixture damping,^{5,6} with zero-mass-loading excitation via automated force hammer impact parallel or perpendicular (for bridge and bridge feet measurements only) to the instrument plane at the bridge corner nearest the lowest string. Zero-mass-loading response measurements were made with a scanning laser Doppler vibrometer. Our support losses are estimated to be $\leq 5\%$ of the total damping, much lower than the $\sim 40\%$ in prior measurements that placed a violin on low-density foam supports and damped the strings with a foam strip.⁵ Suspension was always by some elastic under the in-curved C-bout regions on the lower bout (instrument hung neck down) with an alignment elastic contacting but not supporting the neck. This suspension minimized force hammer torque and hence significantly reduced rigid-body, post-strike motion (seen clearly as large amplitude peaks at the lowest frequencies). Corpus mode frequency errors were estimated as $\pm 0.1\% - 0.5\%$, dependent on the instrument. Total damping, ζ_{tot} , with estimated errors⁵ of $\pm 10\%$, was always expressed as percent-critical damping or %crit [$Q=(100/2\zeta)$ when ζ is expressed as %crit].

The alto and larger instruments were too large to be measured in the violin support fixture used for the treble-mezzo violins. These were moved outside the anechoic chamber into an adjustable support fixture capable of handling large size instruments. Roughly even-spaced grid point measurements were made on the top and back plates, ribs (along three parallel lines—top, bottom and rib midline for $\sim 3/4$ of their length), and major substructures (bridge, tailpiece, and neck-fingerboard-scroll). Substructures were measured in two orthogonal directions. No direct measurement of the strings was attempted, although string driven corpus motion was easily identified due to the frequency placement, harmonic structure and very low damping ($< 0.1\%$ crit).

There were 3200 line mobilities $Y(\omega)$ (=velocity/force) collected for >500 points over 0–4 kHz ($\Delta f=1.25$ Hz, treble-mezzo) or 0–2 kHz ($\Delta f=0.625$ Hz alto-large bass). Small Δf values allow more reliable extraction of damping from curve fitting, especially for low-lying modes. For larger instruments the $\pm 12^\circ$ vibrometer scan limit chosen to ensure accurate surface normal velocity measurements required greater distances, making it difficult to achieve high-quality reflected signals from some parts of some surfaces. Talcum powder or “retro-dust” at problem points generally provided reliable results. Unfortunately the baritone, measured first and then shipped out for use in a scheduled musical performance, did not benefit from numerous small improvements that benefited the other instruments, hence its mobilities were somewhat noisier, a real problem only for the weak modes. Situations with a too-glancing laser beam incidence make it impossible to measure motion reliable, a particular problem

for portions of the ribs. Additional experimental details are given in Refs. 5–7.

B. A0-A1 cavity mode analysis

Although the “main air” mode A0 is the dominant radiator at low frequencies, A1 must also be considered, not only because it can radiate effectively via induced motion,^{7,8} but also because coupling to A0 significantly changes A0’s volume dependence.⁹ Direct radiation through the f -holes for A1 is not expected to be significant since they lie near the nodal line for this mode. Experimental details for cavity mode analysis are quite straightforward, and have been covered in detail elsewhere.⁴

C. Radiation

The acoustic measurements were designed to elicit the *normal mode* acoustic response. A sound characterized as a sharp “bonk” is produced when the instrument bridge is trapped with the force hammer. An approximately 15-min audio binaural DAT recording of room-averaged sound for each instrument was recorded using a sound quality head with high-quality microphones inside the molded pinna and Fourier analyzed to provide a measure of normal mode acoustic radiation. The SQ head was placed 1.1 m along the diagonal from one corner of a quiet, fairly reverberant office, facing the suspended instrument 1.2 m further away. The corner was treated with 15-cm acoustic foam wedges to help reduce first-reflection interference from nearest back wall reflections. Recordings were made with the auto-hammer striking parallel or perpendicular to the instrument plane, and with the top plate facing the SQ head at distances $d > \lambda/2$ for frequencies > 140 Hz. Harmonic string resonances were seen (weakly in general) along with the corpus, cavity, and some coupled corpus-substructure resonances in the recorded sound. The room-averaged relative pressure spectra P_{av} were taken from analysis of >100 averaged sound spectra triggered on the initial transient to minimize room noise contributions, using the same analysis frequency/resolution settings as for experimental modal analysis.

D. Mode classification

The details of mobility and acoustic analysis are well covered in previous publications.^{4–7} Where feasible the normal modes were labeled following the classification scheme of Marshall for the violin.¹⁰ In general the whole instrument does not show significant motion for all modes. Often the substructures were quite passive while the corpus was active, and vice versa. No corpus modes were seen at frequencies below A0, only whole body or substructure modes for the tailpiece or neck-fingerboard-scroll. Typically, except for the smallest instruments, there were >100 modes excited over the measured frequency range. Corpus (tops+ribs+back) mode damping was typically $>0.5\%$ crit, falling slowly with increasing frequency; substructure (tailpiece, neck-fingerboard, bridge) damping was intermediate, $\sim 0.2 - 0.3\%$ crit, while string-induced motions were $< 0.1\%$ crit.

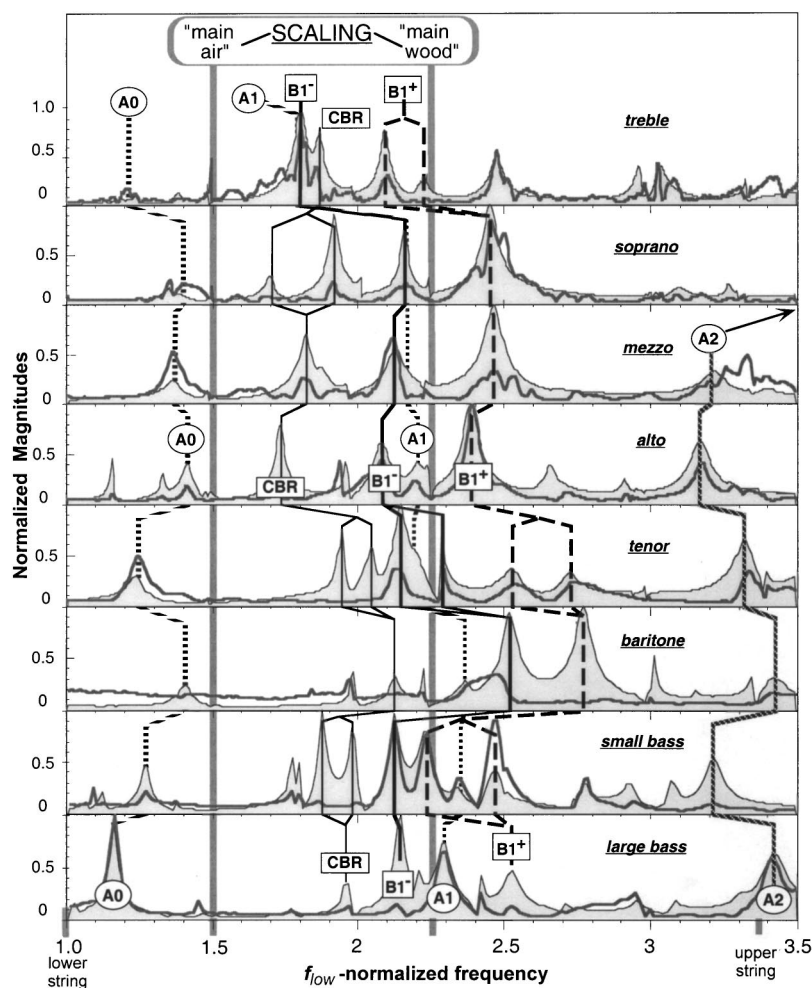


FIG. 1. Comparison of f_{low} -normalized $\langle Y \rangle$ (thin line, shaded) and room-averaged P_{av} (thick line) spectra for the violin octet. Signature A0 and A1 cavity (circled) and corpus CBR, $B1^-$, and $B1^+$ (blocked) modes have lines to guide eye; A2 mode also noted. Desired scaling placement for A0 and B1 at $1.5f_{\text{low}}$ and $2.25f_{\text{low}}$, respectively.

The lower cavity modes A0, A1 (lowest length mode), A2, and A4 (lowest lower-/upper-bout modes, respectively) all showed significant vibro-acoustic interactions with the corpus. In all cases the induced corpus motion was consistent with interior pressure distributions. A0, always below the lowest corpus mode, appeared as a “breathing” mode over the whole instrument; A1, usually in the vicinity of the first corpus bending modes, showed out-of-phase upper and lower bout “breathing.” A2 (and A4) fell significantly above the “main wood” resonance region (see Fig. 1).

If a “string” mode fell close to a corpus mode, the result commonly was a “doppelgänger” mode with similar appearance, damping higher than isolated string modes, but lower than the corpus mode, and animated motions with complex phase relationships between the corpus and various substructures, especially the bridge, in contrast to characteristic normal mode motions. Modes that fell in the open string frequency range had characteristic shapes that were distinguishable even when distorted. These include (1) “whole body” torsion, pitching, yaw modes, (2) tailpiece and neck-fingerboard “whole substructure” or bending modes, (3) cavity-mode (A0, A1, A2, A4) driven corpus motion, and (4) corpus modes such as C-bout “rhomboid” (CBR, labeled for shape of C-bout cross-section¹¹), with a strong shearlike top-back motion and a distinctive double-dagger (\ddagger) nodal line structure on top and back plates, or the first corpus bending modes ($B1^\pm$), also descriptively called

“baseball” modes because their nodal lines resemble the stitching pattern on a baseball.¹² The B1 modes had been identified previously as the “main wood” resonance, and our measurements confirm this identification. The strong, low-lying A0, A1, CBR, $B1^-$, and $B1^+$ modes seen in every instrument are referred to as “signature modes.”

III. RESULTS

Combining all measurements created a comprehensive, interlinked database with information on (1) the normal mode mechanical and acoustical properties of the lower modes for each instrument (including substructures), (2) A0, A1, A2, and A4 cavity modes (direct or indirect identification), (3) vibro-acoustic coupling between A0, A1, A2, and A4 cavity modes and the corpus, and (4) identification of modes with the strongest acoustic output. A summary graphical comparison of rms top+back plate mobilities $\langle Y \rangle$, superimposed on the room-average pressures P_{av} for each instrument over the range of the open strings (a little less than 2 octaves), is presented in Fig. 1 (parallel hammer strike). Each spectrum was amplitude normalized to the maximum $\langle Y \rangle$ or P_{av} for each instrument and then frequency normalized to f_{low} , the pitch of each instrument’s lowest open string for scaling comparison purposes across the entire octet. All signature modes were labeled, and the desired scaled “main air” and “main wood” resonance placements at $1.5f_{\text{low}}$ and

TABLE I. Torsion, pitching, and yaw “whole body” mode frequencies (Hz) for the violin octet. Also given is f_{low} , the pitch of the lowest string.

Mode	Treble	Soprano	Mezzo	Alto	Tenor	Baritone	Small bass	Large bass
Torsion ^a	145	161	141	151	89 ^b	57	55	42 ^b
pitching	272	201	159	109	66	68	42	38
yaw	325	259	206	126	89 ^b	71	51	42 ^b
f_{low} (Hz)	391	261	196	130	98	65	55	41

^aAll with tailpiece substructure motion.

^bMixed mode.

$2.25f_{\text{low}}$ noted with vertical lines. In the individual sections that follow, corpus, substructure, and whole body are discussed briefly.

A. Vibrations

1. Whole body modes

The low-lying “whole body” modes were almost always prominent in every instrument, viz. (i) torsion—rigid body corpus twist around the long axis with some substructure, generally the tailpiece, moving in a complementary way; (ii) pitching—corpus vertical pitching (as in boat or plane) with neck-fingerboard pitching in the opposite sense; and (iii) yaw—rigid body corpus horizontal pivot motion around neck joint with neck-fingerboard pivoting in opposite sense. Occasionally whole body modes appeared as a mixture, however the overall motion is always of the corpus moving in a simple way while a substructure moves in a compensating way to satisfy basic linear and angular momentum conservation laws. Whole body mode frequencies compiled for each instrument are listed in Table I, along with f_{low} from the pitch of each instrument’s lowest open string. These modes all produce little acoustic output, and the yaw mode, generally the highest frequency of the whole body modes, falls close to f_{low} .

2. Substructure modes

While localized motion of the substructures is generally not very interesting in acoustic terms, substantial coupling between substructure modes and the corpus modes was seen throughout the octet. If this coupling occurs with an acoustically important mode, then perceptible changes in sound can occur. The tailpiece and neck-fingerboard were quite active both in a localized motion and coupled with overall global corpus motion. Generally speaking these substructure motions were more prominent at lower frequencies. The bridge motion generally appeared as “going along for the ride,” since the vertical motion of the bridge-feet was taken directly from the top plate motion at two points directly in front, and bridge-top motion often looked string-driven.

The tailpiece also generally underwent “whole substructure” motions of torsion, pitching, and yaw. General pitching or yaw motions were seen for pivot points ranging from the end button through the approximate center of mass of the tailpiece to the string tuner position. The pitching mode presented here was the lowest frequency mode, which usually pivoted vertically about the end-button extremity. To place these modes in their proper relative placement across the

octet all frequencies were normalized to f_{low} . Such f_{low} -normalized “whole” tailpiece modes typically start at ~ 0.3 for the small instruments and increase to ~ 0.8 for the largest instruments. These are shown in Fig. 2. Because they almost uniformly lie below the pitch of the lowest string, and nowhere close to any of the major radiating modes where coupling might be of significance, there is little of acoustic interest in these lowest modes. However, any “tuning” process where tailpiece resonances are manipulated could certainly undermine this presumption. Pitching about other pivot points is of considerably more importance, e.g., pitching about the tailpiece midpoint was associated with the doublet structure observed for the $B1^+$ mode for the tenor and small bass (see later section).

At least two octaves above the lowest string, too high to

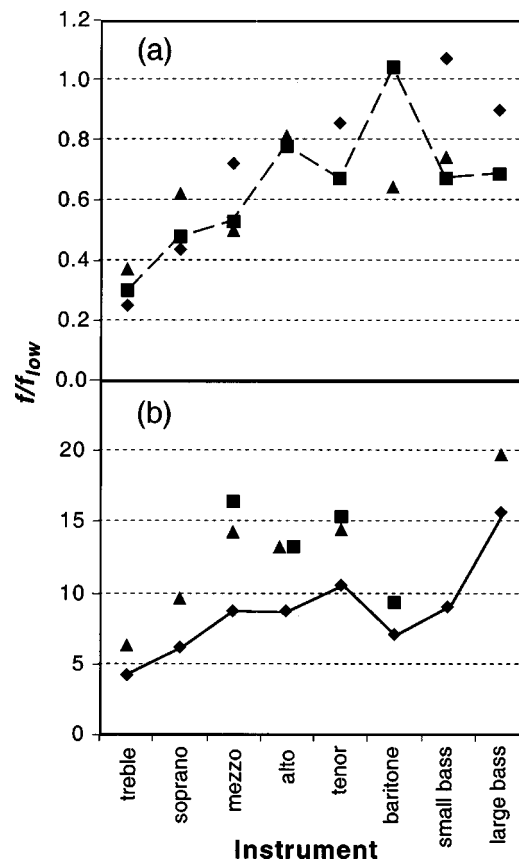


FIG. 2. (a) f_{low} -normalized “whole” tailpiece torsion (◆), pitching (■), and yaw mode (▲) frequencies in the violin octet; (b) as above but for first bending vertical (◆) and horizontal modes (■) and “twist” (▲) tailpiece modes. Not all violin octet instruments showed all modes. A line connects modes seen across the octet.

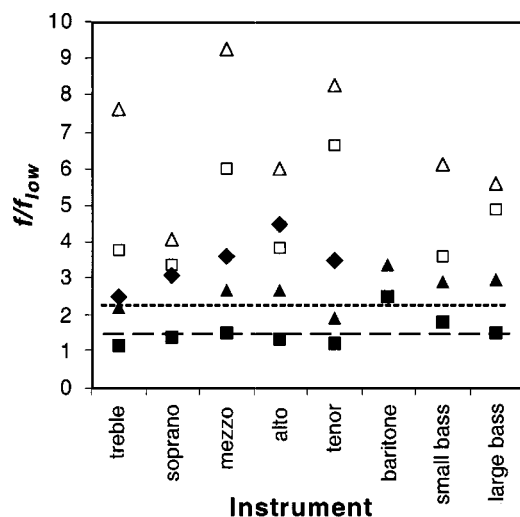


FIG. 3. Pitch-normalized neckfingerboard modes for the violin octet: torsion (◆), first vertical (■) and horizontal (▲) bending, and second vertical (□) and horizontal (△) bending. Desired “main air” ($f/f_{\text{low}}=1.5$) and “main wood” ($f/f_{\text{low}}=2.25$) placements shown with lines.

be involved in the “main wood” resonance, the tailpiece bending modes become apparent. The first vertical bending mode was seen in all instruments, but the horizontal was seen only for the four middle instruments. An additional “twist” mode was observed for most of the instruments at the broadened end near the end button while the rest of the tailpiece was relatively quiet. The f_{low} -normalized frequencies are plotted in Fig. 2 for the first vertical/horizontal bending and twist modes.

Neck-fingerboard vertical and horizontal bending modes were observed along with torsional modes. The torsion mode appeared to migrate into just the fingerboard portion of the substructure for the three largest instruments. Bending modes were also observed for just the fingerboard. The f_{low} -normalized frequencies for the first vertical bending mode usually place it close to the measured A0 position, a situation similar to that which leads to the so-called “A0-B0” coupling for the violin. Hutchins noted a preference for standard violins where the A0 mode was close in frequency to the first neck-fingerboard vertical bending mode¹³ (cum corpus pitching) labeled B0. Our measurements show that the frequency difference A0-B0 was within ± 1.5 semitones for the five smallest instruments. For the three largest instruments A0 was below B0 by at least 5 semitones, and hence any coupling would be weak. The first horizontal bending mode falls close to the “main wood” region and was often involved in B1 mode motions; the B1⁻ doublet in the treble originates from this coupling. The placement of these modes as well as the torsion and second horizontal and vertical bending modes can be seen in Fig. 3.

3. Corpus modes

Modes where the corpus was the dominant or codominant vibratory structure were labeled as corpus modes. Prominent substructure involvement was involved for many of these. The low-lying CBR and B1 modes were strong and identifiable across the octet, but few higher corpus mode nodal patterns were reliably identifiable across the octet. The

first corpus bending modes, established as the violin’s “main wood” resonance and one foundation of Schelleng’s scaling procedure for the violin octet, were clearly identifiable across the octet. Our acoustic and mobility measurements strongly supported this identification, with the unanticipated addition of the cavity mode A1 to the large bass “main wood.”

Horizontal rib motions (perpendicular to corpus long axis) were measured directly; vertical motion was inferred from edge motions of the top and back plates. Few modes had relatively strong rib motions; the CBR mode combined strong vertical (relative to plane of violin) and horizontal rib motion, while B1 motion was mostly vertical. Prominent horizontal rib motions fell into two major classes: (1) in-phase—most prominent for CBR—where both the bassbar- and soundpost-side ribs pivoted about the rib midline, with equivalent areas moving in the same direction, and (2) out-of-phase, with rib areas moving in opposite directions, most prominent for B1. Overall, rib motion was most prominent in the C-bout region. No measurements were made of motions along the violin long axis. It would be expected that some rib motion would occur along this direction, although the end blocks, neck joint, and top and back plate stiffness would tend to constrain this more than in the orthogonal directions.

The in-phase rib motion for the CBR mode suggests that the top and back plates move in the plane of the violin in opposite directions—essentially a “shear” motion. It seems plausible that rib properties—too high/low, too thick/thin, etc.—could be considerably more influential for this mode than for the others. As expected less “shear” motion was observed toward the upper and lower ends of the measured rib sections. Across the octet the CBR mode was always very poorly radiating, although it was always one of the most prominent mechanical modes. The out-of-phase rib motions characteristic of the B1 modes indicate that a transverse shrinking of the top plate in the C-bout regions is accompanied by a corresponding expanding of the back plate. This horizontal motion accompanies the observed vertical motion of both plates.

It was possible also to tell whether in-phase motion of the ribs was consistent with the top plate moving as a unit—or whether edge and center motion differed—by looking at bridge feet horizontal motion in the C-bout region. At lower frequencies the feet of the bridge generally followed the in-plane motion of the top of the ribs, but at higher frequencies this was not necessarily true.

The CBR mode fell below both B1 modes in all cases except the treble. In the treble it appeared between these bending modes. The treble is the narrowest (relatively), and shortest ribbed of the octet instruments and this has some influence in raising the CBR frequency. CBR modes were seen as doublets in the soprano (with tailpiece and neck-fingerboard coupling), tenor (with neck-fingerboard coupling), and small bass (neck-fingerboard). Coupling can wash out the ‡ nodal pattern somewhat, but generally it remains clear.

The B1 modes have the hallmark “baseball” nodal line pattern that switches between top and back plate but remains distinctive. In three instruments there were doublets of one or both of these. The treble, tenor, and small bass instruments

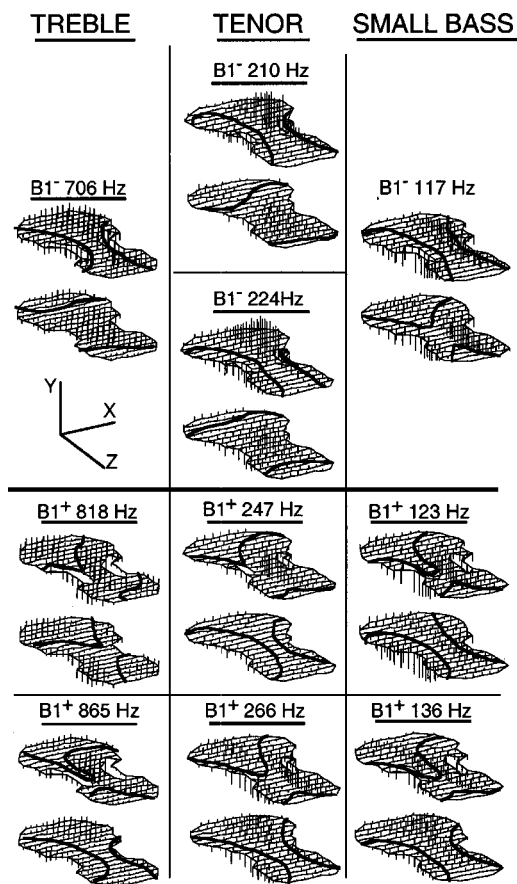


FIG. 4. B1 doublet top and back plate nodal lines superimposed on mobility eigenvector display.

showed strong coupling to some substructure. For example the tenor lower $B1^-$ at 210 Hz, with $\sim 1\%$ critical damping, showed non-“normal mode” bridge top motion indicative of a string (or strings) swinging back and forth rocking the top of the bridge. (The $B1^-$ modes bracketed the tenor A string tuned at 220 Hz.) The upper $B1^-$ at 224 Hz, with $\sim 0.2\%$ critical damping, showed bridge motion following the top plate motion, but evidence of different coupling to the neck-fingerboard first horizontal bending mode (seen isolated at 190 Hz). Under actual playing conditions with fingers on the strings it is not clear whether the 210-Hz mode would show up. On the other hand the tenor $B1^+$ doublet at 247 and 266 Hz showed coupling with a pitching tailpiece motion (pivot at the approximate center-of-mass) that underwent a 180° change of phase between lower and upper $B1^+$. A similar situation held true for the $B1^+$ doublet in the small bass. The treble $B1^-$ doublet showed coupling with the neck-fingerboard first horizontal bending mode that underwent a 180° change of phase between lower and upper doublet

member. The B1 modes all showed the out-of-phase rib motion discussed earlier, so that the vertical motion as seen in Fig. 4 also had a compression/expansion horizontal component along the plate edges in the C-bout region. The B1 mode frequencies are given in Table II.

The baritone was the only instrument in this group whose B1 modes did not bracket the desired “main wood” placement (Fig. 1). Higher-than-normal B1 frequencies argue that the plates for this instrument are too thick. Consistent with this, the position of A0 relative to the desired placement is higher than for any other large instrument. The baritone is also the only instrument whose A1 mode fell below the B1 modes.

B. Radiation

The room-averaged acoustic pressures P_{av} show the expected normal mode radiation peaks except for the case of the baritone, which was distinctly different in the “main air” region (see Fig. 1). The baritone has a band of enhanced acoustic response that ranged roughly from A1 to $B1^-$. It was also the only instrument where A1 was considerably below $B1^-$. As expected the “main air” peak A0 was generally prominent in the P_{av} results, but it was never as large as the B1 peaks except for the large bass. Our measurements show no preference for $B1^-$ or $B1^+$ overall as the strongest radiator in the region of the “main wood” resonance: $B1^-$ was stronger for the treble, mezzo, tenor, baritone, while $B1^+$ was stronger for the soprano, alto, and small bass. Both members of the B1 doublets radiate with roughly similar relative efficiency, using the $P_{av}/\langle Y \rangle$ ratio as a limited check.

Surprisingly, the large bass A1 cavity mode was the strongest radiator in the “main wood” region, far larger than the B1 modes (see Fig. 1). A1 had a prominent mobility peak, close in magnitude to the largest B1 mode. The large bass was also the instrument with the most compliant plates, the largest lower/upper bout volume ratio, and f -holes most below the mid-instrument line, factors that singly or in combination could enhance A1 radiation. Perhaps more surprising was A2, the lowest lower-bout mode, radiating almost as strongly as A1 in the large bass. Although an earlier experiment using heavier-than-air gas exchange in violin, mezzo violin, and viola instruments clearly indicated that A1 could radiate effectively via induced corpus motion when it was dropped below all the corpus modes,⁸ we believe that this is the first example of A1 being the major acoustic radiator in the “main wood” region. Hutchins’ retroactive claim that A1 should be included as part of the “main wood” resonance is partially supported by this observation. It also introduces a cautionary note in the evaluation of the “main wood” reso-

TABLE II. First corpus bending mode frequencies for the violin octet.

Mode	Treble	Soprano	Mezzo	Alto	Tenor	Baritone	Small bass	Large bass
$B1^-$	706	561	416	271	210 224	164	117	88
$B1^+$	818 ^a 865	638	482	311	247 266	181	123 136	104

^aStrong soundpost side bending, weaker bassbar side motion.

nance for two common larger instruments, the cello and double bass, which might possibly be expected to show significant contributions from A1.

Substructure coupling might also have some consequences for radiation. Violins typically have only about a half-dozen strong acoustic radiators among the ~ 25 – 30 corpus modes below 1 kHz,¹⁴ generally only one of which is a cavity mode, A0. B1 modes are in general efficient acoustic radiators, *and* strongly excited when energy is put into the instrument at the bridge, a simple prescription for strong radiation.^{15,16} If substructure coupling suddenly “doubles” the number of strongly radiating corpus modes in the “main wood” resonance region, some strengthening of instrument tone may be possible in this region, because vibrational energy would have a greater likelihood of being channeled into acoustic energy across a wider frequency range. Hutchins’ A0-B0 mode coupling, demonstrated live at the Tokyo String Quartet concert at the 1994 ASA meeting, is a good example of the efficacy of coupling for violin tone.

C. Damping

Great care was taken to ensure that reliable total damping ζ_{tot} —comprised of internal mechanical heat, radiative, and fixture losses⁷—could be extracted from the mobility measurements. With energy loss to the support fixture $\leq 5\%$ ζ_{tot} , $\zeta_{\text{tot}} \approx \zeta_{\text{int}} + \zeta_{\text{rad}}$. The overall estimated error for damping values was $\pm 10\%$ due to fit-related variations, but considerably larger for a weak peak on the shoulder of a strong one. ζ_{tot} values for A0, CBR, A1, B1[−], B1⁺, A2, and A4 modes were collected for all instruments and an average (\pm standard deviation) computed. The A0 mode always had the highest damping, with an average $\zeta_{\text{tot}} = 1.34 \pm 0.40$ %crit ($Q=37$). The A1 and A2 mode average dampings were similar, $\zeta_{\text{tot}} = 0.86 \pm 0.14$ ($Q=58$) and $\zeta_{\text{tot}} = 0.89 \pm 0.22$ %crit ($Q=57$), respectively, and quite close to the B1 mode. B1[−] had an averaged $\zeta_{\text{tot}} = 0.81 \pm 0.31$ %crit ($Q=62$), while for B1⁺ $\zeta_{\text{tot}} = 0.88 \pm 0.25$ %crit ($Q=57$). The CBR mode had an averaged $\zeta_{\text{tot}} = 0.61 \pm 0.28$ %crit ($Q=82$).

In favorable cases such as for B1 reasonable estimates of radiation damping can be computed from total damping results. The fraction of vibrational energy lost to radiation by the B1 modes, F_{rad} , can be calculated from $F_{\text{rad}} = \zeta_{\text{rad}} / \zeta_{\text{tot}} \approx 1 - \zeta_{\text{int}} / \zeta_{\text{tot}}$. By assuming that the prominent adjacent, *nonradiating* CBR mode, where $\zeta_{\text{tot}} \approx \zeta_{\text{int}}$, can provide a reasonable value of ζ_{int} for the B1 modes, the damping values given above provide a straightforward estimate that approximately 28% of the B1 vibrational energy is due to acoustic radiation, in good agreement with Schelleng’s estimate of “28% or more” (Ref. 2, but cf. erratum).

D. Scaling predictions

Schelleng based his violin scaling procedure on two fundamental assumptions: (1) *similarity of instrument shape* and (2) *flat-plate scaling*. Similarity in a normal mode context assumes that the governing physical properties of density, damping, elastic moduli, and shape are consistent, reflecting centuries-old empirical tradition by the maker. Of course there were noticeable changes in instrument proportions,

f-hole size and placement, and bridge shape and placement across the octet. The applicability of flat plate scaling to arched instruments is obviously of great practical interest. The scaling procedure was applied to only two violin resonances—the “main air” and the “main wood.” The desired placement of the “main air” was $1.5f_{\text{low}}$. For the six smallest instruments this coincides with the lower of the two middle strings. The small and large bass are tuned in fourths but desired placement pitchwise was unchanged. The desired placement of the “main wood” is $2.25f_{\text{low}}$. For the six smallest instruments this coincides with the upper of the two middle strings.

Schelleng considered the A1 cavity mode, which he grouped with all higher modes, as being of “little or no value” acoustically,² hence it was not mentioned in his scaling. However, A1 has undergone significant reappraisal since 1963. The aforementioned A0-A1 coupling effect on the A0 volume dependence, its ability to radiate strongly when dropped below the corpus modes by gas exchange, Hutchins’ retroactive grouping with the “main wood” resonance,¹ its acoustic importance for the large bass, plus its claimed value as a principal determinant of a violin’s quality,¹⁷ all argue for its inclusion here (although obviously “main wood” thereby becomes an inappropriate label).

1. “Main air”

As can be seen in Fig. 1 A0 always fell below the desired placement. Unsurprisingly placement was lowest for the largest and smallest instruments. Accurate predictions of the frequency of A0 were practically impossible due to at least two fundamental weaknesses in the original theoretical model: wall compliance effects and coupling with higher air modes. Furthermore, the large instruments’ original scaling suffered from normalization to an incorrect cello A0 frequency.¹ Both of these weaknesses have been at least partly addressed. Shaw’s 2 degree-of-freedom (DOF) network model¹⁸ is a significantly improved theoretical treatment for A0 and A1 that naturally incorporates their coupling. Employing an optimized version of this 2DOF model and adding a semi-empirical cavity wall compliance correction term, this 2DOF model was capable of predicting A0 and A1 frequencies within $\pm 9\%$ over the octet.^{4,9} For a group of instruments whose length varies by a factor of ~ 4.5 , *f*-hole area by ~ 10 , top plate thickness by ~ 3 , and volume by ~ 128 , this is a considerable improvement over any previous Helmholtz resonance equation result. Because the four largest instruments all had their rib heights much reduced from the original scaling predictions,¹ their A0 placement was probably more indicative of empirical than physical scaling.

2. “Main wood”

As can be seen from Fig. 1 generally the B1 modes fall in the desired position, straddling the desired placement $2.25f_{\text{low}}$ above the lowest pitch. However, B1 modes in the treble fall below desired placement, while in the baritone both lie above. The baritone misplacement is actually more surprising since it falls between other instruments where placement is close to desired. Perhaps in reducing rib height

another important scaling parameter, such as proportional rib mass, introduced undesired changes. (Part of Schelleng's scaling related total rib mass to plate mass.)

To improve agreement with desired scaling placement the most reasonable baritone modification—based on octet systematics seen in Fig. 1—would probably be to reduce top and back plate thickness somewhat, leaving the ribs alone. Our reasoning is that the CBR mode has a strong “shear” appearance and falls at approximately the desired frequency relative to that seen for the other instruments. This implies that the rib stiffness is approximately where “desired.” On the other hand, the B1 modes are far too high and these will drop down by thinning the plates.¹⁹ This will also drop the mass of the plates and bring the rib mass scaling into better agreement. Of course thinning the plates has the secondary effect of dropping A0 and A1 somewhat.

3. A1 and the “main wood”

A1 in the large bass induces a large corpus response, which produces an equally large room-averaged acoustic response, dwarfing that of the B1 modes. This was exceptional across the octet, although a few violins show lesser, similar behavior.⁷ A1 scales primarily by corpus length, hence the practical and ergonomic reasons that led to the larger octet members being progressively shortened relative to a “pure” length scaling (cf. Fig. 9 of Ref. 1) should lead to A1 falling progressively higher in frequency relative to the “main wood.” However, A1 falls in the “main wood” region, not because of scaling (none was ever attempted for A1) but as a result of the larger instruments becoming increasingly more compliant,¹ serendipitously compensating for an increasingly too-short cavity.⁴ Overall A1 usually falls among the B1 modes in terms of frequency placement. More importantly, the acoustic prominence of A1 for certain instruments is a practical justification for including it in the “main wood” resonance, which should then be relabeled, possibly adopting Schelleng's “principal body resonance” or some other broader, more inclusive term.

4. Summary

Schelleng's use of “flat plate” theory for carved arched instruments—using the average B1 mode frequency as the standard—can be justified *ex post facto*, although numerous departures from strict “similarity in shape” were encountered. For example, the arch/length ratio was not constant over the octet. It clustered about two values, with the four smaller instruments being about 20% larger than the four larger. Unfortunately this essential parameter still remains untouched in a scaling sense. Scaling of “main air” was unsuccessful due to an application of the simple-port, rigid-cavity Helmholtz resonator theory to a system for which it was clearly not appropriate.

In Fig. 5 actual A0 and B1 (averaged) mode placements are compared with the scaling standards set for the Hutchins–Schelleng violin octet. A1 is included for reference. Overall A0 is always low, especially so for the smallest and largest instruments, with only the alto falling within a semitone of the $1.5f_{\text{low}}$ scaling criterion. B1 is within a semi-

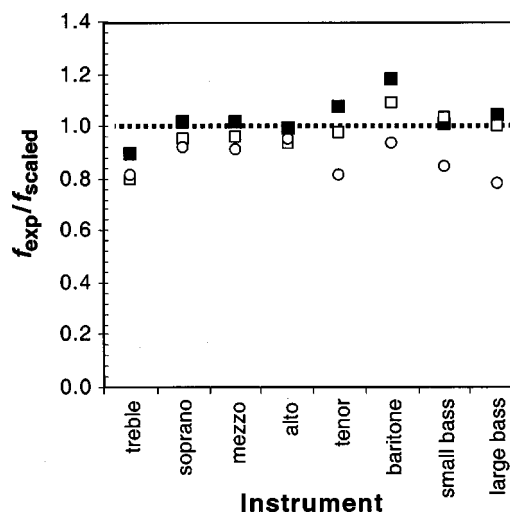


FIG. 5. Ratio of experimental to desired frequency placement for the “main air” A0 (○) and averaged “main wood” B1 (■), incorporating A1 (□) for the violin octet.

tone of the $2.25f_{\text{low}}$ scaling criterion for a majority of the instruments, with only the treble (barely) and baritone failing to bracket the desired position. Only the treble fails to have A1 within a semitone of $2.25f_{\text{low}}$. Overall this group of arched instruments should be judged as a successful application of “flat plate” scaling theory, especially so for a pioneering effort.

IV. CONCLUSIONS

Schelleng certainly understood the manifold difficulties inherent to any scientific study of the violin,²

“The violin family presents many unsolvable problems; its shape and the peculiarities of its materials were certainly not selected with regard to convenience in analysis. This however only emphasizes the need for understanding the simplicities that do exist and may even condone a certain amount of oversimplification.”

In this spirit we conclude that the “flat-plate-violin” approach to scaling the “main wood” resonance of the violin into new pitch regimes is of substantial practical value. The B1 modes of the violin appeared in all instruments, usually close to the desired placement, and commonly among the very strongest acoustic radiators. A1 appears also to have some basis for inclusion in this resonance region, due to its prominent acoustic presence for the largest instruments.

Individual failures in scaling observed here (e.g., the baritone) could just as readily be ascribed to an incomplete characterization of the materials used, or to the design envelope being modified for ergonomic reasons. Coupling effects were pervasive throughout the octet, inter- and intra-, cavity, and corpus modes. The lack of coupling mechanisms—wall compliance among them—was a weakness in Schelleng's original “main air-main wood” scaling. Hence its failure was, in retrospect, inevitable, but there surely was no failure in the underlying premise. Much improved A0-A1 cavity mode scaling is possibly already with an inherently coupled,

wall-compliance-corrected network model, pointing hopefully toward future improvements in other equally important areas.

ACKNOWLEDGMENTS

I am indebted to Carleen Hutchins for the loan of a complete violin octet for an extended period and for her support over many years. Without the support of the National Science Foundation (Grant No. DMR-9802656), which we gratefully acknowledge here, the VIOCADEAS Project, of which this research is only a part, would have been impossible. I wish to thank Gene Maddux and Mark French for providing a little retro-dust to sprinkle over otherwise impossible surfaces, John Kieffer for his extensive help in data acquisition, Amy Bissinger for her help in analyzing the modal analysis data, and Ken Jacobs for his work on the automated force hammer and large instrument support fixture.

- ¹C. M. Hutchins, "A 30-year experiment in the acoustical and musical development of violin-family instruments," *J. Acoust. Soc. Am.* **92**, 639–650 (1992).
- ²J. C. Schelleng, "The violin as a circuit," *J. Acoust. Soc. Am.* **35**, 326–338 (1963); cf. erratum, **35**, 1291 (1963).
- ³F. A. Saunders, "The mechanical action of violins," *J. Acoust. Soc. Am.* **9**, 81–98 (1937).
- ⁴G. Bissinger, "Wall compliance and violin cavity modes," *J. Acoust. Soc. Am.* **113**, 1718–1723 (2003).
- ⁵K. Ye and G. Bissinger, "Attaining 'free-free' normal mode frequency and damping conditions for the violin," *Proc. 18th Intern. Modal Analysis Conf.-Soc. Exp. Mechanics*, Bethel, CT, 2000, pp. 398–403.

- ⁶G. Bissinger and K. Ye, "Automated hammer-impact modal analysis with a scanning laser vibrometer: working example—a violin," *Proc. 18th Intern. Modal Analysis Conf.-Soc. Exp. Mechanics*, Bethel, CT, 2000, pp. 943–949.
- ⁷G. Bissinger and J. Keiffer, "Radiation damping, efficiency, and directivity for violin normal modes below 4 kHz," *ARLO* **4**, 7–12 (2003) (<http://ojps.aip.org/ARLO/top.jsp>).
- ⁸G. Bissinger and C. M. Hutchins, "A1 cavity-mode-enhanced fundamental in bowed violin and viola sound," *Catgut Acoust. Soc. J.* **3** (Series II), 11–13 (1988).
- ⁹G. Bissinger, "A0 and A1 coupling, arching, rib height, and f-hole geometry dependence in the 2-degree-of-freedom network model of violin cavity modes," *J. Acoust. Soc. Am.* **104**, 3608–3615 (1998).
- ¹⁰K. D. Marshall, "Modal analysis of a violin," *J. Acoust. Soc. Am.* **77**, 695–709 (1985).
- ¹¹G. Bissinger and A. Gregorian, "Relating normal mode properties of violins to overall quality," *Catgut Acoust. Soc. J.* (submitted).
- ¹²"Baseball" modes appellation for B1[±] due to Dr. Eric Arnold (student of G. Weinreich, private communication), circa 1980.
- ¹³C. M. Hutchins, "Effects of an air-body coupling on the tone and playing qualities of violins," *Catgut Acoust. Soc. J.* **44**, 12–15 (1985).
- ¹⁴G. Weinreich, "Violin radiativity: concepts and measurements," *Proc. Stockholm Music Acoust. Conf., R. Swed. Acad.* **46**(2), 99–110 (1985).
- ¹⁵G. Bissinger, "Some mechanical and acoustical consequences of the violin soundpost," *J. Acoust. Soc. Am.* **97**, 3154–3164 (1995).
- ¹⁶G. Bissinger, "Simulating violin sounds—may the (driving) force be with you," *Catgut Acoust. Soc. J.* **4**(6), 5–6 (2002).
- ¹⁷C. M. Hutchins, "A measurable controlling factor in the tone and playing qualities of violins," *Catgut Acoust. Soc. J.* **1**(4) (Series II), 10–15 (1989); "The future of violin research," *Catgut Acoust. Soc. J.* **2**(1) (Series II), 1–7 (1992).
- ¹⁸E. A. G. Shaw, "Cavity resonance in the violin: network representation and the effect of damped and undamped rib holes," *J. Acoust. Soc. Am.* **87**, 398–410 (1990).
- ¹⁹H. Meinel, "On the frequency curves of violins," *Akust. Z.* **2**, 22–33 (1937).

Laryngeal biomechanics and vocal communication in the squirrel monkey (*Saimiri boliviensis*)

Charles H. Brown^{a)}

Department of Psychology, University of South Alabama, Mobile, Alabama 36688

Fariborz Alipour

Department of Speech Pathology and Audiology, Wendell Johnson Speech and Hearing Center,
The University of Iowa, Iowa City, Iowa 52242

David A. Berry

School of Medicine, University of California at Los Angeles, Los Angeles, California 90095

Douglas Montequin

Department of Speech Pathology and Audiology, Wendall Johnson Speech and Hearing Center,
The University of Iowa, Iowa City, Iowa 52242

(Received 28 February 2002; revised 2 May 2002; accepted 2 October 2002)

The larynges of eight squirrel monkeys were harvested, dissected, mounted on a pseudotracheal tube, and phonated using compressed air. Patterns of vocal fold oscillation were compared with sound spectrograms of calls recorded from monkeys in our colony. Four different regimes of vocal fold activation were identified. Regime 1 resembled typical human vowel production, with regular vocal-fold vibration, a prominent fundamental frequency, and an accompanying series of harmonic overtones. This regime is likely to give rise to squirrel monkey “cackles,” as well as a variety of other harmonically structured calls. In regime 2, the pattern of vibrations exhibited the presence of two or more unrelated frequencies (biphonation). This regime of glottal activity resembled the biphonation observed in many exemplars of “twitter” and “kecker” calls. The vocal folds oscillated continuously in regime 3, but produced glottal pulses whose amplitudes waxed and waned rhythmically. This phenomenon resulted in the percept of a series of discrete pulses, and may give rise to “errs,” “churrs,” and other calls composed of a rapid sequence of acoustic elements. In regime 4, the period of each oscillation was quasi-irregular. Shrieks and other broadband calls or call elements that lack an apparent fundamental frequency may be produced in this manner. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1528930]

PACS numbers: 43.80.Ka, 43.40.Dx, 43.70.Bk [WA]

I. INTRODUCTION

Like humans, primates are conspicuously vocal, and are capable of modifying both the shape of the airway and the source characteristics of the oscillations of the vocal folds (Brown and Cannito, 1995a, 1995b). Furthermore, humans and primates share a common plan in the development of the layers of the vocal folds. A well-developed vocal ligament layer, composed of inextensible collagenous fibers, is sandwiched between the body of the vocal folds, the thyroarytenoid muscle, and the mucosa, the epithelial layer of the folds (Hast, 1983; Hirano and Sato, 1993). Because of the inextensible property of the ligament, vocalizers may potentially taunt the ligament layer (via contraction of the cricothyroid muscle) while simultaneously allowing the mucosal layer of the vocal folds to remain comparatively lax. This particular combination, the capacity for a taunt ligament and lax mucosa, is not present in all mammals, and it may be an adaptation favored to heighten the capacity for sustained phonation (Titze, 1994), as is characteristic of singing, speech, and some primate vocalizations. Though theoretical models of animal voicing pertinent to primate phonation

have been developed (Mergell *et al.*, 1999), no studies of airflow or direct measurements of the biomechanics of voicing have been published for any nonhuman primate.

The present study adopts the excised larynx preparation to measure airflow, voicing regimes, and their implications for vocal communication in the squirrel monkey. Of all the primates, the squirrel monkey has likely received the most intensive study of vocal behavior (Newman, 1985). Vocal behavior has been studied under field (Winter, 1972; Boinski and Newman, 1988; Boinski, 1991; Boinski and Mitchell, 1992, 1995) and laboratory conditions (Winter *et al.*, 1966; Schott, 1975; Smith *et al.*, 1982). The vocal control centers in the brain have been mapped (Jürgens, 1976a, b, 1988, 1998; Jürgens and Pratt, 1979; Jürgens and Zwirner, 2000), and the effect of lesions on vocal behavior have been studied for a wide variety of sites (Jürgens *et al.*, 1978; Thoms and Jürgens, 1981; Jürgens *et al.*, 1982; Kirzinger, 1985; Kirzinger and Jürgens, 1985). The wealth of knowledge of vocal behavior and the neurobiology of vocal control make the squirrel monkey an attractive candidate for studies of airflow and phonation. Because the squirrel monkey's vocal fold length, at 4–5 mm, approximates the vocal fold dimensions of the human neonate (Hirano *et al.*, 1983; Smith and Gray, 1994), this work may be pertinent to researchers interested in

^{a)}Electronic mail: cbrown@usamail.usouthal.edu

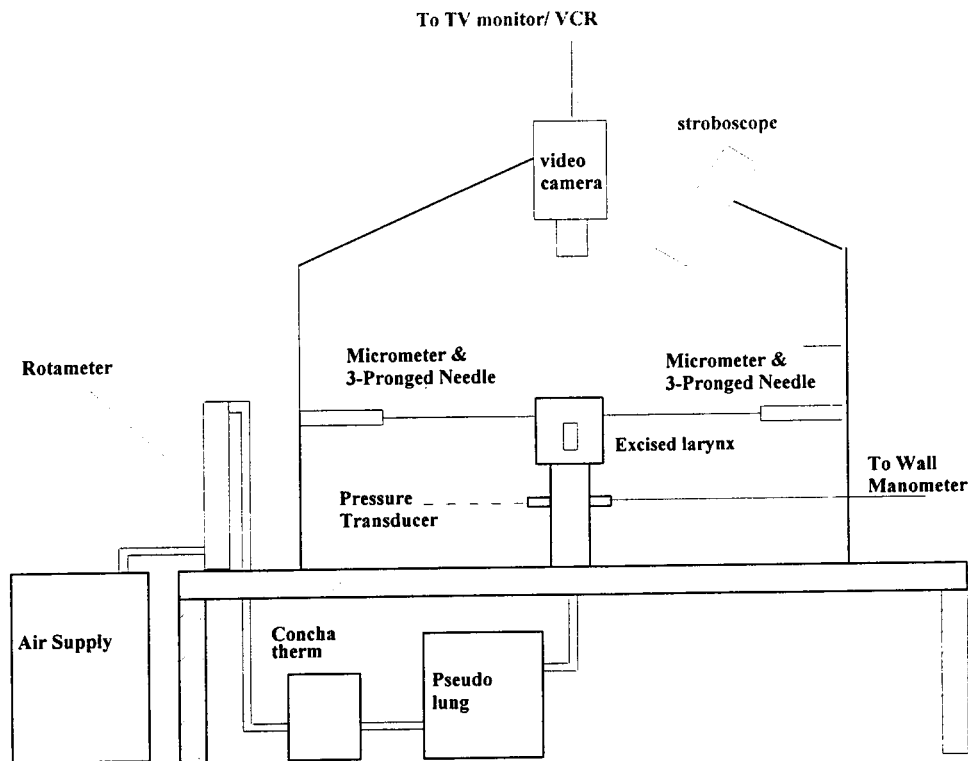


FIG. 1. Schematic of the excised larynx apparatus.

infant cries, the biomechanics of infant voice characteristics, and the development of motor control systems underlying vocal and verbal behavior.

II. METHODOLOGY

A. Subjects

Excised squirrel monkey larynges were obtained from the Squirrel Monkey Breeding and Research Resource, University of South Alabama. The Squirrel Monkey Breeding and Research Resource, housing approximately 500 animals, is the largest squirrel monkey colony in the United States, with a low annual mortality of about 5%. The larynges of eight monkeys were harvested from animals which suffered a natural spontaneous death. No monkeys were euthanized for the purpose of conducting this research. Larynx ID 1630, 4510, and 90780 were extracted from adult female Bolivian squirrel monkeys (*Saimiri boliviensis boliviensis*). Larynx ID 2683 was removed from an adult male Bolivian squirrel monkey. Larynx ID 1232 was harvested from an adult female Guyanese squirrel monkey (*Saimiri sciureus sciureus*). The remaining three larynges were harvested from the Peruvian subspecies (*Saimiri boliviensis peruviansis*). Of these, Larynx ID 410 was harvested from an adult male, while larynx ID 742 and 90004 were obtained from adult female specimens.

Adult male squirrel monkeys, also obtained from the Squirrel Monkey Breeding and Research Resource, served as subjects for tape recordings of exemplars of the squirrel monkey call repertoire.

B. Apparatus and procedure

Experiments were conducted in an IAC single-walled booth in which the interior surfaces were covered with Sonex foam to reduce acoustic reflections. Each larynx was dissected and trimmed. The trachea and larynx were mounted on a pseudotracheal tube, positioned on a laboratory bench, and the tracheal axis was oriented in the vertical position, exposing the glottis to a camera and recording apparatus (Fig. 1). Adduction was controlled by two micrometers connected to the arytenoids by surgical thread sutured to the arytenoid tissue, or by needle prongs which could be oriented to press against the left and right arytenoids. In one condition, the length of the vocal folds was not manipulated, and the larynx was permitted to phonate freely without any attachments to the thyroid tissue. In the second condition, vocal fold length was manipulated. Two methods were employed: In one method, length changes were controlled by a micrometer attached directly to the thyroid prominence above the anterior commissure. In the second method, surgical thread was sutured to the thyroid cartilage and a weight and pulley arrangement was employed to apply a known force to stretch the vocal folds. No attempt was made to apply asymmetrical adjustments to differentially lengthen the left and right vocal folds.

The pseudotracheal tube received air from the building's compressed and filtered air supply. The air was heated to 37 °C via a Concha Therm III Servo Control Heater (RCI laboratories, Arlington Heights, IL), and was humidified to approximately 100% relative humidity. The mean air pressure below the glottis was monitored with a wall-mounted water manometer (Dwyer No. 1230-8), and the mean flow rate was monitored with an in-line flowmeter (Gilmont rota-

meter model J197). The top view of the larynx and vocal folds was videotaped with a Hitachi video camera (KP-D50) and Panasonic recorder (AG-1960) for later image analysis. For stroboscopic images, a Pioneer DS-303ST stroboscope was employed. The sound pressure level of the audio signal was measured with a Quest (model 2700) sound level meter positioned 20 cm above the glottis, and audio recordings of the signal were obtained with a Shure (model 48) microphone also positioned 20 cm above the glottis. The pressure and microphone signals were recorded on a Sony model PC-108M Digital Audio Tape (DAT) recorder, and simultaneously filtered, sampled, digitized (12-bit A/D, 44.2 kHz sample rate), and stored on a Gateway personal computer.

Sound samples from the excised larynx preparation were compared with tape recordings of call exemplars recorded from subjects housed in the Squirrel Monkey Breeding and Research Resource. These recordings were conducted with a Sennheiser MKH816T directional microphone, Sennheiser BP-2 microphone power supply, and a Sony TCD-D10 digital audio tape recorder. The distance from microphone to monkey was typically 1 m.

The digitized audio samples were analyzed with TFR signal processing software (Avvaz Innovations). The spectral analysis was calculated with a fast Fourier transform utilizing the following parameters: preemphasis 0, Hamming window, 256 FFT frequency bands, 50% overlap. Because the optimal parameters for signal analysis are dependent upon the frequency, temporal and spectral properties of the signal in question, and because squirrel monkeys vocalize over a very broad range, when warranted, we modified these parameters to heighten the resolution of selected features in samples of interest.

The area of the maximum glottal opening was measured for selected subglottal pressures by scoring videotaped images of the oscillations of the vocal folds. For each step in subglottal pressure, the stroboscope strobe rate was adjusted to capture the pattern of oscillations of the vocal folds. The maximum opening of the glottis was determined through a frame-by-frame inspection of several glottal cycles, and the profile of the maximum glottal opening was traced on transparency film overlaid upon the video monitor. The relative area of the glottal opening was derived by overlaying the transparency film on graph paper, and counting the number of 1 mm squares encompassed by the glottal profile. The magnification factor of the video system was determined by comparing the video images of a millimeter grid positioned adjacent to the glottis with the size of a millimeter grid positioned on the video monitor. The inverse of the magnification factor was used to convert the relative area of the glottal opening to absolute values.

III. RESULTS

Each larynx was readily phonated, and each larynx exhibited several different regimes of phonation. Because the squirrel monkey preparation has not been studied before, there were no established guidelines for a physiologically reasonable range of subglottal pressures. Hence, we incremented subglottal pressures cautiously. However, in our second subject, we incremented subglottal pressure to

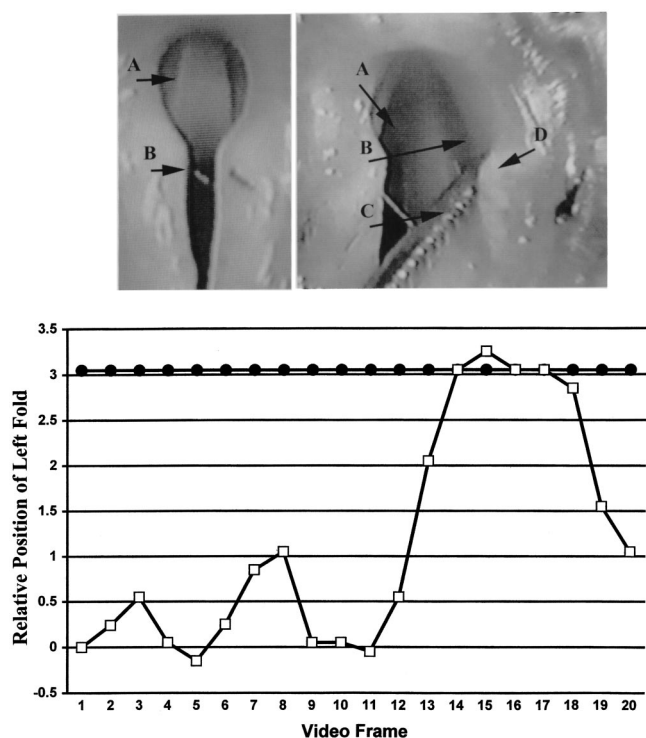


FIG. 2. The ventricular and vocal folds in the squirrel monkey. A superior perspective of the ventricular fold (B) overlaying the vocal fold (A) is shown in the top left panel. Top right panel: the prong of a hemostat (C) was inserted under the right ventricular fold (D) revealing the ventricle (B) and the superior surface of the right vocal fold (A). The relative position of the left ventricular fold (solid circle) and left vocal fold (open square) is shown for 20 successive video frames during phonation. The position of both folds was captured by a strobe light, and the position of each fold was measured in arbitrary screen units measured from a grid superimposed on a video monitor. Zero on the ordinate represents the midline, and values above zero represent relative locations to the left of midline.

40 cm-H₂O, and in this particular larynx, pressure at this level appeared to damage the vocal folds (we were unable to replicate samples obtained at lower levels). Thus, we discontinued experiments with this larynx, and kept subglottal pressures below 38 cm-H₂O with our remaining larynges. Summed across all eight larynges, we obtained 501 samples of phonation.

A. Tissue vibration induced by airflow

In Fig. 2 (top left panel) we show that, from a superior perspective, the ventricular folds (B) overlay and obstruct visual observation of much of the length of the vocal folds (A). Only about the anterior third of the length of the vocal folds could be observed in the present specimen (larynx 90780). This degree of visual obstruction is representative of our sample. In Fig. 2 (top right panel) a prong of a hemostat (C) was inserted under the ventricular fold (D), and the tissue was reflected to reveal the ventricle (B) exposing the superior surface of the vocal fold (A). In shape, the vocal folds resembled the profile shown in Fig. 3. In agreement with Kahane *et al.* (2002) we observed a ridge on the superior margin of the vocal folds. However, it was unclear if this structure conformed to the geometry of the vocal membrane (Mergell *et al.*, 1999), as described for squirrel monkeys by Starck and Schneider (1960). The relative profile of the

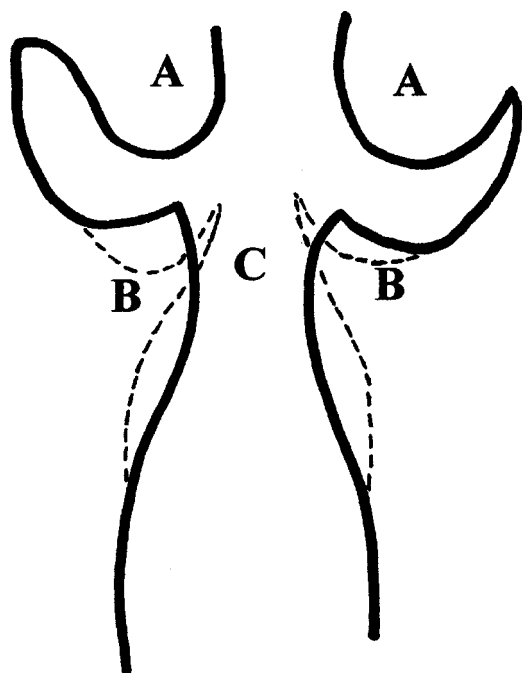


FIG. 3. The approximate profile of the cross section of the squirrel monkey larynx inspired by measurements from Kahane *et al.* (2002). The glottis (C) is medial to the vocal folds (B), and positioned inferior to the ventricular folds (A). For comparison, the relative profile of the vocal fold and vocal membrane complex described by Starck and Schneider (1960) for the squirrel monkey are depicted by the dashed line.

squirrel monkey vocal folds and vocal membrane reported by Starck and Schneider (1960) scaled to the Kahane *et al.* (2002) measurements are depicted by the dash line in Fig. 3.

When the flow of compressed air induces phonation in the larynx, vibrations are established in the vocal folds (A, Fig. 2, top left panel), but not in the ventricular folds (B). This is shown in the bottom panel of Fig. 2. The relative positions of the left ventricular fold (solid circle) and the left vocal fold (open square) were plotted for 20 successive video frames illuminated by a stroboscope. At frame one, the vocal fold was at midline, and the ventricular fold was at three units lateral to midline. Over the first 11 frames the vocal fold was captured at positions near the midline, and at frames 12–20, the strobe captured the vocal fold at positions lateral to the midline. The important point is that the ventricular folds were stationary during phonation, and that oscillations of the vocal folds were the source of phonation.

In Fig. 4 we show stroboscopic pictures of vocal fold oscillation for eight successive video frames for larynx ID 90780 following the removal of the ventricular folds. For this sequence an intermediate subglottal pressure was used (26 cm of H^2O). As may be observed in reference to a millimeter grid displayed in the first frame of Fig. 4, the length of the vocal folds for larynx 90780 was approximately 5 mm, and the maximum width of the glottis at this subglottal pressure was about 1 mm. These values are representative of the eight larynges we studied.

B. Phonation threshold pressure

The subglottal phonation threshold pressure varied between 5 and 8 cm- H^2O under the lax condition (no attempt

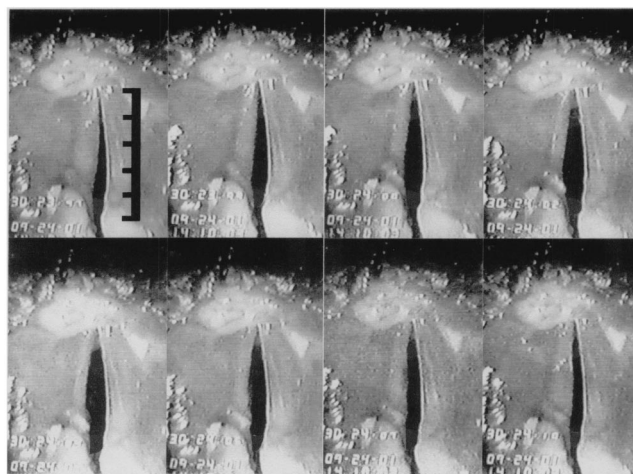


FIG. 4. Eight successive video frames of larynx 90780 oscillating from expiratory airflow at a subglottal pressure of 26 cm H^2O . A millimeter grid was superimposed on the first frame.

to increase adduction or lengthen the vocal folds). These values are similar to measurements obtained in our laboratory with dogs (Alipour *et al.*, 1997). Phonation threshold pressure increased with increments in vocal fold tension and adduction. In Fig. 5 we show the relationship between phonation threshold pressure and vocal fold tension for low adduction (larynx 1223, solid diamonds) and moderate adduction (larynx 2683, open squares) conditions. Under these conditions, phonation threshold pressure varied between 10 and 28 cm- H^2O .

C. Maximum glottal area

The maximum glottal area varied as a function of subglottal pressure. In Fig. 6 we display the maximum glottal area for larynx 90780 (solid diamonds) and larynx 2683 (open triangles) as a function of subglottal pressure for the lax condition. At the onset of phonation the maximum glottal area was about 2.0 mm², and the maximum glottal area increased as subglottal pressure was incremented reaching a maximum area of 3.6 mm² at a subglottal pressure of

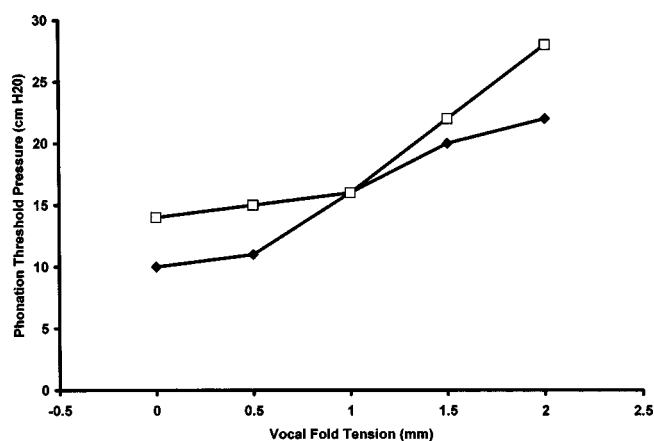


FIG. 5. Phonation threshold pressure as a function of vocal fold tension. Tension was incremented by changing the position of micromanipulators to stretch the vocal folds in $\frac{1}{2}$ -mm steps. Two levels of adduction were employed: low adduction (larynx 1223, solid diamond) and moderate adduction (larynx 2683, open square).

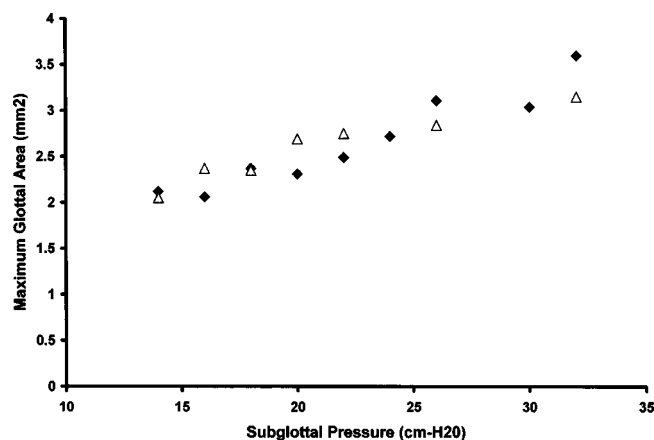


FIG. 6. Maximum glottal area as a function of subglottal pressure. Measurements of glottal area were conducted under the lax condition with no tension applied to increase adduction or vocal fold length for larynx 90780 (solid diamond) and larynx 2683 (open triangle). The correlation coefficient between maximum glottal area and subglottal pressure was 0.96 for each larynx.

32 cm-H²O. The correlation coefficient between maximum glottal area and subglottal pressure was 0.96 for each larynx, indicating the relationship between maximum glottal area and subglottal pressure was approximately linear.

D. Airflow

In Fig. 7 we show the relationship between airflow and subglottal pressure for larynx 4510 and 90004. In each case, phonation was initiated at subglottal pressures of between 5 and 8 cm-H²O, and phonation was sustained as subglottal pressure was increased to a maximum value of 30–35 cm-H²O. We conducted airflow measurements under three conditions: no tension on the vocal folds (solid diamonds), adduction increased (shaded squares), and vocal folds shortened (open triangles). In each of these conditions,

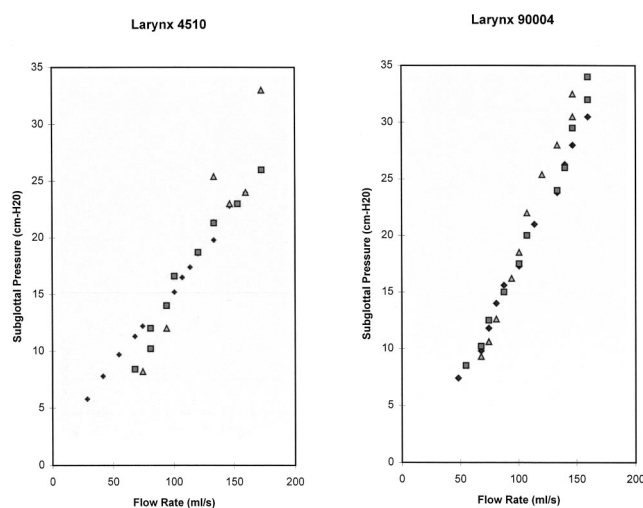


FIG. 7. The relationship between airflow and subglottal pressure for larynx 4510 and 90004. Airflow measurements were conducted under three conditions: no tension on the vocal folds (solid diamonds), adduction increased (shaded squares), or vocal folds shortened (open triangles). The correlation coefficient between airflow and subglottal pressure varied between 0.96 and 0.99 for the three conditions.

the relationship between flow rate and pressure was fairly linear. The lowest correlation coefficient between airflow and subglottal pressure was 0.96. This value was obtained for the vocal fold shortened condition for subject 4510. The mean coefficient for all six correlations was 0.99. Hence, under all three conditions, no tension on the vocal tissue, adduction increased, and vocal folds shortened, the relationship between flow and pressure was very orderly. These results are similar to our findings with canines (Alipour *et al.*, 1997).

During phonation, airflow rates for squirrel monkeys ranged between 20 and 200 ml/s. These values are, in comparison, rather small relative to the values reported for canines. In 20–30-kg dogs, airflow rates typically range between 200 and 800 ml/s (Alipour *et al.*, 1997). However, the size of the squirrel monkey trachea and larynx are about $\frac{1}{8}$ the dimension of the corresponding structure in canines, and the flow rates observed here are congruent with the size differences between these preparations.

E. Sound pressure level

At a propagation distance of 20 cm, the sound pressure level at the threshold of phonation was approximately 52 dB. We were able to obtain stable phonation for amplitudes as low as 45 dB SPL by gradually reducing subglottal pressure once phonation was established. The maximum sound pressure level we recorded was 81 dB. Hence, within our pool of 501 samples, we observed a 36-dB dynamic range in sound pressure levels. Because we were concerned about the risk of damaging the vocal folds with physiologically excessive subglottal pressures, as noted above, we kept subglottal pressures below 38 cm-H²O. It remains possible that squirrel monkeys occasionally vocalize with still higher pressures, and their dynamic range may exceed the values we observed here.

F. Regimes of vocal fold oscillation

Four different regimes of squirrel monkey voicing were observed, and evidence of each regime was observed in exemplars of natural vocalizations.

1. Regime 1: Periodic phonation with overtones

Figure 8(a) shows the spectrogram for a representative exemplar of periodic phonation. This regime of voicing resembles that underlying typical human vowel production with stable vocal fold vibration, a prominent fundamental frequency, and an accompanying series of harmonic overtones. “Cackles” [Fig. 8(b)] and a number of other harmonically structured calls are attributed to the glottal source characteristics shown in Fig. 8(a).

2. Regime 2: Biphonation

Figure 9(a) shows a representative sample in which the frequency peaks are not a product of a simple harmonic series. Four main frequency peaks are apparent in Fig. 9(a) at 1500, 2411, 4909, and 5770 Hz, respectively. Given the ratio of the frequencies, and the respective amplitudes of the

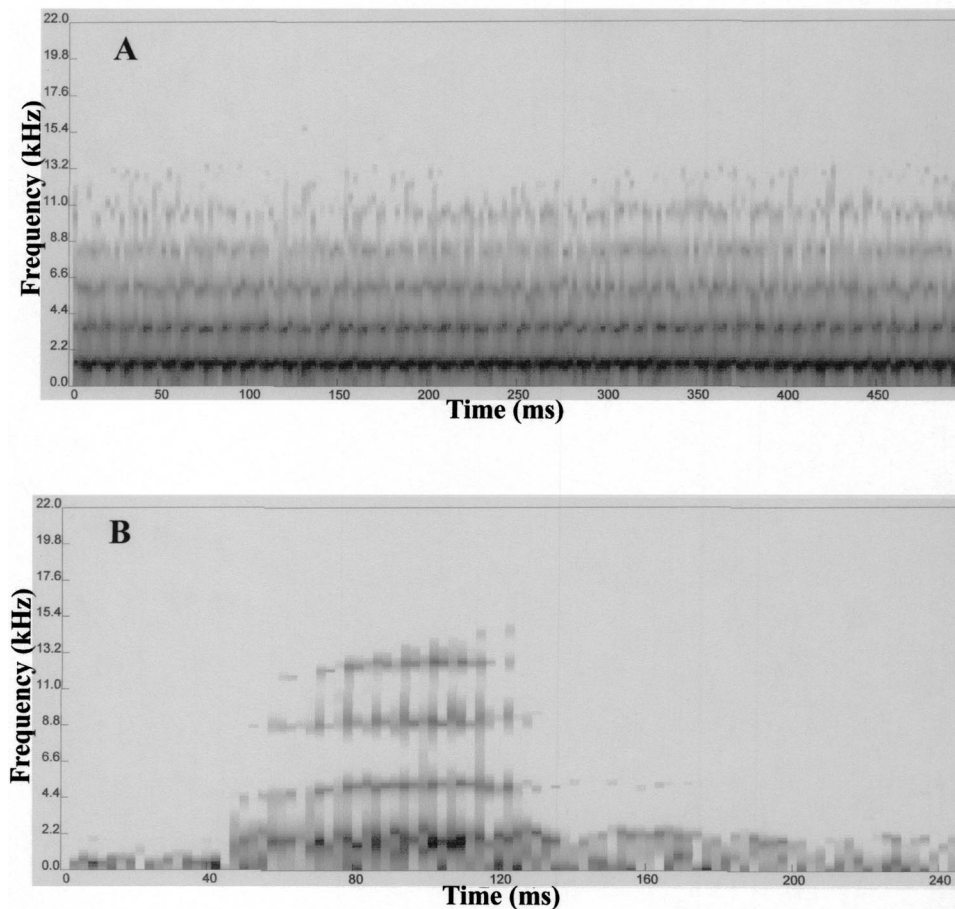


FIG. 8. (a) Sound sample recorded from an excised larynx exhibiting a prominent fundamental frequency, and an accompanying series of harmonic overtones representative of a limit cycle attractor. (b) An exemplar of a squirrel monkey cackle exhibiting a prominent fundamental frequency, and an accompanying series of harmonic overtones. Note that the resonate characteristics of the supralaryngeal vocal tract would influence the relative amplitudes of the frequency peaks in the spectrum).

peaks, this sample must be the product of biphonation, a phenomenon where two or more uncoupled oscillators have simultaneously been set in vibration.

Squirrel monkeys produce a variety of calls including many exemplars in the “twitter,” “trill,” “chuck,” and “kecker” classes [see Fig. 9(b)] which exhibit biphonation (frequency bands which are not multiples of a single fundamental frequency, or uncoupled frequency contours for different frequency bands). Hence, biphonation is apparent in a variety of calls in the squirrel monkey vocal repertoire. Thus, in squirrel monkeys biphonation, *per se*, it is not indicative of a pathology or health disorder (Riede *et al.*, 1997; Riede and Stolle-Malorny, 1999), but rather is a characteristic feature of some classes of vocalizations.

3. Regime 3: Staccato phonation

Unexpectedly, we also obtained exemplars in which the sample appeared to be composed of a sustained series of pulses. A representative sample of this phenomenon is displayed in Fig. 10(a). In this example, each pulse, typically 50–60 ms in duration, exhibited a frequency peak at 2700 Hz accompanied with a subharmonic at 1350 Hz. The interval between pulses varied between 10 and 50 ms. The pulse rate in this example was 8–10 Hz. Close inspection of the waveform [Fig. 11(a)] shows that phonation never ceased, rather the amplitude of the wave waxed and waned, yielding the impression of distinct segments.

Squirrel monkeys produce a wide variety of pulsed calls including “errs,” “churrs,” and “trills” which share the

common property of being composed of a rapid sequence of elements. The spectrogram in Fig. 10(b) shows two pulses of a churr segmented by a 12-ms interpulse interval [the waveform for this call is shown in Fig. 11(b)]. It is likely that many pulsed calls are the product of a voicing regime in which the oscillations of the vocal folds rhythmically modulate the amplitude of the airflow waveform.

4. Regime 4: Aperiodic phonation

In Fig. 12(a) we show the spectrum of a sample characterized by aperiodic or irregular vocal fold vibration. A fundamental frequency cannot be identified for samples produced by this regime, and because the periods for each successive oscillation of the folds are irregular, no adjustment of the strobe can be found which will “freeze” this pattern of vibration. These findings are consistent with prior reports of aperiodic or chaotic regimes of vocal fold oscillation (Titze *et al.*, 1993; Berry *et al.*, 1994). Shrieks, and a variety of other broadband calls, or call elements, emitted by squirrel monkeys (and many call exemplars emitted by most other species of primates) also lack an apparent fundamental frequency [Fig. 12(b)], and have recently been attributed to a chaotic regime of voicing (Wilden *et al.*, 1998).

In Table I we show that at low subglottal pressures regimes 2 and 4 were absent, while at high subglottal pressures regime 3 was nearly absent. At intermediate subglottal pressures exemplars of all four regimes were apparent. Hence, exemplars of periodic phonation were apparent over the en-

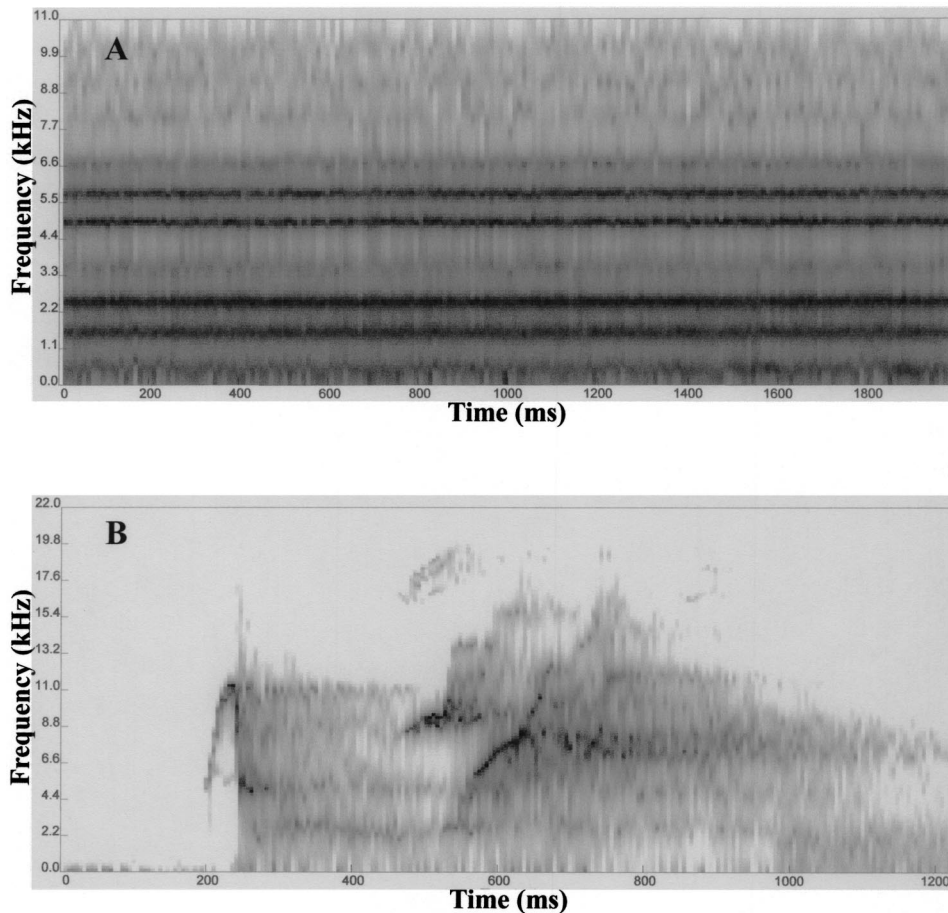


FIG. 9. (a) Sound sample recorded from an excised larynx exhibiting biphonation. The frequency peaks were at 1550, 2411, 4909, and 5770 Hz, respectively. (b) A sequence of two squirrel monkey chuck calls in which biphonation was exhibited in the second chuck. The spectrum shows that the frequency bands are not harmonically related.

tire range of subglottal pressures studied, while the other three voicing regimes tended to be exhibited over a narrower range of pressures.

G. Regime instability

The abrupt transition from one regime of vocal fold oscillation to another is termed bifurcation, and under certain conditions, two alternative regimes of oscillation are equally probable (Berge *et al.*, 1986; Kaplan and Glass, 1995). Under these conditions the regime of oscillation may unpredictably jump, or bifurcate, from one to another with no change in the physiological parameters (subglottal pressure, adduction, vocal fold length, and so forth) underlying the vocal production (Wilden *et al.*, 1998). Figure 13 shows a typical example of regime instability within a sample. At the onset of the sample, the peak frequency was 1033 Hz. The peak frequency jumped to 1550 Hz at about 80 ms, jumped back to 1033 Hz at 90 ms, and jumped to 1378 Hz at about 115 ms. Chaotic oscillations with no apparent peak frequency were exhibited between 250 and about 310 ms during the sample, and were followed by the reemergence of the frequency peak at 1033 Hz. It is possible that abrupt transitions in the source features of natural vocalizations are also due to a similar phenomenon.

IV. DISCUSSION

The results show that in squirrel monkeys, as is the case for humans, voicing was produced by the vibrations of the

vocal folds. In the present study we never observed oscillations in the ventricular folds with any combination of subglottal pressure, vocal fold elongation, or adduction employed. Furthermore, we observed no changes in squirrel monkey voicing associated with the presence or absence of the ventricular folds. Though classic studies of the primate larynx (Starck and Schneider, 1960) suggested that the squirrel monkey exhibits a prominent vocal membrane projecting medially from the body of the vocal fold (see Fig. 3), we were unable to confirm the presence of the vocal membrane in our specimens (see Fig. 2). The video images obtained in the present study are consistent with the recent profile of the squirrel monkey larynx reported by Kahane *et al.* (2002), and the degree of anatomical variation in the shape of the primate larynx is presently unclear. In some species there is evidence of a polymorphism in the presence or absence of the vocal membrane (Riede *et al.*, 2000), and the emerging data suggests that different primate specimens show marked variation in the morphology of the vocal folds (Riede, 2002). Thus, strain or species differences may contribute to the morphological variation observed for the squirrel monkey vocal folds. Furthermore, differences in tissue preparation protocols have been shown to markedly alter the apparent geometry of the vocal folds (Tayama *et al.*, 2002). Some histologic processes may result in shrinkages as large as 30%, and it is possible that the mucosal layer of the vocal folds in the Starck and Schneider (1960) study may have experienced proportionally greater dehydration and shrinkage relative to that in the Kahane *et al.* (2002) study. Changes in histologi-

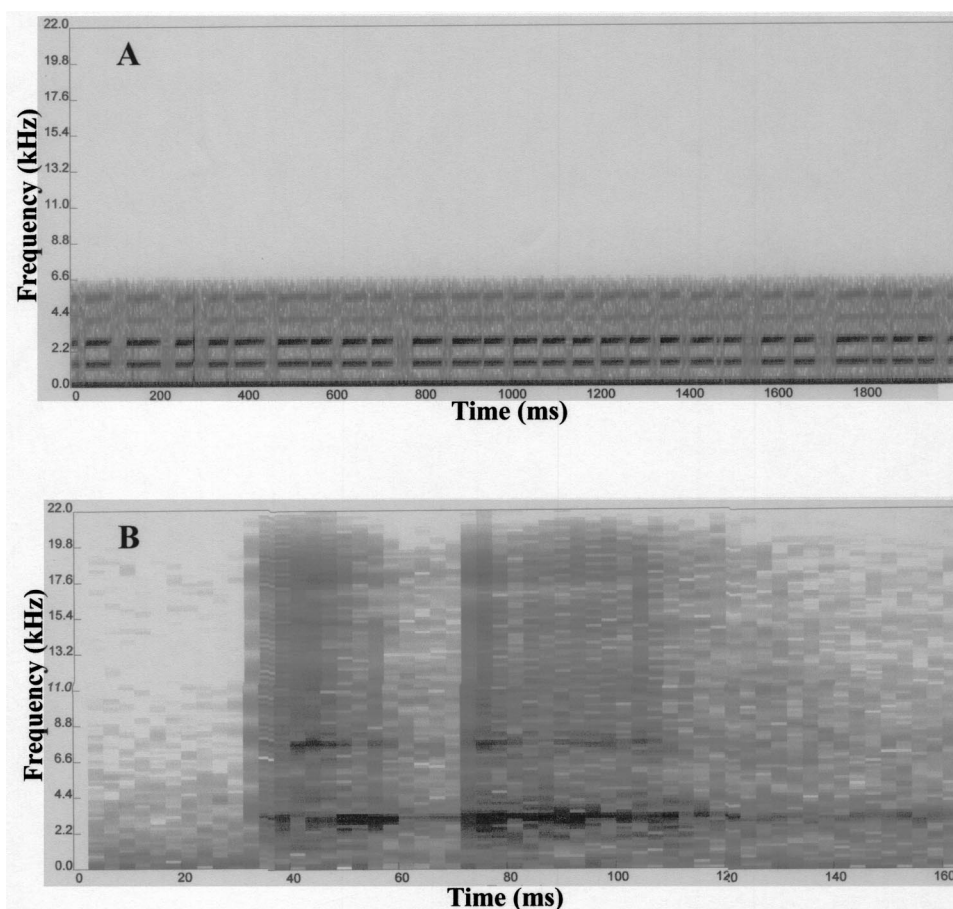


FIG. 10. (a) Sound sample recorded from an excised larynx exhibiting a sustained series of pulses. The typical interval between pulses was 10–15 ms. (b) An exemplar of a squirrel monkey churr exhibiting two pulses. The interval between the two pulses of the churr was 12 ms, and is close to the interval between the pulses shown in the excised larynx sample displayed in 10a.

cal techniques over the 40 years spanning these studies may account for the morphological variation observed for the squirrel monkey vocal folds. However, as noted above, there may be strain or species differences as well, and additional research will be needed to resolve these questions.

The basic characteristics of squirrel monkey phonation were consistent with studies in our laboratory conducted with canines. Phonation threshold pressures observed here for squirrel monkeys (5 to 8 cm-H₂O) are similar to the values observed for dogs (Alipour *et al.*, 1997), and measurements of the maximum glottal area and airflow are consistent with our findings from canines given the difference in scale. Compared to adult humans, or canines, the squirrel monkey larynx is small; the vocal folds are about 4–5 mm in length. During voicing the maximum area of the squirrel monkey glottis was about 3.5 mm², and the maximum airflow was about 200 ml/s. The squirrel monkey trachea and larynx are about $\frac{1}{8}$ the proportion of the canine tissues normally studied in excised larynx preparations (Berry *et al.*, 1996; Alipour *et al.*, 1997), and the performance parameters of the squirrel monkey larynx are in keeping with the size differences between these preparations.

Measurements of sound pressure level showed that the squirrel monkey larynx exhibited a dynamic range of 36 dB. The squirrel monkey dynamic range can be compared with that for humans and other primates. Many singers can achieve a range in variation of acoustic power of about 15 dB, and some highly skilled vocalists can achieve acoustic power variations approaching 30–40 dB (Titze, 1994). While

the squirrel monkey samples are not given at a fixed pitch, as is the case for singers, the data nevertheless suggests that, in spite of the small vocal fold length, and the limitations that this parameter would seem to impose on the upper end of the vocal intensity range, the squirrel monkey range in vocal intensity is in the same general ballpark as that for humans. Careful acoustic measurements of vocal amplitude (in which the influence of environmental acoustic has been controlled) have been reported for two species of Old World monkeys: blue monkeys (*Cercopithecus mitis*) and gray-cheeked mangabeys (*Lophocebus albigena*), and the data suggest that the dynamic range in vocal amplitude for blue monkeys and mangabeys is 31 and 33 dB, respectively (Brown, 1989). These values are close to the squirrel monkey data reported here.

The squirrel monkey larynx exhibited four different regimes of voicing including periodic phonation with overtones, biphonation, staccato phonation, and aperiodic oscillation. Each regime arose without any attempt to manipulate the larynx to adopt an “unnatural” regime of oscillation. Because the squirrel monkey larynx is comparatively small, we were sensitive to the possibility that we may inadvertently apply asymmetric forces to the left and right side of the arytenoid process and, as a consequence, artificially disrupt the coupling between the left and right vocal folds. This situation would heighten the likelihood of the expression of regimes of phonation that were an artifact of our methods, and not a natural property of the squirrel monkey’s voice range profile. To guard against this possibility, five larynges

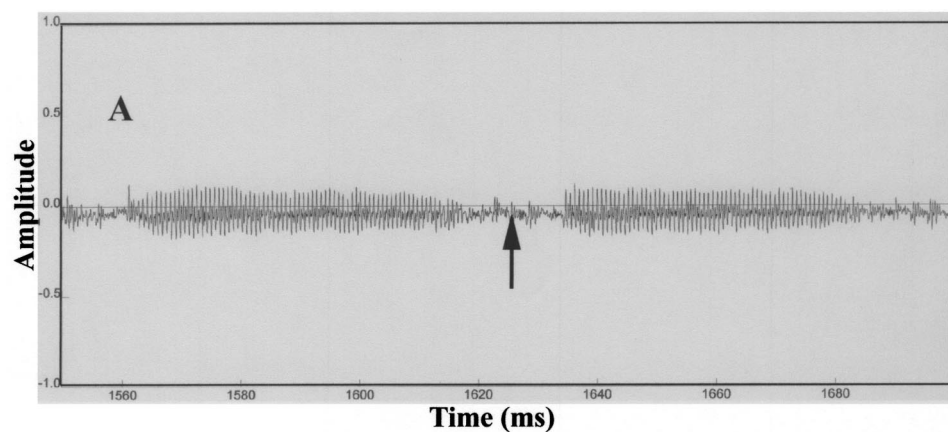


FIG. 11. (a) Waveform for two pulses recorded from the excised larynx [Fig. 10(a)]. (b) Waveform for the two pulses recorded from a churr vocalization [Fig. 10(b)]. The arrows denote the interval between pulses, and the waveforms show that voicing is continuous, and the amplitude of voicing waxes and wanes yielding the impressing of discrete pulses.

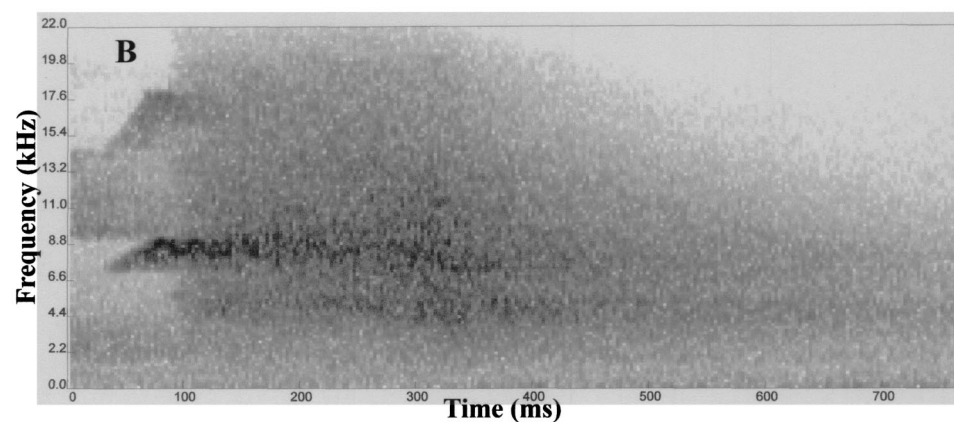
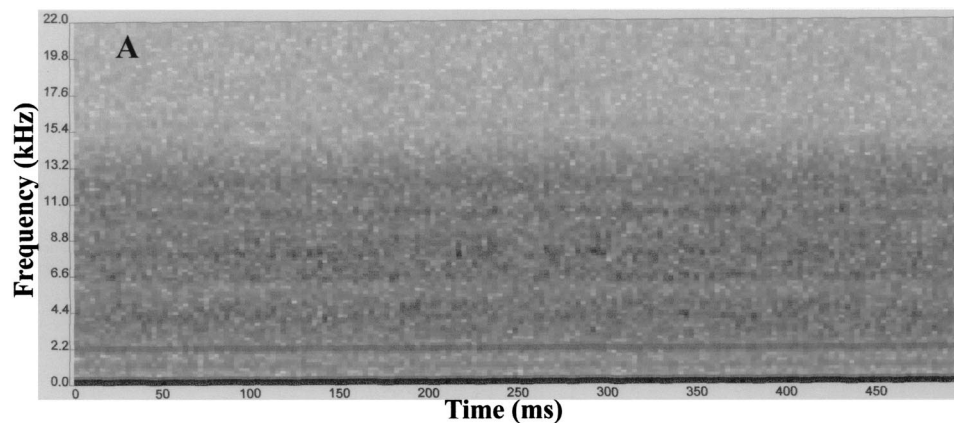
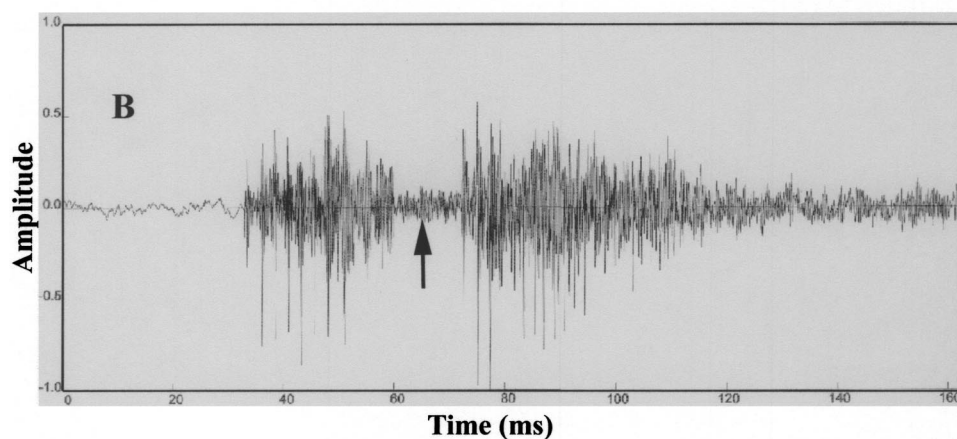


FIG. 12. (a) Sound sample recorded from an excised larynx exhibiting a chaotic pattern of vibrations of the vocal folds. A fundamental frequency is not discernable in the spectrum. (b) Exemplar of a squirrel monkey shriek. The call frequency spectrum is very broad in bandwidth, and evidence of a fundamental frequency is not apparent during the interval of 200–700 ms.

TABLE I. The relationship between voicing regimes and subglottal pressure.

	Subglottal pressure		
	Low (6–16 cm H ₂ O)	Moderate (18–28 cm H ₂ O)	High (30–40 cm H ₂ O)
Regime 1 (Periodic phonation)	Present	Present	Present
Regime 2 (Biphonation)	Absent	Present	Present
Regime 3 (Staccato phonation)	Present	Present	Rare
Regime 4 (Aperiodic phonation)	Absent	Present	Present

were phonated, in one control condition, with no attachments to the thyroid and arytenoid tissues, and clear examples of the voicing regimes described here were obtained. Thus, we propose that the voicing regimes observed here were not an artifact of our procedures, but instead were a natural property of the squirrel monkey voice profile. Furthermore, each regime of oscillation resembled the source features of one or more categories of calls found in the squirrel monkey vocal repertoire. This observation suggests that squirrel monkeys, and possibly other primates, electively select the regime of oscillation appropriate to the target utterance. If this is true, the motor control of the parameters of voicing that alter the regimes of oscillation may be an important topic for understanding primate communication.

Regime 1 was characterized by a single fundamental frequency accompanied by a series of harmonics of the fundamental. It was observed at low, moderate, and high subglottal pressures. It is possible that rather subtle variations in the shape of the glottal flow waveform are important for squirrel monkey communication. Compared to the normal human voice, the amplitudes of the harmonics shown in Fig. 8(a) were relatively high; this is indicated by the shallow spectral slope of only 7 dB per octave. A shallow spectral slope is associated with glottal flow waveforms characterized by a small open quotient, a large skewing quotient, or both (Titze, 1994). We predict that relative to the normative pa-

rameters for human vowel production, squirrel monkeys will exhibit glottal flow waveforms with attributes that have been selected to heighten call audibility. In this respect, we hypothesize that the squirrel monkey open quotient will be comparatively small. As the open quotient is reduced in duration, glottal power increases (Titze, 1994), and greater glottal power would heighten audibility. Furthermore, given a fixed glottal flow rate, the maximum number of glottal cycles generated for a fixed volume of airflow would increase as the open quotient is reduced. Hence, utterance duration, the number of glottal cycles per exhalation, and glottal power are all favored by relatively small open quotients. Thus, small open quotients may be a characteristic feature of the glottal airflow waveform for calls produced with regime 1 source characteristics.

Regime 2 was characterized by biphonation, and was observed at moderate and high subglottal pressures. Biphonation results when the left and right vocal folds are uncoupled, and the folds oscillate asynchronously and independently from each other. Asynchronous tension on the vocal folds associated with unilateral paralysis may result in biphonation of this type. Recent work has shown that in adult humans the vocal folds of healthy normal speakers may become uncoupled and oscillate independently as a natural phenomenon over a portion of their range (Tigges *et al.*, 1997). Thus, biphonation is not necessarily indicative of uni-

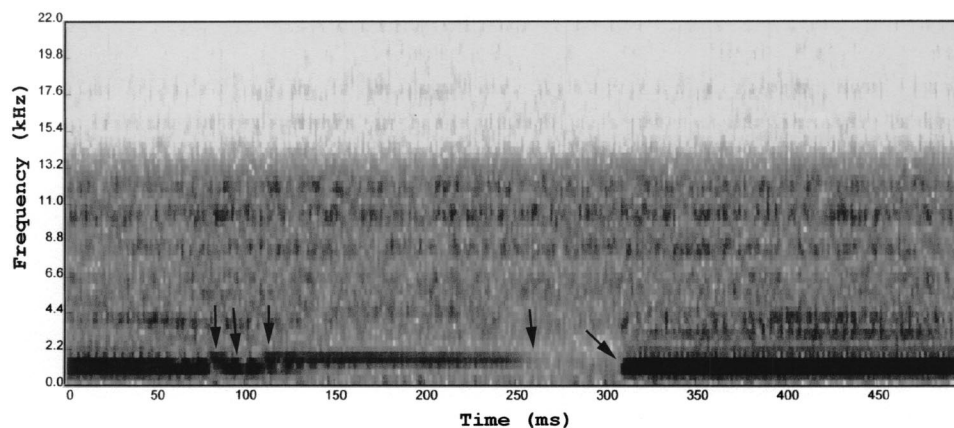


FIG. 13. Sound sample recorded from an excised larynx exhibiting spontaneous bifurcations between multiple attractors. The sound spectrogram is displayed in the top panel, the waveform in the middle panel, a segment of the waveform is displayed in the bottom left panel, and the spectrum is shown in the bottom right panel. The peak frequency was 1033 Hz at the onset of the sample. The frequency peak jumped to 1550 Hz at 80 ms, jumped to 1033 Hz at 90 ms, and bifurcated to 1378 Hz at 115 ms. A segment of chaotic oscillations was exhibited between 250 and 310 ms, and was followed by the reemergence of the frequency peak at 1033 Hz. The bifurcations are denoted by the arrows.

lateral paralysis or some other underlying pathology. There are a variety of other phenomena that may also give rise to biphonation. For example, the horizontal and vertical vibratory modes within a single fold may become desynchronized and result in biphonation (Berry *et al.*, 1994). Hence, additional research is necessary to identify the sources of biphonation in primate phonation. It should be emphasized that in the case of squirrel monkeys, it is possible that biphonation is purposefully exploited to produce an additional vocal category. It is unclear if different patterns of biphonation are under vocal control, and are of potential communicative significance, or alternatively if the patterns of variation produced during biphonation are random events and are insignificant communicatively. Studies of call perception are needed to address this question.

Regime 3 was characterized by staccato phonation. The regime was observed at low and moderate subglottal pressures, and was only rarely observed at high subglottal pressures. The human larynx is also capable of rapid repetitions of notes or elements. Vocalic imitations of the sound of machine gun fire, the bleat of a lamb, the whinny of a horse, as well as giggles and chuckles involve rapid reiterations of acoustic elements. In the musical literature, *trillo* is a staccato singing style, or vocal ornament, involving pulses of the same note repeated at rates ranging between about 2 to 12 Hz (Hakes *et al.*, 1990; Brown and Scherer, 1992). In staccato imitative sounds, as well as in *trillo*, phonation never ceases despite the percept of “silence” between the pulses (Hamlet and Palmer, 1974; Brown and Scherer, 1992), and, in this respect, both monkeys and humans are able to produce staccato phonation in which voicing is continuous, but the percept, at least for human listeners, is that of a stream of segmented discrete elements, pulses, or notes.

Brown and Scherer (1992) hypothesized that rapidly alternating patterns of abduction and adduction were necessary for the production of *trillo*. However, the discovery of *trillo*-like iterations of pulses in an excised larynx preparation raises the possibility that the oscillation in amplitude of the glottal pulses, observed in *trillo* and other examples of staccato phonation, is a passive property of the biomechanics of the vocal folds under certain parameters of subglottal pressure, adduction, and so forth, and is not necessarily due to active processes involving pulsed patterns of muscle activation. It is possible, of course, that different forms of staccato phonation are produced by different mechanisms, and further research would have to explore this possibility.

Giovanni *et al.* (1999) have shown that if the two vocal folds differed in their relative stiffness, then the frequency and amplitudes of oscillation of the vocal folds would differ, and may result in the development of a regime of oscillation in which the amplitude of the signal will wax and wane according to the nonlinear combination of the oscillations of the two folds. The amplitude oscillations of the samples reported by Giovanni and his associates resemble those observed here for the staccato voice samples. It is possible that a similar phenomenon underlies the trilled and pulsed calls issued by squirrel monkeys. Though at present we cannot rigorously specify how the larynx produces these source features, there is good reason to assume that the ability to

modify and control these source parameters is important for the squirrel monkey's vocal communication system, and this intriguing possibility merits additional exploration.

Regime 4 was characterized by chaotic (or aperiodic) oscillations of the vocal folds, and was observed at moderate and high subglottal pressures. Wilden *et al.* (1998) were the first to introduce the field of nonlinear dynamics to bioacoustical studies of animal vocalizations, and their work emphasized the possibility that chaotic vibrations of the vocal folds would yield sound spectrograms similar to those attributed to turbulence. Thus, the broad bandwidth, aperiodic character of many animal vocalizations may not be due to aerodynamic instabilities, but may rather be due to irregular regimes of vocal fold vibration. Though chaos results in noiselike regimes of vocal production, one or more of the following characteristics are typically associated with this phenomenon: the sound amplitude is high (subglottal pressure is high), the onset of the nonperiodic segment is abrupt and often preceded by a periodic segment, hints of a fundamental frequency may be apparent as a residual periodic structure (or periodic windows with harmonic structures) interspersed within chaotic segments, adjacent episodes of other nonlinear phenomena are common including frequency jumps, biphonation, subharmonics, and so forth (Wilden *et al.*, 1998). These attributes help distinguish between exemplars produced by turbulence from those produced by a chaos.

Under some conditions spontaneous bifurcations from one regime of oscillation to another regime were observed to occur without any change in voice control parameters (Fig. 13). If this phenomenon were prevalent over large segments of the voice profile range, vocalizers would have little ability to control the source features of an utterance, any intended modulation in fundamental frequency or amplitude could easily lead to unexpected bifurcations to completely different voicing regimes, and the risk of producing a call with “unintended” source features would be significant. Bifurcation diagrams (Berry *et al.*, 1996), for changes in subglottal pressure and symmetric variations in vocal fold length (or adduction), may reveal the relative area (and hence, probability) of the region exhibiting regime instability. Additional data are needed to determine the significance of this problem for squirrel monkey vocal communication.

In three previous studies that focused on the nonlinear behavior of vocal fold vibration, different regimes of oscillation were experimentally induced by asymmetrically manipulating the stiffness of the two vocal folds (Yanagi and McCaffery, 1992; Giovanni *et al.*, 1999), or the tension and length of the two folds (Berry *et al.*, 1996). Berry and his colleagues (1996) noted that only occasionally were bifurcations from periodic to nonperiodic regimes of oscillation observed for symmetric folds. In contrast, in the present study, the various regimes of oscillation occurred readily and spontaneously without any attempt to produce asymmetries in the stiffness, length, or tension of the vocal folds. In order to produce adequate acoustic power, the amplitude of oscillations of the squirrel monkey's vocal folds may tend to be relatively high. That is, to compensate for the comparatively small vocal fold length, squirrel monkeys may tend to vocal-

ize with high-amplitude oscillations, and one result of the heightened amplitude-to-length ratio may be an increase in the number of different regimes of oscillation. The present results are certainly consistent with this perspective. However, there may be other characteristics of the squirrel monkey's larynx that influence the number of different regimes of oscillation exhibited, and their respective characteristics. In this respect additional research is needed to explore the tissue characteristics of the squirrel monkey larynx, and the significance of these tissue characteristics to the biomechanics of voicing. The squirrel monkey produces some calls that cannot readily be attributed to the voicing regimes observed here, and it is possible that additional research will identify additional regimes of oscillation. In this regard, it is possible that primates may asymmetrically alter the stiffness or tension of the vocal folds, and asymmetrical changes may result in new regimes of oscillation.

We are aware of no studies where the excised larynx preparation has been adopted to the infant human larynx. However, the literature on the voice characteristics of infants and young children suggest that many of the phenomena reported here are prominent. Biphonation, frequency jumps, and harmonic doubling are found in the voicing patterns of both infants and young children (Lind, 1965; Sirviö and Michelsen, 1976; Robb and Saxman, 1988). Furthermore, measurements of intraoral and tracheal air pressure suggest that, relative to adults, children tend to vocalize at comparatively high pressure levels (Stathopoulos and Weismer, 1985; Stathopoulos and Sapienza, 1993), and, taken together, these findings suggest that human infants, like squirrel monkeys, may tend to generate comparatively high amplitude oscillations of the vocal folds, and the heightened amplitude-to-length ratio may be associated with an increase in the number of different regimes of vibration.

In this context, it appears likely that the expression of unstable regimes of phonation, or unintended bifurcations, may pose a serious impediment for the emergence of speech. As speech begins to develop, infants and young children apparently learn to predominantly use periodic phonation. How infants and young children develop the capacity to confine voicing to periodic phonation is unknown. As a starting point, it would be interesting to determine the relative size of the parameter space supporting periodic phonation compared to other regimes in a larynx with a short vocal fold length, and determine the amplitude-to-length ratios associated with bifurcations from one regime to another for both symmetric and experimentally induced asymmetries in the vocal folds.

Alipour, F., Scherer, R. C., and Finnegan, E. (1997). "Pressure-flow relationships during phonation as a function of adduction," *J. Voice* **11**, 187–194.
 Berge, P., Pompeau, Y., and Vidal, C. (1986). *Order Within Chaos* (Wiley, New York).
 Berry, D. A., Herzel, H., Titze, I. R., and Krischer, K. (1994). "Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions," *J. Acoust. Soc. Am.* **95**, 3595–3604.
 Berry, D. A., Herzel, H., Titze, I. R., and Story, B. H. (1996). "Bifurcations in excised larynx experiments," *J. Voice* **10**, 129–138.
 Boinski, S. (1991). "The coordination of spatial position: a field study of the vocal behavior of adult female squirrel monkeys," *Anim. Behav.* **41**, 89–102.

Boinski, S., and Mitchell, C. L. (1992). "Ecological and social factors affecting the vocal behavior of adult female squirrel monkeys," *Ethnology* **92**, 316–330.
 Boinski, S., and Mitchell, C. L. (1995). "Wild squirrel monkey (*Saimiri sciureus*) "caregiver" calls: Contexts and acoustic structure," *Am. J. Primatol.* **35**, 129–137.
 Boinski, S., and Newman, J. D. (1988). "Preliminary observations on squirrel monkey (*Saimiri oerstedii*) vocalizations in Costa Rica," *Am. J. Primatol.* **14**, 329–343.
 Brown, C. H. (1989). "The measurement of vocal amplitude and vocal radiation pattern in blue monkeys and grey-cheeked mangabeys," *Bioacoustics* **1**, 253–271.
 Brown, C. H., and Cannito, M. P. (1995a). "Modes of vocal variation in Sykes's monkey (*Cercopithecus albigena*) squeals," *J. Comp. Psychol.* **109**, 398–415.
 Brown, C. H., and Cannito, M. P. (1995b). "Articulated and inflected primate vocalizations: Developing animal models of speech," in *Disorders of Motor Speech: Assessment, Treatment, and Clinical Consideration*, edited by D. Robbin and D. Beukleman (Brookes, Baltimore), pp. 43–63.
 Brown, L. R., and Scherer, R. C. (1992). "Laryngeal adduction in trillo," *J. Voice* **6**, 27–35.
 Giovanni, A., Ouaknine, M., Guelfucci, B., Yu, P., Zanaret, M., and Triglia, J.-M. (1999). "Nonlinear behavior of vocal fold vibration: the role of coupling between the folds," *J. Voice* **13**, 465–476.
 Hakes, J., Doherty, E., and Shipp, T. (1990). "Trillo rates exhibited by professional early music singers," *J. Voice* **4**, 305–308.
 Hamlet, S. L., and Palmer, J. M. (1974). "Investigation of laryngeal trills using the transmission of ultrasound through the larynx," *Folia Phoniatr.* **26**, 362–377.
 Hast, M. H. (1983). "Comparative anatomy of the larynx: Evolution and function," in *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control*, edited by I. R. Titze and R. C. Scherer (Denver Center for the Performing Arts, Denver), pp. 3–14.
 Hirano, M., and Sato, K. (1993). *Histological Color Atlas of the Human Larynx* (Singular, San Diego).
 Hirano, M., Kurita, S., and Nakashima, T. (1983). "Growth, Development, and Aging of Human Vocal Folds," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego), pp. 22–43.
 Jürgens, U. (1976a). "Projections from the cortical larynx area in squirrel monkey," *Exp. Brain Res.* **25**, 401–411.
 Jürgens, U. (1976b). "Reinforcing concomitants of electrically elicited vocalizations," *Exp. Brain Res.* **26**, 203–214.
 Jürgens, U. (1988). "Central control of monkey calls," in *Primate Vocal Communication*, edited by D. Todt, P. Geodeking, and D. Symmes (Springer-Verlag, Berlin), pp. 162–167.
 Jürgens, U. (1998). "Neuronal control of mammalian vocalization, with special reference to the squirrel monkey," *Naturwissenschaften* **85**(8), 376–388.
 Jürgens, U., and Pratt, R. (1979). "Role of the periaqueductal grey in vocal expression of emotion," *Exp. Brain Res.* **167**, 367–378.
 Jürgens, U., and Zwirner, P. (2000). "Individual hemispheric asymmetry in vocal fold control of the squirrel monkey," *Behav. Brain Res.* **109**(2), 213–217.
 Jürgens, U., Hast, M., and Pratt, R. (1978). "Effects of laryngeal nerve transection on squirrel monkey calls," *J. Comp. Physiol.* **123**, 23–29.
 Jürgens, U., Kirzinger, A., and von Cramon, D. (1982). "The effects of deep-reaching lesions in the cortical face area of phonation. A combined case report and experimental study," *Cortex* **18**, 125–140.
 Kahane, J., Becker, M., and Brown, C. (2002). "Gross anatomy and histology of the squirrel monkey vocal tract and larynx," presented at the 19th Congress of the International Primatological Society, Beijing, China (unpublished).
 Kaplan, D., and Glass, L. (1995). *Understanding Nonlinear Dynamics* (Springer, Berlin).
 Kirzinger, A. (1985). "Cerebellar lesion effects on vocalization of the squirrel monkey," *Behav. Brain Res.* **16**, 177–181.
 Kirzinger, A., and Jürgens, U. (1985). "The effects of brain stem lesions on vocalization in the squirrel monkey," *Brain Res.* **358**, 150–162.
 Lind, J. (1965). *Newborn Infant Cry* (Almqvist and Wiksells Boktrycken, Uppsala).
 Mergell, P., Fitch, W. T., and Herzel, H. (1999). "Modeling the role of

- nonhuman vocal membranes in phonation," J. Acoust. Soc. Am. **105**, 2020–2028.
- Newman, J. D. (1985). "Squirrel monkey communication," in *Handbook of Squirrel Monkey Research*, edited by L. A. Rosenblum and C. L. Coe (Plenum, New York), pp. 99–126.
- Riede, T. (2002). Private communication.
- Riede, T., Herzel, H., Mehald, D., Seider, W., Trumler, E., Böhme, G., and Tembrock, G. (2000). "Nonlinear phenomena in the natural howling of a dog-wolf mix," J. Acoust. Soc. Am. **108**, 1435–1442.
- Riede, T., and Stolle-Malorny, A. (1999). "The vocal change of a kitten with craniocerebellar trauma—a case study," Bioacoustics **10**, 131–141.
- Riede, T., Wilden, I., and Tembrock, G. (1997). "Subharmonics, biphonations, and frequency jumps—common components of mammalian vocalizations or indicators for disorders," Z. Säugetierkunde **62** (Suppl. 2), 198–203.
- Robb, M. P., and Sařman, J. (1988). "Acoustic observations in young children's non-cry vocalizations," J. Acoust. Soc. Am. **83**, 1876–1882.
- Schott, D. (1975). "Quantitative analysis of the vocal repertoire of squirrel monkeys (*Saimiri sciureus*)," Z. Tierpsychol. **38**, 225–250.
- Sirviö, P., and Michelsson, K. (1976). "Sound-spectrographic cry analysis of normal and abnormal newborn infants," Folia Phoniatr. **28**, 161–173.
- Smith, M., and Gray, S. D. (1994). "Laryngeal framework surgery in children," NCVS Status and progress Report-5, University of Iowa, Iowa City, Iowa, November 1993, pp. 91–98.
- Smith, H. J., Newman, J. D., and Symmes, D. (1982). "Vocal concomitants of affiliative behavior in squirrel monkeys (*Saimiri sciureus*)," in *Primate Communication*, edited by C. T. Snowdon, C. H. Brown, and M. R. Petersen (Cambridge U.P., Cambridge), pp. 30–49.
- Starck, D., and Schneider, R. (1960). "Respirationsorgane," in *Primatologia III/2*, edited by H. Hofer, A. H. Schultz, and D. Starck (S Karger, Basel).
- Stathopoulos, E. T., and Sapienza, C. (1993). "Respiratory and laryngeal measures of children during vocal intensity variation," J. Acoust. Soc. Am. **94**, 2531–2543.
- Stathopoulos, E. T., and Weismer, G. (1985). "Oral airflow and intraoral air pressure: A comparative study of children, youths and adults," Folia Phoniatr. **37**, 152–159.
- Tayama, N., Chan, R. W., Kaga, K., and Titze, I. R. (2002). "Functional definitions of vocal fold geometry for laryngeal biomechanical modeling," Ann. Otol. Rhinol. Laryngol. **111**, 83–92.
- Thoms, G., and Jürgens, U. (1981). "Role of the internal laryngeal nerve in phonation: an experimental study in squirrel monkey," Exp. Neurol. **74**, 187–203.
- Tigges, M., Mergell, P., Herzel, H., Wittenberg, T., and Eysholdt, U. (1997). "Observation and modeling of glottal biphonation," Acustica **83**, 707–714.
- Titze, I. R. (1994). *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ).
- Titze, I. R., Baken, R., and Herzel, H. (1993). "Evidence of chaos in vocal fold vibration," in *Vocal Fold Physiology: Frontiers in Basic Science*, edited by I. Titze (Singular, San Diego), pp. 143–188.
- Wilden, I., Herzel, H., Peters, G., and Tembrock, G. (1998). "Subharmonics, biphonation, and deterministic chaos in mammal vocalizations," Bioacoustics **9**, 171–196.
- Winter, P. (1972). "Observations of the vocal behavior of free-ranging squirrel monkeys," Z. Tierpsychol. **31**, 1–7.
- Winter, P., Ploog, P., and Latta, J. (1966). "Vocal repertoire of the squirrel monkey (*Saimiri sciureus*), its analysis and significance," Exp. Brain Res. **1**, 359–384.
- Yanagi, E., and McCaffery, T. V. (1992). "Study of vibratory pattern of the vocal folds in the excised canine larynx," Arch. Otolaryngol. Head Neck Surg. **118**, 30–36.

Variation in chick-a-dee calls of a Carolina chickadee population, *Poecile carolinensis*: Identity and redundancy within note types

Todd M. Freeberg,^{a)} Jeffrey R. Lucas, and Barbara Clucas^{b)}

Department of Biological Sciences, Lilly Hall, Purdue University, West Lafayette, Indiana 47907

(Received 21 February 2002; accepted for publication 14 January 2003)

Chick-a-dee calls of chickadee species are structurally complex because calls possess a rudimentary syntax governing the ordering of their different note types. Chick-a-dee calls were recorded in an aviary from female and male birds from two field sites. This paper reports sources of variation of acoustical parameters of notes in these calls. There were significant sex and microgeographic differences in some of the measured parameters of the notes in the calls. In addition, the syntax of the call itself influenced characteristics of each of the notes. For example, calls with many introductory notes began with a note of higher frequency and longer duration, relative to calls with few introductory notes. Furthermore, the number of introductory notes influenced frequency and duration components of notes later in the call. Thus, single notes are predictive of the note composition of the signaler's call. This suggests that a receiver might gain the meaning in the call even if it hears only part of the call. Further, single notes within these complex calls can contain information enabling receivers to predict the sex of the signaler, and whether it is from the local population. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1559175]

PACS numbers: 43.80.Ka, 43.80.Ev, 43.80.Lb [WA]

I. INTRODUCTION

There are several potential conflicts between various functions of the vocal signals of animals. For example, the need to generate signals that effectively transmit information about individual identity will possibly limit the amount of song or vocal sharing that would facilitate the maintenance of group or population markers. Furthermore, when information is being transferred by vocal signals, it is often necessary for the signaler to provide redundant information, particularly as the noise level in the environment increases (Bradbury and Vehrencamp, 1998; Wiley, 1983). Redundancy of information can be achieved by using other signal modalities in parallel with vocal signals, by repeating the same or similar vocal signals, or by providing the same information in different ways within the same vocal signal (Bateson, 1968; Endler, 2000; Hauser, 1996; Smith, 1977; Uetz, 2000). This latter notion could be facilitated by the use of different note, syllable, or phrase types within a vocal system that were probabilistically associated with one another, such that the detection of one note type by a receiver might allow it to predict, for example, other note types in the signal to follow. As with the trade-offs between individual and group identification in vocal signals, redundancy of information within signals may limit the amounts of different information regarding identity or external referents that could be conveyed.

The chick-a-dee call system in avian species of the ge-

nus *Poecile* is produced in a variety of primarily social contexts (Ficken *et al.*, 1978; Hailman, 1989; Smith, 1972, 1991). The chick-a-dee call has a high acoustic variability with a rudimentary syntax that governs the ordering of its note types (Hailman, 1989). In black-capped chickadees, *Poecile atricapillus*, for example, there are four note types (A, B, C, and D) that follow a general A-B-C-D ordering. Any given note type may or may not be present in a given chick-a-dee call and, if present, can occur multiple times. These calls are said to be "combinatorial" in that they follow a simple and relatively fixed syntax but nonetheless could vary enormously in the number of different call types that could be produced (Hailman and Ficken, 1986; Hailman *et al.*, 1987). The chick-a-dee call system of Carolina chickadees, *P. carolinensis*, also obeys these simple syntactic rules. The chick-a-dee call of Carolina chickadees typically begins with relatively high-frequency A notes ["high-see" and related notes in Smith (1972); see also Hailman (1989)], occasionally followed by a single B note (variants of "high-tee" note) or by more structurally complex C notes ("chick" notes), and ends with lower-frequency D notes ("dee" notes). The numbers and even presence of each note type is variable in these calls, although calls virtually always follow an "A-B-D" or "A-C-D" ordering [Fig. 1(a)]. This variable, yet rule-governed, system of notes in chick-a-dee calls means that the call system theoretically could convey a large amount of information (Hailman *et al.*, 1985, 1987).

The presence of identity information in the chick-a-dee call system has primarily been studied in black-capped chickadees. Mammen and Nowicki (1981) found that frequency and duration properties of the first D note in calls, as well as characteristics of the complete call such as call duration and number of D notes (which relate to the "syntax" or note composition of the call), varied among different

^{a)}Author to whom correspondence should be addressed. Department of Psychology, Austin Peay Building 303A, University of Tennessee, Knoxville, TN 37996. Electronic mail: tfreeber@utk.edu

^{b)}Current address: Animal Behavior Graduate Group and Department of Psychology, University of California at Davis, One Shields Avenue, Davis, CA 95616. Electronic mail: baclucas@ucdavis.edu

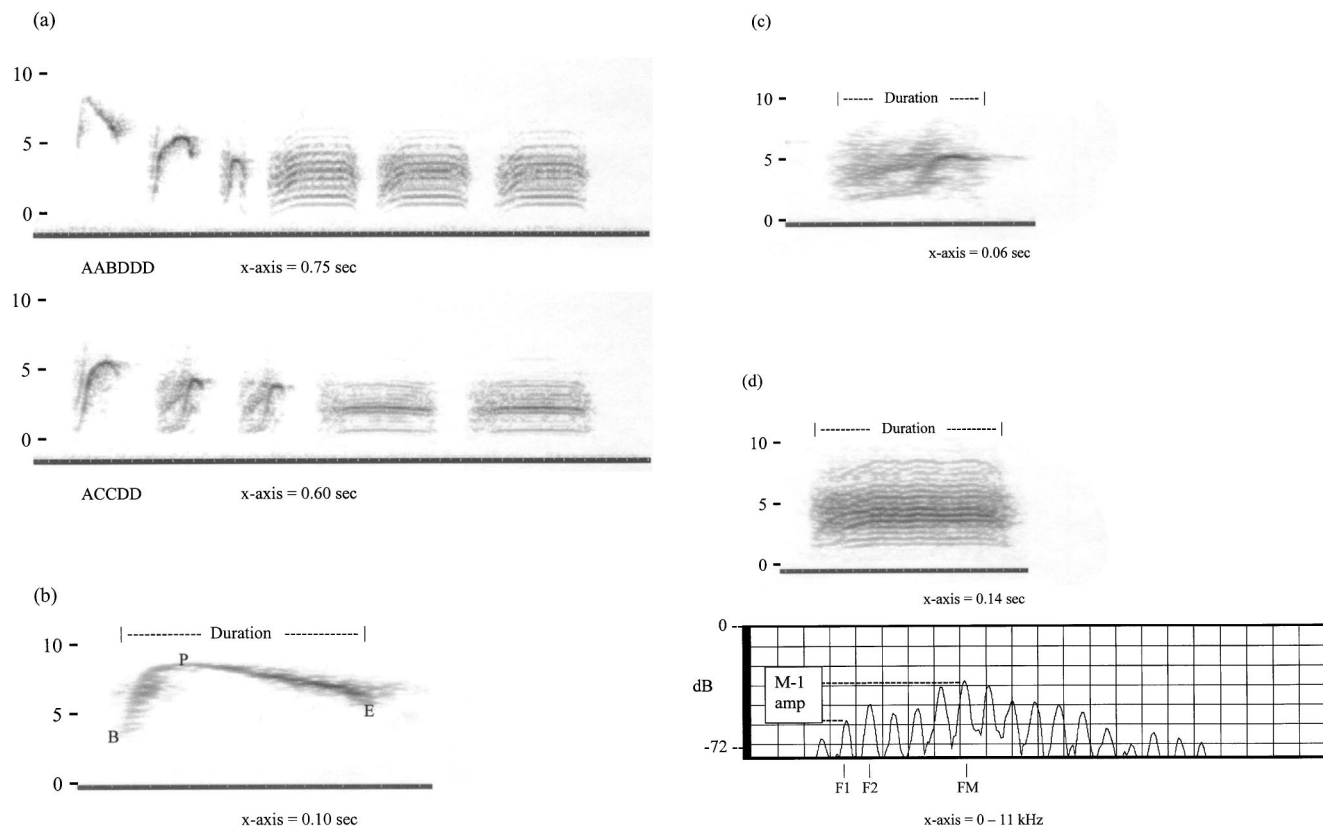


FIG. 1. Examples of note types comprising chick-a-dee calls of Carolina chickadees, and illustration of some of the parameters measured in the study. For each example, frequency (kHz) is depicted on the Y axis. (a) Sonograms of chick-a-dee calls containing 2 A, 1 B, and 3 D notes (top figure) and 1 A, 2 C, and 2 D notes (bottom figure). The X axis in the top figure is 750 ms and in the bottom figure is 600 ms. (b) Example of some A note measures. B = beginning frequency. P = peak frequency. E = end frequency. (c) Example of C note measure. (d) Example of some D note measures. The bottom figure in (d) is an amplitude spectrum with relative amplitude (dB) on the Y axis and frequency (Hz) on the X axis. F1 = frequency of the first peak. F2 = frequency of the second peak. FM = frequency of the maximum amplitude peak. M-1 amp = the relative amplitude difference between the maximum peak amplitude and the first peak amplitude.

flocks. That study did not consider the possibility that acoustic properties of notes might be affected by the note composition of the entire call. Therefore, a potential confound between note composition and note properties may bias conclusions drawn about flock differences. We consider this possibility here using the calls of Carolina chickadees.

To our knowledge, no analysis of variation in note parameters within calls in Carolina chickadees exists comparable to the work of Mammen and Nowicki (1981) on black-capped chickadees. Some recent work has focused on the production and perception of chick-a-dee calls in different situations, including testing whether different note compositions in calls correspond with different social or physical contexts. For example, Carolina chickadees signalers produce chick-a-dee calls at higher rates during times of temporary energetic stress (Lucas *et al.*, 1999). Further, Carolina chickadees respond differently to field playbacks of chick-a-dee calls that vary in note composition, suggesting that calls with different note compositions might convey different meanings to receivers (Freeberg and Lucas, 2002). Finally, in other chickadee species, different usage of note types in different contexts has been documented (Mexican chickadees, *Poecile sclateri*, Ficken *et al.*, 1994; mountain chickadees, *P. gambeli*, Gaddis, 1985).

Here we present data on sources of acoustic variation in

several characteristics of the notes of calls we recorded from female and male Carolina chickadees from two capture sites (separated from one another by at most 5 km). Geographic variation in vocalizations has long been studied in songbirds (e.g., Baker and Cunningham, 1985; Kroodsma and Miller, 1996; Mundinger, 1982), and sex differences in acoustic characteristics of the same vocal types have been demonstrated in some species (Ballintijn and ten Cate, 1997; Okanoyak, 1993; Yamaguchi, 1998a). One of our goals was to establish a first step towards uncovering micro-level variation (between sites close in proximity, and between birds from the same sites that differ in sex) in the structures of this call. We asked whether characteristics of individual note types in chick-a-dee calls (after syntax effects are statistically removed) might provide information about the identity of the signaler. We characterized notes in our analyses using a number of frequency and amplitude measurements, in addition to whole-note properties such as duration, internote interval, entropy, continuity, and modulation (see "C notes" and "D notes" below).

In addition to testing for this micro-level of geographic variation, we sought to determine whether properties of the first A, C, and D notes in a call predicted (or were predicted by) syntactic characteristics of the rest of the call. In chick-a-dee calls, a note can occur more than once in a string of the

same note type—for example, a call containing A, C, and D notes can vary from ACD to AAAAAACD to AC-CCCCCD to ACDDDDDDDD to AACCCDDDDDD, and so on. It has been hypothesized that each of the four note types may encode different information about signaler motivation to approach or withdraw from an area (e.g., Ficken *et al.*, 1994; Freeberg and Lucas, 2002; Gaddis, 1985). Further, the number of notes may encode information about the intensity of those tendencies (Hailman *et al.*, 1985, 1987; Smith, 1972). Thus, an AAACCC call may signal the same general message as an AAACCCCCCCC call, but the latter call would indicate a higher intensity on the part of the signaler for the particular message conveyed by C notes. In the present study, we sought to test whether the properties of certain notes could predict other notes in the call. If so, a receiver might be able to determine probabilistically a call's composition of notes, even if the receiver were to hear only part of the call. This would suggest an interesting form of redundancy of information in these structurally complex calls.

II. METHODS

A. Experimental animals and recording of chick-a-dee calls

We captured individual Carolina chickadees from flocks at the Ross Biological Reserve (hereafter, Ross) and the Martell Forest Reserve (hereafter, Martell) in West Lafayette, IN. Ross and Martell females and males were captured for recording from late fall to early spring, from 1997 to 1999. We tried to capture only adult birds [aged according to Pyle (1997)], and time of year for recording birds was balanced across sex and microgeographic capture site (Table I). All birds used in this study were banded with unique combinations of colored leg rings to permit individual identification. In addition, we had banded a number of the birds in months or years prior to their use in this study and this served as an aid for aging these previously banded birds. Birds were captured using baited seed stands with Potter (treadle) traps. After capture, birds were placed in individual 1-m³ cages in the laboratory. All birds were maintained with mixed seeds, grit, shredded carrot, one to three mealworms, and fresh vitamin-treated water daily. Lights in the aviary were set to the natural light/dark cycle that the birds would experience in the wild, given the time of year when the birds were captured and recorded. After recording (or, for some birds, after their use in a different study), birds were released at their site of capture.

We recorded all the birds in the same 1-m³ cage, to standardize the recording conditions across birds, in a room whose walls were made of acoustic tiling covered with polyurethane foam padding to minimize reverberation. All birds were recorded within a week of capture. During recording, chickadees could hear chickadees housed in the same room and in adjoining rooms. We recorded chick-a-dee calls with Saul Mineroff directional electret microphones, placed within 1 m of the bird, on Maxell XLII tape using a Marantz PMD 222 portable cassette recorder.

We attempted to record at least ten chick-a-dee calls

TABLE I. Capture site, sex, age, and recording date of Carolina Chickadees. AHY = "After Hatch Year" (see Pyle, 1997), indicating a bird that had been hatched at least as early as the summer of the year prior to recording. ASY = "After Second Year," indicating a bird that had been hatched at least as early as two summers before recording. U = "Unknown," indicating a bird we were unable to age successfully, and so could not determine whether it was older than HY ("Hatch Year").

Capture site	Sex	Bird no.	Age	Recording date
Ross	F	5	AHY	12/97
Ross	F	7	AHY	03/98
Ross	F	11	AHY	10/99
Ross	F	14	AHY	09/99
Ross	F	16	U	10/99
Ross	F	22	AHY	10/99
Ross	M	1	AHY	11/97
Ross	M	9	AHY	04/98
Ross	M	15	ASY	12/98
Ross	M	19	AHY	02/99
Ross	M	20	AHY	09/98
Martell	F	2	AHY	11/97
Martell	F	4	AHY	11/97
Martell	F	6	AHY	02/98
Martell	F	10	AHY	09/98
Martell	F	13	ASY	02/99
Martell	F	17	AHY	09/98
Martell	F	21	AHY	12/98
Martell	F	23	U	12/98
Martell	F	25	AHY	02/99
Martell	M	3	AHY	11/97
Martell	M	8	AHY	03/98
Martell	M	12	AHY	09/98
Martell	M	18	U	12/98
Martell	M	24	AHY	02/99

from each individual for later analyses. We obtained an average of 8.7 (range 3–15) calls from six females and an average of 12.6 (range 10–20) calls from five males from the Ross capture site, and an average of 10.6 (range 3–20) calls from nine females and an average of 11.8 (range 10–18) calls from five males from the Martell capture site. In our population, sex is strongly predicted by wing chord, the length of the bird's unflattened wing from the bend of the wing to the end of the longest primary feather [the cutoff between the sexes in our area is 62 mm, with males being larger than females (Thirakhupt, 1985)]—we used this size criterion to determine sex in this study.

B. Analyses of chick-a-dee calls

Recorded calls were digitized using the Cool Edit sound analysis program (version 96 run on the Windows 95 platform), with a sampling rate of 22 050 at 16-bit resolution. We measured several aspects of the notes in the calls, some of which were based on measures reported in earlier studies of black-capped chickadees (Mammen and Nowicki, 1981; Nowicki, 1989; Nowicki and Nelson, 1990). The first two note types (A and B notes) tend to be frequency modulated, relatively pure-tone whistled notes. We measured acoustic parameters of all of the introductory whistled notes in each call, although most of our analyses below focused on only the first introductory note (in this north-central Indiana population, this is virtually always an A note, rather than a B note, and, in the sample here, all were A notes).

We also measured properties of all of the C notes and the first, second, and last D notes within calls containing one or more of each note type, although, as was the case with A notes, most of our analyses concentrate on parameters of the first C note and first D note in calls. We used different acoustic parameters for the three different note types (A, C, and D) because the notes are structurally quite distinct (see Fig. 1).

1. A notes

Six parameters were used to characterize A notes [Fig. 1(b)]: *duration* (ms), *beginning frequency* (at the onset of the note, Hz), *peak frequency* (highest frequency in the note, Hz), *end frequency* (at the offset of the note, Hz), *position of peak amplitude* ([time at peak amplitude-note start time]/note duration), and *relative shape* ([time at peak frequency-note start time]/note duration).

2. C notes

Six parameters were used to characterize C notes [Fig. 1(c)]: *duration* (ms), *peak frequency* (Hz), *entropy*, *continuity*, *modulation*, and *pitch* (Hz). All of these parameters were obtained by using Sound Analysis [Version 1.4; see Tchernichovski *et al.* (2000)]. *Peak frequency* is a frequency measure derived from the spectrogram. *Entropy* provides a measure of the randomness or amount of “noise” in the waveform of a note; white noise, which by definition is completely random, has an entropy of 1, and a pure tone, which is by definition completely nonrandom, has an entropy of 0. *Continuity* is a measure of the extent to which frequency contours in a note type are stable across consecutive sampling windows. As Tchernichovski *et al.* (2000) note, continuity and entropy are correlated but not perfectly so—adding unbroken harmonics to a sound will not change continuity but will cause an increase in entropy. *Modulation* provides a measure of the relative frequency changes within the note. *Pitch* indicates the fundamental frequency of the note.

3. D notes

Twelve parameters were used to characterize D notes [Fig. 1(d)]. Six of these 12 parameters were whole-note characteristics: *duration* (ms), *internote interval* preceding the D note (ms), *entropy*, *continuity*, *modulation*, and *pitch* (the last four were obtained using Sound Analysis, described above). The other six parameters were characteristics of the frequencies and amplitudes of the stack of overtone-like structures that make the note. Frequency parameters (all Hz) were *freq max peak*, the frequency of the overtone with the maximum amplitude; *freq first peak*, the frequency of the lowest overtone within 30 dB of the maximum amplitude peak; and *freq second peak*, the frequency of the second overtone within 30 dB of the maximum amplitude peak. This 30-dB criterion for first peak and second peak measures is based on Nowicki (1989). Relative amplitude parameters were *position of maximal amplitude* ([time at peak amplitude-note start time]/note duration), *max to first peak amplitude difference* (the difference in amplitude from the max peak to the first peak

measured at 50% of the duration of the note), and 75%–50% *amplitude modulation* (the average of measures for first, second, and max peaks measured from 50% to 75% of the duration of the note).

C. Statistical analyses

The bulk of the analysis presented here focuses on the first A, first C, and first D notes in calls. However, to provide some context for the structure of these notes relative to the rest of the notes in the calls, we measured A and C note parameters for every A and C note in the calls, and D note parameters for the first, second, and last D note (when more than one D note occurred). We report general trends within strings of notes below, but carried out statistical tests only for the first A, C, and D notes in strings of A, C, and D notes, respectively. Since calls vary in the number of notes present within a string of the same note type (one of the components of the rudimentary syntax of the chick-a-dee call system), the power we have to detect syntax-specific note properties will obviously be highest with a detailed analysis of the first note in a sequence. Additionally, we analyzed in detail parameters for only these first notes of strings because there is a precedent for analyzing the first note in a series of the same note type (e.g., Mammen and Nowicki, 1981; Nowicki, 1989), and because one of the goals of the study was to determine whether information in the first note of a series of notes was predictive of the other notes in the call, as well as of sex or capture site.

We evaluated the source of variation in the properties of the chick-a-dee call in three ways. First, we ran a factor analysis of the parameters used to describe each note [PROC FACTOR; SAS Institute (1990)]. We then performed repeated-measures ANCOVA on the first factor for the first A note and first C note, and on the first two factors for the first D note in each call [PROC MIXED; SAS Institute (1990)]. Five independent variables were used in each ANCOVA: sex, capture site, number of A + B notes, number of C notes, and number of D notes in the call. Two-way interactions were added to the model and removed in order of increasing *F* statistic until all remaining interactions were significant ($\alpha=0.05$). For analysis of A notes, *duration* was log-transformed, and for analyses of D notes, *internote interval*, *pitch*, and *freq first peak* were log-transformed, to normalize the residuals of the ANCOVA models.

The factor scores were generally a robust index of the properties of a note, but they obviously did not explain all the variability in our measured note parameters. To ensure that no parameter-specific trends were missed, we also ran separate repeated-measures ANCOVAs for each measured note variable to determine the extent to which a bird's sex or capture site influenced specific properties of notes of the calls. Because we measured several variables for each of the notes (6 for A and C notes, and 12 for D notes), we corrected for multiple tests in these significance tests of single parameters using a sequential Bonferroni adjustment (Rice, 1989). The adjusted α value is provided with each of the reported significant tests. We carried out these sequential Bonferroni adjustments separately for sex and capture site.

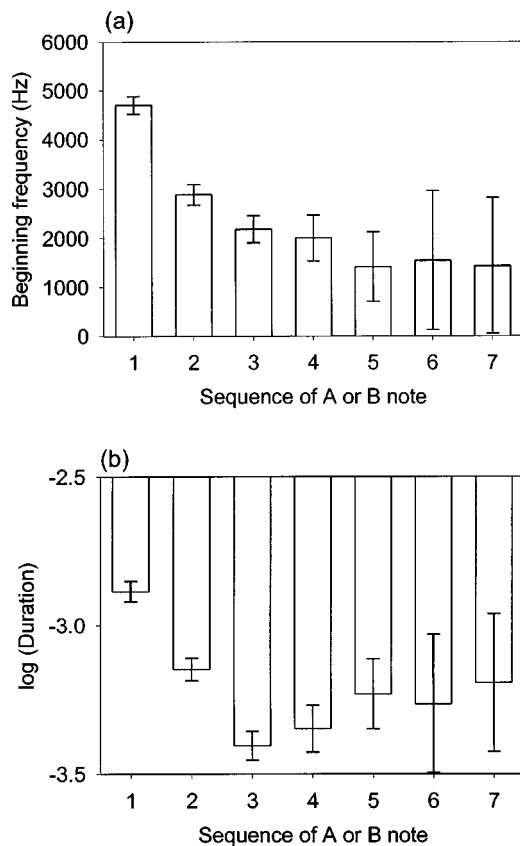


FIG. 2. Frequency and duration (log-transformed) characteristics of A notes within a string of A notes. Data are illustrated as means \pm s.e.

III. RESULTS

In total, we analyzed 269 calls and 674 notes. Martell calls ($N=154$) contained on average (\pm s.d.) 2.1 ± 1.1 A + B notes (range in analyzed calls 0–8), 0.4 ± 0.9 C notes (range 0–6), and 3.6 ± 2.2 D notes (range 1–14). Ross calls ($N=115$) contained on average 2.4 ± 1.2 A + B notes (range 0–8), 0.2 ± 0.6 C notes (range 0–3), and 3.1 ± 1.9 D notes (range 1–12).

A. A notes

To provide some context for understanding variation in note structure between calls, it is important to realize that strings of introductory whistled notes (A and B notes) tend to start with high-frequency, long-duration notes. Subsequent notes show a decrease in both beginning frequency ($F_{6,43} = 21.1, p < 0.0001$) and duration (log-transformed, $F_{6,43} = 17.3, p < 0.0001$), as is generally the case with the chick-a-dee call in *Poecile* species (Fig. 2; see also Hailman, 1989). These trends are similar between sexes and between populations. This is illustrated by a lack of significant interaction between sex and the position of a note in a sequence of A/B notes (frequency: $F_{4,39} = 0.22, p = 0.92$; duration: $F_{4,58} = 0.59, p = 0.67$), and no significant interaction between population and note position (frequency: $F_{4,58} = 0.49, p = 0.74$; duration: $F_{4,58} = 2.4, p = 0.06$). Note that multiple notes of the same type are always clustered together in a string, and are never interdigitated with other note types in the call.

A factor analysis of the first A note in calls indicates that

TABLE II. Factor loadings for the first two factors of a factor analysis of the properties of the first A note in the chick-a-dee call. See text for definitions of properties.

Parameter	Factor 1	Factor 2
Log (Duration)	0.416	−0.312
Position of peak amplitude	0.584	−0.292
Beginning frequency	0.805	−0.260
Peak frequency	0.882	−0.023
End frequency	0.646	0.688
Shape	−0.629	0.530
Eigenvalue	2.75	1.01
Variance explained	46%	17%

the six parameters used to characterize the notes are all correlated with one another (Table II). Factor 1 (46% of the variance explained) is positively correlated with note duration, position of peak amplitude, and note frequencies (beginning, peak, and end), and is negatively correlated with note shape. Thus, longer notes (compared to shorter notes) tend to have higher frequencies, an amplitude peak relatively late in the note, and the peak frequency relatively early in the note. Only 17% of residual variance is explained by factor 2, and this residual variance is most strongly correlated with shape and end frequency (Table II).

We used factor 1 as an index of the properties of each note in a repeated-measures ANCOVA. Both sex category and capture site had an effect on factor 1 through a significant interaction between these parameters (sex: $F_{1,18} = 1.2, P = 0.29$; capture site: $F_{1,18} = 0.5, P = 0.48$; sex \times capture site: $F_{1,18} = 8.8, P = 0.008$; Fig. 3). A multiple comparisons analysis shows that the interaction results from two different effects: (1) a significant effect of sex in birds from the Ross capture site ($t_{18} = 2.8, P = 0.013$), with male birds having lower factor 1 scores, and (2) a significant population difference between males ($t_{18} = 2.6, P = 0.02$), with Martell males showing higher scores than Ross males. Analysis of individual parameters found the duration of the first A note (log-transformed to normalize variance) to be significantly different between populations ($F_{1,19} = 9.9$, adjusted $\alpha = 0.0083, P = 0.0054$): Martell A notes are longer than Ross A notes. Furthermore, there was a significant interaction between capture site and sex on beginning frequency ($F_{1,18} = 10.86$, adjusted $\alpha = 0.0083, P = 0.004$). A multiple comparisons analysis shows that the interaction re-

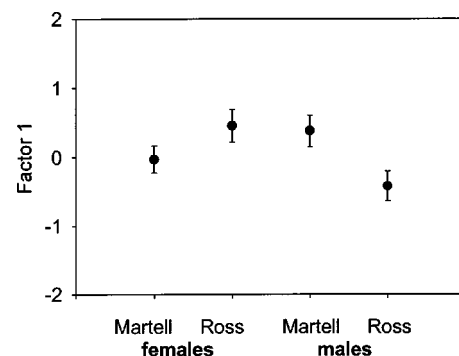


FIG. 3. Factor 1 scores for Martell and Ross females and males. Data are illustrated as means \pm s.e.

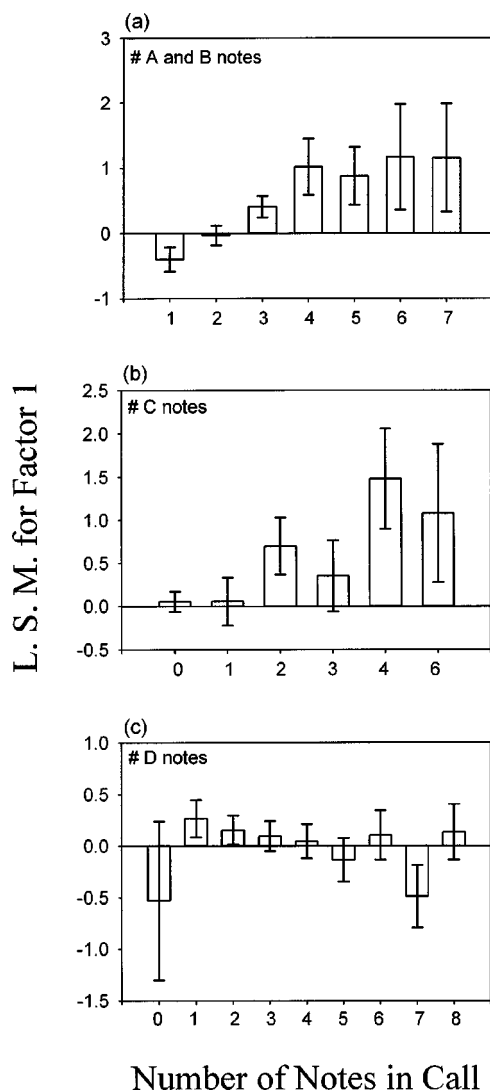


FIG. 4. Factor 1 for the first A note as a function of the number of introductory whistled A and B notes in the call (a), the number of C notes in the call (b), and the number of D notes in the call (c). Data are illustrated as least squares means \pm s.e.

sults from two different effects: (1) a significant population difference between males ($t_{18}=3.17$, $P=0.005$), with Martell males showing higher frequency A notes than Ross males, and (2) a significant effect of sex in birds from the Ross capture site ($t_{18}=3.04$, $P=0.007$), with male birds having lower frequency A notes than females.

Following the first A note in a call, the subsequent note composition of the call had a significant effect on the structure of that first A note. Factor 1 scores increased significantly with an increase in the number of introductory whistled notes (A and B notes) in the call [$\beta=0.39\pm0.06$; $F_{1,17}=43.5$, $P<0.0001$; Fig. 4(a)], and showed a marginally significant increase with an increase in the number of C notes [$\beta=0.20\pm0.09$; $F_{1,20}=5.0$, $P=0.048$; Fig. 4(b)]. We did not detect an effect of the number of D notes in the call on factor 1 scores for the first A note [$F_{1,20}=1.2$, $P=0.30$; Fig. 4(c)].

B. C notes

Unlike series of A notes, the structure of the first C note does not appear to change appreciably within a string of C

TABLE III. Factor loadings for the first two factors of a factor analysis of the properties of the first C note in the chick-a-dee call. See text for definitions of properties.

Parameter	Factor 1	Factor 2
Duration	-0.849	-0.283
Pitch	0.282	0.604
Continuity	-0.860	0.409
Entropy	0.891	-0.324
Peak frequency	0.901	0.096
Modulation	0.467	0.307
Eigenvalue	3.36	0.82
Variance explained	80%	18%

notes. In other words, the parameters we measured for C notes are not significantly different whether the note is the first, second, or last in a long string of C notes.

A factor analysis of the six parameters used to characterize C notes generated one factor with an eigenvalue greater than 1 (80% variance explained: Table III). Factor 1 was negatively correlated with C note duration and continuity, and positively correlated with entropy and peak frequency. Thus, short C notes (compared to long C notes) tend to be noisy (low continuity and high entropy), and have a high peak frequency.

Repeated measures ANCOVA suggests that factor 1 scores do not vary with bird sex category or with the site from which the birds were sampled (Factor 1: sex, $F_{1,10}=0.84$, $P=0.38$; site, $F_{1,10}=1.41$, $P=0.26$). Analysis of single parameters of the C notes found a significant effect of capture site on duration of the note ($F_{1,10}=31.51$, adjusted $\alpha=0.0083$, $P=0.0002$): Martell C notes are longer in duration than Ross C notes.

Also unlike series of A notes, the structure of the first C note does not appear to change appreciably with the syntax of the call. We detected no effect on factor 1 of the number of introductory whistled notes [number A+B notes: $F_{1,5}=0.97$, $P=0.37$; Fig. 5(a)], number of C notes [$F_{1,3}=0.21$, $P=0.68$; Fig. 5(b)], or number of D notes [$F_{1,2}=1.78$, $P=0.31$; Fig. 5(c)].

C. D notes

As with series of C notes, the structure of D notes does not appear to change greatly in a series of D notes. We collected measurements for the first, second, and last D note in each call (assuming the call had more than a single D note in it). Only note duration changed significantly, with the first D note being significantly shorter than the following notes ($F_{2,47}=16.1$, $p=0.0001$; Fig. 6). These trends were similar between sexes and between populations, as indicated by a lack of significant interaction between sex and the position of the note in the sequence of notes ($F_{2,45}=0.02$, $p=0.98$) and no significant interaction between population and note position ($F_{2,45}=2.58$, $p=0.09$).

A factor analysis of the 12 parameters used to characterize the first D note in each call generated two factors with eigenvalues greater than 1.0. Factor 1 (46% of the variance explained) primarily represents whole-note traits, showing positive correlations with the internote interval (log-

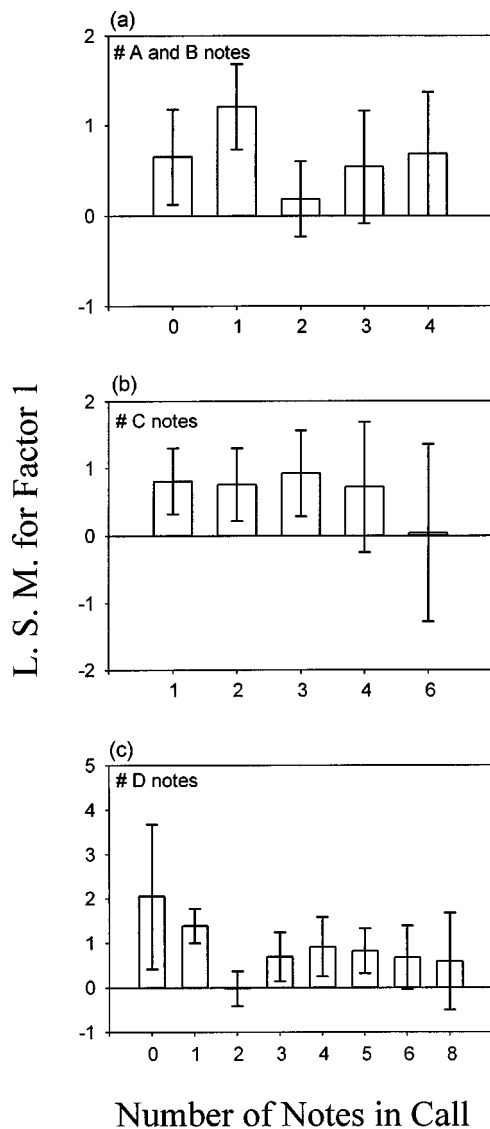


FIG. 5. Factor 1 for the first C note as a function of the number of introductory whistled A and B notes in the call (a), the number of C notes in the call (b), and the number of D notes in the call (c). Data are illustrated as least squares means \pm s.e.

transformed), entropy, and modulation, and negative correlations with note duration, continuity, and the amplitude difference between the first and maximal frequency peaks (Table IV). Thus, long notes (compared to short notes) tend to have high continuity, low entropy, low modulation, a large ampli-

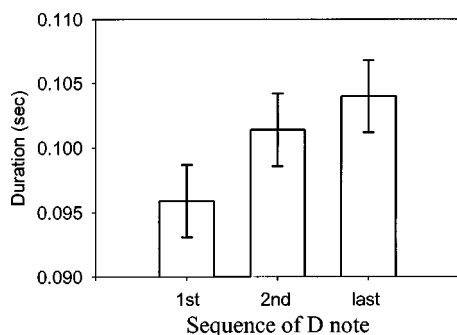


FIG. 6. Duration of D notes as a function of the sequence of D notes. Data are illustrated as means \pm s.e.

TABLE IV. Factor loadings for the first two factors of a factor analysis of the properties of the first D note in the chick-a-dee call. See text for definitions of properties.

Parameter	Factor 1	Factor 2
duration	-0.478	0.056
Log (Freq 1st peak)	0.249	0.857
Freq 2nd peak	0.134	0.892
Freq max. peak	0.074	0.512
Position of maximal amplitude	0.015	0.159
Log (Inter-note interval)	0.529	0.128
75%–50% amplitude modulation	0.135	-0.270
Max to first peak amplitude difference	-0.598	-0.183
Entropy	0.610	0.074
Continuity	-0.852	0.282
Log (Pitch)	-0.284	-0.113
Modulation	0.598	-0.287
Eigenvalue	2.54	2.09
Variance explained	46%	38%

tude difference between the maximal amplitude frequency peak and the first peak, and a short internote interval. Factor 2 (38% variance explained) is strongly correlated with the frequency properties of the note, showing a positive correlation with the first, second, and maximal frequencies (Table IV).

Repeated measures ANCOVAs indicate that neither factor 1 nor factor 2 scores vary with bird sex or with the site from which the birds were sampled (factor 1: sex, $F_{1,22} = 0.90$, $P = 0.35$; capture site, $F_{1,22} = 2.0$, $P = 0.17$; factor 2: sex, $F_{1,22} = 0.06$, $P = 0.81$; capture site, $F_{1,22} = 2.7$, $P = 0.12$). While the composite properties derived from the factor analysis did not differ between birds of different capture sites or sex categories, two individual note parameters did differ between capture sites. There was a significant effect of capture site on amplitude modulation ($F_{1,22} = 13.5$, adjusted $\alpha = 0.0042$, $P = 0.0013$) and an effect of capture site on frequency modulation ($F_{1,22} = 5.2$, adjusted $\alpha = 0.0045$, $P = 0.0026$). D notes from the Martell site show stronger frequency modulation and stronger amplitude modulation than D notes from the Ross site.

Note composition had a significant effect on the properties of the first D note. For factor 1 scores, this includes the number of introductory whistled notes preceding the first D note [number of A+B notes, $F_{1,21} = 16.3$, $P = 0.0006$; Fig. 7(a)], the number of C notes preceding the first D note [$F_{1,11} = 8.8$, $P = 0.013$; Fig. 7(b)], and the number of D notes following the first D note [$F_{1,23} = 19.9$, $P = 0.0002$; Fig. 7(c)]. For all note types, an increase in the number of notes increased the factor 1 score (number of A+B notes: $\beta = 0.32 \pm 0.08$; number of C notes: $\beta = 0.18 \pm 0.06$; number of D notes: $\beta = 0.20 \pm 0.04$). Thus, as the number of any single note type increased, the first D note became noisier (higher entropy, lower continuity), shorter in duration and with a longer internote interval, and more strongly frequency modulated.

Note composition had a different effect on factor 2 scores than it had on factor 1 scores. The number of introductory whistled notes and the number of C notes had no

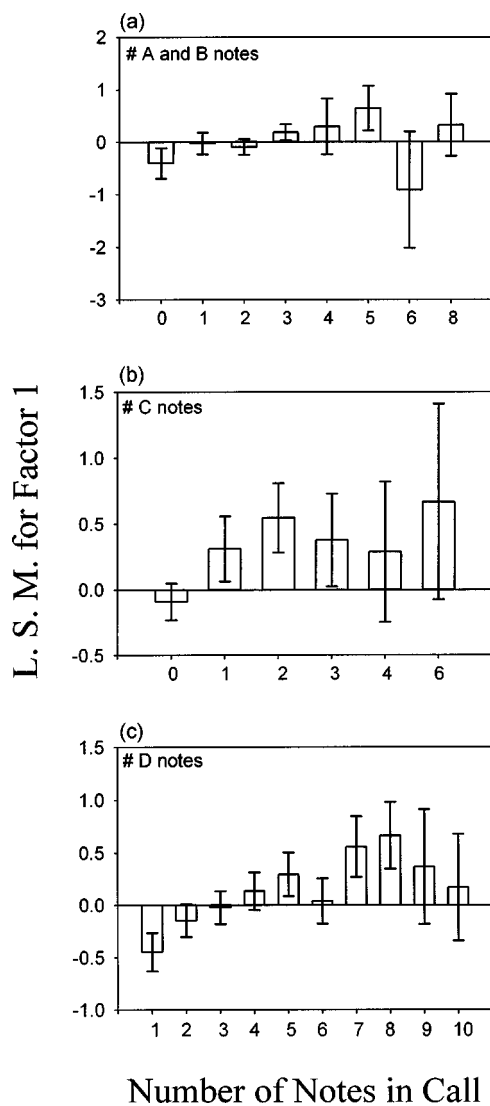


FIG. 7. Factor 1 for the first D note as a function of the number of introductory whistled A and B notes in the call (a), the number of C notes in the call (b), and the number of D notes in the call (c). Data are illustrated as least squares means \pm s.e.

effect on factor 2 scores (A+B notes: $F_{1,21}=0.13$, $P=0.72$; C notes: $F_{1,11}=3.4$, $P=0.09$). In contrast, factor 2 scores decreased significantly with an increase in the number of D notes ($F_{1,23}=5.3$, $P=0.031$; $\beta=-0.066\pm0.029$). Thus, first D notes tended to be lower in frequency when followed by a large number of D notes, compared to first D notes followed by few D notes.

IV. CONCLUSIONS

In the present study, we found that single elements within chick-a-dee calls (the first note in a string of A, C, or D notes) were predictive of a bird's local capture site. In one case we found a significant interaction between capture site and sex for a note type parameter, suggesting that at least some sex-specific characteristics of elements of a call can vary. However, at least for this Indiana population of Carolina chickadees, it appears that chick-a-dee calls are structurally much more sexually monomorphic than are vocalizations that have been tested for sexual differences in other

species—such as song in Northern cardinals (Yamaguchi, 1998a), contact calls in Bengalese finches (Okanoyak, 1993), and the vocal signals of collared doves (Ballintijn and ten Cate, 1997).

Most of the studies done to determine markers of identity in chick-a-dee calls of black-capped chickadees have focused on properties of the D notes (e.g., Mammen and Nowicki, 1981; Nowicki, 1989). Here, we found evidence of properties of D notes being predictive of a bird's capture site, but we also found that parameters in A and C notes were predictive of capture site. Factor analyses of A notes indicated an interaction between capture site and sex of bird for factor 1. Looking at individual parameters, both A note and C note duration predicted capture site. These microgeographic differences (5 km or less) suggest that chickadees learn these variations in note properties, as chickadees can easily disperse this distance within their lifetimes. Evidence of call learning has been documented in black-capped chickadees (Hughes *et al.*, 1998; Nowicki, 1989), and, in addition, the Nowicki (1989) study suggested that adult birds could modify call parameters. Call learning and vocal plasticity in both females and males would appear to be advantageous for chickadee species in that flocks may differ in parameters of their chick-a-dee calls, so newly-independent young chickadees dispersing some distance from their parents' territory might have to modify their calls. Moreover, even adult individuals may be in different flocks from one winter to the next (see Dixon, 1963; Mammen and Nowicki, 1981; Odum, 1941, 1942; Weise and Meyer, 1979).

A previous study on micro-level geographic differences in black-capped chickadee calls (Mammen and Nowicki, 1981) failed to control for the note composition of calls when considering identity markers, despite the fact that note composition characteristics of the chick-a-dee calls differed between some of the flocks studied. Our results suggest that note composition has to be statistically controlled to remove the statistical confound between note composition (syntax) and note structure. As such, the nature of flock differences in the structure of the D note of black-capped chickadees needs to be reexamined.

The first A and D notes in a string of A and D notes were largely predictive of the note type composition of the call as a whole—i.e., of other notes (both of the same and of different note types) in the call. This suggests possible redundancy of information in this call system—whatever the meaning of a complete chick-a-dee call, the receiver potentially could ascertain that meaning by hearing only a portion of the call. Signal redundancy can result from multiple signals that refer to the same information [e.g., visual and vibrational mating cues (Uetz, 2000)] or from multiple dimensions of a single signal each of which refers to the same information, such as individual identity cued by multiple overtones and several frequency-modulated elements within a note (Jouventin *et al.*, 1999). The chick-a-dee call provides a third type of redundancy where the properties of single notes provide cues about the syntactical structure of the entire call. Several studies have suggested that syntax (or more generally, note composition) is a critically important property of the chick-a-dee call (Ficken *et al.*, 1994; Freeberg and Lucas, 2002;

Smith, 1972). These studies provide a context for why redundancy might evolve in the chick-a-dee vocal system.

Redundancy increases the ability of a receiver to detect correctly the relevant properties of a vocal signal (Wiley and Richards, 1982). A recent study on penguin vocal patterns illustrates this point well. King penguins, *Aptenodytes patagonicus*, increase both the number of calls given and the number of syllables per call under windy conditions, compared to vocalizations given under calm conditions (Lengagne *et al.*, 1999). Thus, the high background noise generated by wind causes these birds to give signals that are more redundant, and this increases the probability that receivers of those signals will be able to detect signal-specific cues (here individual recognition cues). We should emphasize that while note repetition is a form of redundancy in species like the king penguin, in the chick-a-dee call note repetition is more likely to be a part of the information encoded in the signal (here the entire call) than a simple redundant repetition of information encoded in a single note (see Smith, 1972).

In some chickadee species, longer strings of D notes have been hypothesized to function in “mobbing” behaviors in the context of a predator (Odum, 1942). In addition to hearing the long string of harsher D notes used in these “scolding” calls (or if due to background noise, the entire string of D notes is not heard), the results from our study indicate that a receiver might gain the information from individual parameters of the first D note heard [Fig. 7(c)]. Alternatively, if individual parameters of the first A note in a call (or of the string of A notes in general) convey information about alarm level, as suggested by a study of a structurally similar note type in Mexican chickadees, *P. sclateri*, by Ficken (1989), then a call possessing redundant information in the form of individual parameters of the first D note could convey a similar meaning [Fig. 7(a)]. As A notes are highly directional note types and appear to degrade in the environment more quickly than do the other note types (Witkin, 1977), this redundancy could be of adaptive significance (see also Marten and Marler, 1977).

It is important to stress, however, a limitation of the present study—we have assessed note type variation and call differences only from the standpoint of signal production. Playback tests are needed to determine whether receivers do respond to the note differences we have detected (e.g., Ratcliffe and Otter, 1996; Searcy and Nowicki, 1999; Yamaguchi, 1998b). This has been done, for example, at the level of flock differences in chick-a-dee calls in black-capped chickadees (Nowicki, 1983).

Physiological and anatomical constraints may provide a mechanistic explanation for this “redundancy” hypothesis in notes in chick-a-dee calls. A signaler about to produce a long string of A notes may need to begin that string at a higher frequency and with different relative amplitude characteristics, due to requirements of muscular processes or neuroanatomical structure at the level of the syrinx (Gaunt, 1996; Lambrechts, 1996). Physical constraints on vocal production have been documented in a number of avian taxa (Larsen and Goller, 1999; Podos, 1997; Suthers, 1999). Limits on the production of chick-a-dee calls in black-capped chickadees have been suggested. Hailman *et al.* (1987) argued that call

length and a general frequency decrease over the course of the call are likely constrained by a process similar to “breath groups” in human speech, whereby fundamental frequencies of speech decrease as alveolar pressure drops before another breath is taken. Learning experiments could be useful in determining whether frequency drops in chick-a-dee calls (and the gargle vocalization of *Poecile* species), and therefore parameters of individual notes, are constrained by physiological process (Lambrechts, 1996), or are constrained by the acoustical and social environment of young birds (Hughes *et al.*, 1998).

The chick-a-dee call in many species of the genus *Poecile* is a combinatorial communicative system. Because of this combinatorial nature, it has been hypothesized that call types with different numbers and proportions of individual notes could convey different meanings to call receivers (Hailman and Ficken, 1986; Hailman *et al.*, 1985). Hailman and Ficken (1996) suggested that chickadee signalers use chick-a-dee calls under situations of mild alarm, but in a diversity of contexts. Different *Poecile* species, furthermore, may use similar chick-a-dee calls in different contexts (Hailman and Ficken, 1996). These relationships between call composition and calling context in *Poecile* species would seem to be a fruitful area for comparative studies. Data from the present study suggest that the redundancy between call composition and single note properties might be widely distributed in the *Poecile*—comparative work on identity and redundancy in this complex call system might help elucidate the relationships between call composition and meaning.

ACKNOWLEDGMENTS

This research was carried out under approved protocols of the Purdue University Animal Care and Use Committee, and followed the *Guidelines for the Use of Animals in Research* of the Association for the Study of Animal Behavior and the Animal Behavior Society and the *Guiding Principles in the Care and Use of Animals* of the American Physiological Society. We thank Billy Bragg, Steve Earle, and Lucinda Williams for helpful advice, and Adam Boyko, James Kellam, Terry Ord, and three anonymous reviewers for helpful comments on earlier drafts of the manuscript.

- Baker, M. C., and Cunningham, M. A. (1985). “The biology of bird-song dialects,” *Behav. Brain Sci.* **8**, 85–133.
- Ballintijn, M. R., and ten Cate, C. (1997). “Sex differences in the vocalizations and syrinx of the collared dove (*Streptopelia decaocto*),” *Auk* **114**, 22–39.
- Bateson, G. (1968). “Redundancy and coding,” in *Animal Communication: Techniques of Study and Results of Research*, edited by T. A. Sebeok (Indiana U.P., Bloomington, IN), pp. 614–626.
- Bradbury, J. W., and Vehrencamp, S. L. (1998). *Principles of Animal Communication* (Sinauer, Sunderland, MA).
- Dixon, K. L. (1963). “Some aspects of social organization in the Carolina chickadee,” in *Proceedings of the XIII International Ornithological Conference*, Vol. 1, pp. 240–258.
- Endler, J. A. (2000). “Evolutionary implications of the interaction between animal signals and the environment,” in *Animal Signals: Signalling and Signal Design in Animal Communication*, edited by Y. Espmark, T. Amundsen, and G. Rosenqvist (Tapir Academic, Trondheim, Norway), pp. 11–46.
- Ficken, M. S. (1989). “Acoustic characteristics of alarm calls associated with predation risk in chickadees,” *Anim. Behav.* **39**, 400–401.

- Ficken, M. S., Ficken, R. W., and Witkin, S. R. (1978). "Vocal repertoire of the black-capped chickadee," *Auk* **95**, 34–48.
- Ficken, M. S., Hailman, E. D., and Hailman, J. P. (1994). "The chick-a-dee call system of the Mexican chickadee," *Condor* **96**, 70–82.
- Freeberg, T. M., and Lucas, J. R. (2002). "Receivers respond differently to chick-a-dee calls varying in note composition in Carolina chickadees (*Poecile carolinensis*)," *Anim. Behav.* **63**, 837–845.
- Gaddis, P. K. (1985). "Structure and variability in the vocal repertoire of the mountain chickadee," *Wilson Bull.* **97**, 30–46.
- Gaunt, A. S. (1996). "A hypothesis concerning the relationship of syringeal structure to vocal abilities," *Auk* **100**, 853–862.
- Hailman, J. P. (1989). "The organization of major vocalizations in the Paridae," *Wilson Bull.* **101**, 305–343.
- Hailman, J. P., and Ficken, M. S. (1986). "Combinatorial animal communication with computable syntax: chick-a-dee calling qualifies as 'language' by structural linguistics," *Anim. Behav.* **34**, 1899–1901.
- Hailman, J. P., and Ficken, M. S. (1996). "Comparative analysis of vocal repertoires, with reference to chickadees," in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Cornell U.P., Ithaca, NY), pp. 136–159.
- Hailman, J. P., Ficken, M. S., and Ficken, R. W. (1985). "The 'chick-a-dee' calls of *Parus atricapillus*: a recombinant system of animal communication compared with written English," *Semiotica* **56**, 191–224.
- Hailman, J. P., Ficken, M. S., and Ficken, R. W. (1987). "Constraints on the structure of combinatorial 'chick-a-dee' calls," *Ethology* **75**, 62–80.
- Hauser, M. D. (1996). *The Evolution of Communication* (MIT, Cambridge, MA).
- Hughes, M., Nowicki, S., and Lohr, B. (1998). "Call learning in black-capped chickadees (*Parus atricapillus*): the role of experience in the development of 'chick-a-dee' calls," *Ethology* **104**, 232–249.
- Jouventin, P., Aubin, T., and Lengagne, T. (1999). "Finding a parent in a king penguin colony: the acoustic system of individual recognition," *Anim. Behav.* **57**, 1175–1183.
- Kroodsma, D. E., and Miller, E. H. (1996). *Ecology and Evolution of Acoustic Communication in Birds* (Cornell U.P., Ithaca, NY).
- Lambrechts, M. M. (1996). "Organization of birdsong and constraints on performance," in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Cornell U.P., Ithaca, NY), pp. 305–320.
- Larsen, O. N., and Goller, F. (1999). "Role of syringeal vibrations in bird vocalizations," *Proc. R. Soc. London, Ser. B* **266**, 1609–1615.
- Lengagne, T., Aubin, T., Lauga, J., and Jouventin, P. (1999). "How do king penguins (*Aptenodytes patagonicus*) apply the mathematical theory of information to communicate in windy conditions?" *Proc. R. Soc. London, Ser. B* **266**, 1623–1628.
- Lucas, J. R., Schraeder, A., and Jackson, C. (1999). "Carolina chickadee (*Aves*, Paridae, *Poecile carolinensis*) vocalization rates: effects of body mass and food availability under aviary conditions," *Ethology* **105**, 503–520.
- Mammen, D. L., and Nowicki, S. (1981). "Individual differences and within-flock convergence in chickadee calls," *Behav. Ecol. Sociobiol.* **9**, 179–186.
- Marten, K., and Marler, P. (1977). "Sound transmission and its significance for animal vocalization, I: temperate habitats," *Behav. Ecol. Sociobiol.* **2**, 271–290.
- Mundinger, P. C. (1982). "Microgeographic and macrogeographic variation in the acquired vocalizations of birds," in *Acoustic Communication in Birds, Volume 2: Song Learning and its Consequences*, edited by D. E. Kroodsma and E. H. Miller (Academic, New York), pp. 147–208.
- Nowicki, S. (1983). "Flock-specific recognition of chickadee calls," *Behav. Ecol. Sociobiol.* **12**, 317–320.
- Nowicki, S. (1989). "Vocal plasticity in captive black-capped chickadees: the acoustic basis and rate of call convergence," *Anim. Behav.* **37**, 64–73.
- Nowicki, S., and Nelson, D. A. (1990). "Defining natural categories in acoustic signals: comparison of three methods applied to 'chick-a-dee' call notes," *Ethology* **86**, 89–101.
- Odum, E. P. (1941). "Annual cycle of the black-capped chickadee-2," *Auk* **58**, 518–535.
- Odum, E. P. (1942). "Annual cycle of the black-capped chickadee-3," *Auk* **59**, 499–531.
- Okanoyak, K. T. (1993). "Acoustical and perceptual structure of sexually dimorphic distance calls in Bengalese finches (*Lonchura striata domestica*)," *J. Comp. Psychol.* **107**, 386–394.
- Podos, J. (1997). "A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Embirizidae)," *Evolution* (Lawrence, Kans.) **51**, 537–551.
- Pyle, P. (1997). *Identification Guide to North American Birds* (Slate Creek, Bolinas, CA).
- Ratcliffe, L., and Otter, K. (1996). "Sex differences in song recognition," in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Academic, New York), pp. 339–355.
- Rice, W. R. (1989). "Analyzing tables of statistical tests," *Evolution* (Lawrence, Kans.) **43**, 223–225.
- SAS Institute, Inc. (1990). *SAS/STAT User's Guide, Version 6, 4th Edition* (SAS, Cary, NC).
- Searcy, W. A., and Nowicki, S. (1999). "Functions of song variation in song sparrows," in *The Design of Animal Communication*, edited by M. D. Hauser and M. Konishi (MIT, Cambridge, MA), pp. 577–595.
- Smith, S. M. (1991). *The Black-Capped Chickadee: Behavioral Ecology and Natural History* (Comstock, Ithaca, NY).
- Smith, S. T. (1972). "Communication and other social behavior in *Parus carolinensis*," *Publ. Nuttall Ornithol. Club* **11**, 1–125.
- Smith, W. J. (1977). *The Behavior of Communicating: An Ethological Approach* (Harvard U.P., Cambridge, MA).
- Suthers, R. A. (1999). "The motor basis of vocal performance in songbirds," in *The Design of Animal Communication*, edited by M. D. Hauser and M. Konishi (MIT, Cambridge, MA), pp. 37–62.
- Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., and Mitra, P. P. (2000). "A procedure for an automated measurement of song similarity," *Anim. Behav.* **59**, 1167–1176.
- Thirakhuat, K. (1985). "Foraging ecology of sympatric parids: individual and population responses to winter food scarcity," Ph.D. thesis. Purdue University, West Lafayette, IN.
- Uetz, G. W. (2000). "Signals and multi-modal signaling in spider communication," in *Animal Signals: Signalling and Signal Design in Animal Communication*, edited by Y. Espmark, T. Amundsen, and G. Rosenqvist (Tapir Academic, Trondheim, Norway), pp. 387–405.
- Weise, C. M., and Meyer, J. R. (1979). "Juvenile dispersal and development of site-fidelity in the black-capped chickadee," *Auk* **96**, 40–55.
- Wiley, R. H. (1983). "The evolution of communication: information and manipulation," in *Animal Behaviour, Volume 2: Communication*, edited by T. R. Halliday and P. J. B. Slater (Blackwell Scientific, Oxford), pp. 156–189.
- Wiley, R. H., and Richards, D. G. (1982). "Adaptations for acoustic communication in birds: sound transmission and signal detection," in *Acoustic Communication in Birds*, edited by D. E. Kroodsma, E. H. Miller, and H. Ouellet (Academic, New York), pp. 131–181.
- Witkin, S. R. (1977). "The importance of directional sound radiation in avian vocalization," *Condor* **79**, 490–493.
- Yamaguchi, A. (1998a). "A sexually dimorphic learned birdsong in the Northern cardinal," *Condor* **100**, 504–511.
- Yamaguchi, A. (1998b). "Can a sexually dimorphic learned birdsong be used for male-female recognition?" *Behaviour* **135**, 833–844.

Time-frequency model for echo-delay resolution in wideband biosonar

Nicola Neretti^{a)}

Brain Sciences, Brown University, Providence, Rhode Island 02912 and Institute for Brain and Neural Systems, Brown University, Providence, Rhode Island 02912

Mark I. Sanderson

Department of Neuroscience, Brown University, Providence, Rhode Island 02912

Nathan Intrator

Institute for Brain and Neural Systems, Brown University, Providence, Rhode Island 02912 and School of Computer Science, Tel-Aviv University, Ramat Aviv 69978, Israel

James A. Simmons

Department of Neuroscience, Brown University, Providence, Rhode Island 02912

(Received 9 August 2002; revised 17 December 2002; accepted 30 December 2002)

A time/frequency model of the bat's auditory system was developed to examine the basis for the fine ($\sim 2 \mu\text{s}$) echo-delay resolution of big brown bats (*Eptesicus fuscus*), and its performance at resolving closely spaced FM sonar echoes in the bat's 20–100-kHz band at different signal-to-noise ratios was computed. The model uses parallel bandpass filters spaced over this band to generate envelopes that individually can have much lower bandwidth than the bat's ultrasonic sonar sounds and still achieve fine delay resolution. Because fine delay separations are inside the integration time of the model's filters ($\sim 250\text{--}300 \mu\text{s}$), resolving them means using interference patterns along the frequency dimension (spectral peaks and notches). The low bandwidth content of the filter outputs is suitable for relay of information to higher auditory areas that have intrinsically poor temporal response properties. If implemented in fully parallel analog-digital hardware, the model is computationally extremely efficient and would improve resolution in military and industrial sonar receivers. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1554693]

PACS numbers: 43.80.Lb, 43.64.Bt [WA]

I. INTRODUCTION

The behavior of echolocating bats that emit frequency-modulated (FM) biosonar sounds shows that they create a detailed three-dimensional representation of their immediate environment from processing echoes of these sounds (Neuweiler, 2000; Popper and Fay, 1995). The images these bats perceive incorporate the shapes of objects at their correct locations over the operating range of their sonar [e.g., ~ 5 m for big brown bats (Kick, 1982)]. Experimental evidence indicates that FM bats are capable of perceiving objects with a resolution of the order of millimeters and fractions of a millimeter (Simmons *et al.*, 1995, 1996, 1998). This performance is amazing because the neural representations that support the bat's images are formed by auditory midbrain and cortical neurons whose responses have a temporal precision of hundreds of microseconds to several milliseconds at best (Casseday and Covey, 1995; O'Neill, 1995; Dear *et al.*, 1993; Pollak and Casseday, 1989). Not only is the sharpness of the bat's images better than individual neurons seem able to sustain, but the computations required to place the spatial information the bats perceive into images from sonar signals are very intensive, involving large numbers of parallel temporal calculations. For example, the big brown bat (*Eptesicus*

fuscus) transmits FM signals containing two to three harmonics that collectively span the band from 20 to over 100 kHz. These broadcasts are beamed broadly, so they ensoundify objects in nearly every direction, especially towards the bat's front. Consequently, echoes are effectively received from all the objects in the field of "view," which makes it necessary to explicitly form simultaneous distinct representations for each object so that they become segregated in perception (Simmons *et al.*, 1996). The range of hearing in big brown bats does not extend beyond 100 kHz (Dalland, 1965; Koay *et al.*, 1997), so frequencies in the upper half of the third harmonic probably do not contribute to echo processing. The representation of each object derived from any single broadcast has to be integrated with the corresponding representations from previous and subsequent broadcasts so that the object's path can be tracked and its shape reconstructed as a single coherent picture from the succession of echoes (Grinnell, 1995; Kalko and Schnitzler, 1998). To create the bat's images, detailed information about the wideband, intrinsically high-resolution time-series FM waveforms of broadcasts and echoes first has to be converted into a representation capable of being transferred efficiently into higher auditory, presumably cortical areas, without using a bandwidth for transmission in any one neural channel that exceeds the surprisingly limited temporal response properties of the higher-level neurons that accept this information. Information necessary for measuring the arrival-times and

^{a)}Address for correspondence: Institute for Brain and Neural Systems, Box 1843, Brown University, Providence, RI 02912. Electronic mail: nicola_neretti@brown.edu

arrival-time separations of multiple echoes has to survive the compression of the representation caused by the abrupt decrease in the temporal precision of neural responses between the auditory brainstem and the midbrain (Haplea *et al.*, 1994; Casseday and Covey, 1995). This would allow the bat to perceive the objects with high resolution using its seemingly low-resolution computational elements. The desired representation presumably takes advantage of the ability of parallel neurons arranged in neuronal maps to carry the detailed information using multiple, parallel low-pass representations.

We propose a model of biosonar processing which performs a set of detailed measurements on echo returns in the brainstem and sends the outcome of these measurements using multiple low-bandwidth neuronal channels to higher areas of the auditory system, in the midbrain and cortex. This paper focuses on retention of high resolution for the delay of closely spaced echoes in a computational representation suitable for incorporating delay differences in range-axis images. In the context of delay resolution, we demonstrate the connection between the signal-to-noise ratio of the echoes and the resolution of echo separation, as would be embodied in neural responses in higher cortical areas. In general, higher resolution requires wider bandwidth, and, for a given echo resolution, lower signal-to-noise ratios require broader—rather than narrower-band processing. For detection, systems are designed instead to focus on narrower signal bands when noise is strong to concentrate the energy available for overcoming the noise.

We want to clearly state that *resolution* and *accuracy* refer to *separate* features of an echo-delay estimator, and the terms should not be used interchangeably (see Schnitzler and Henson, 1980). The accuracy of a delay estimator is the uncertainty in its estimate for the arrival-time of echoes from a single reflecting point. In contrast, delay resolution refers to the minimum separation between two nearly simultaneous, overlapping echoes where each reflection is assigned its own delay estimate instead of being interpreted as a single echo.

II. MODEL

A. Background

Localization of objects from their echoes is a fundamental problem for analysis of acoustic signals. It is the basis of object exploration and scene analysis in sonar systems. Wideband active acoustic exploration of scenes relies on transmitting a series of pings, which impinge on objects and then return from different edges and surfaces. Distances to different parts of objects can be determined from the arrival-times of the individual replicas of the incident sound included in the overall return. The problem for understanding the performance of FM bats in such tasks as discrimination of airborne mealworms from disks (Griffin *et al.*, 1965) is how to conceive of the information necessary to reconstruct the appearance of objects being carried in the responses of neurons in the bat's auditory pathways.

When a sonar signal hits an object that is composed of several scattering points or planes (called “glints” in wide-band parlance), there are multiple returns from the object. The delay between those returns gives an accurate indication

of the structure in depth of object surfaces, or the separation of the glints in range. Using such delays at both ears, the bats may be able to achieve a complete reconstruction of the object in the range-crossrange plane. The resolution of the sonar is determined by the smallest detectable temporal difference between echoes. Higher temporal resolution between echoes leads to a higher depth and shape resolution of objects. For high temporal resolution, the transmitted signal should be wideband. There also is a competing requirement for being able to detect targets at long range from weak echoes. The maximum range for object detection is governed by the energy of the signal—that is, by its amplitude and duration. Longer duration leads to higher integrated energy and thus longer operating range. Echolocating bats apply the classical “chirp-radar” solution to this problem—they transmit FM sounds whose bandwidth is kept high by the frequency span of the FM sweep, while energy is increased by increasing the signal's duration in conditions of low echo strength. The bat's inner ear receives the FM sweeps through parallel bandpass filters (Kössl and Vater, 1995; Ruggero, 1992) whose outputs are smoothed to create an integration-time of several hundred microseconds for echo reception (Simmons *et al.*, 1995, 1996). The resulting auditory representation is a spectrogramlike *time-frequency* distribution for the energy in the sound made up of the envelopes of the bandpass-filtered, smoothed segments of the FM sweeps. The integration-time of the bat's auditory spectrograms is relatively long compared to the intrinsic time resolution of the bat's signals, which for big brown bats is crudely indexed to be about 12 μ s from the reciprocal of the broadcast bandwidth (80 kHz). The much longer integration-time of the bat's auditory filters means that echoes arriving closer together in time than several hundred microseconds will merge together to form a single sound at separations far larger than the resolution limit of the signals themselves. Big brown bats nevertheless can assign two closely spaced echoes their own arrival-times at far smaller separations. They readily distinguish echoes only 10–30 μ s apart as separate, for example (Simmons *et al.*, 1995). The bat's limit of resolution at moderate signal-to-noise ratios is about 2 μ s (Simmons *et al.*, 1998). When two echoes arrive closer together than the integration-time, they mix to create an interference pattern of peaks and notches in the spectrum, where the frequency spacing of the notches is the reciprocal of the time separation of the echoes. If the big brown bat can resolve echoes as separate if they are only 10–30 μ s apart, it must use the interference pattern in *frequency* to estimate the separation in *time*. This conclusion implies that the bat uses its spectrogramlike time-frequency representation as more than just the format of its initial auditory code—it uses the properties of the two dimensions of time and frequency as the basis for computations to assemble spatial images of objects (Simmons *et al.*, 1996). Here, we examine the resolution attainable from a time-frequency representation of the big brown bat's FM signals for two echoes arriving at close time separations and different signal-to-noise ratios.

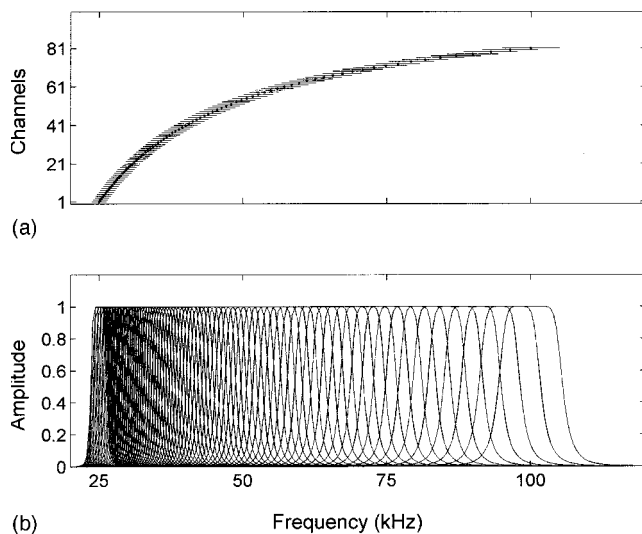


FIG. 1. Frequency response of the 81 bandpass filters (b); corresponding bandwidths and center frequencies (a).

B. Outline of the model

Our model is based on the generation of multiple time/frequency templates representing ideal responses of the bat's auditory filters for different two-glint echo separations in the absence of noise. The model is evaluated by supplying input signals at different time separations and signal-to-noise ratios to test the ability of the time-frequency templates for resolving two-glint delay separations. The structure of the simulations resembles a psychophysical experiment on echo-delay resolution by bats. The filters that generate the frequency axis of the time/frequency representation are analogous to the cochlear block of the SCAT model (Saillant *et al.*, 1993). This model describes a comprehensive imaging system based on auditorylike time-frequency representations (Matsuo *et al.*, 2001; Peremans and Hallam, 1998; Saillant *et al.*, 1993). We modeled the action of the inner ear by a bank of bandpass filters followed by an envelope-smoothing process. The frequency selectivity of the basilar membrane in the bat's cochlear is simulated by 81 tenth-order IIR Butterworth filters, with a constant $Q=10$, and center frequencies hyperbolically spaced between 25 and 100 kHz (Fig. 1). The excitation of the hair cells and primary auditory neurons is modeled by half-wave rectification and low-pass filtering using an IIR Butterworth first-order filter with a cutoff frequency of 3 kHz.

Figure 2 illustrates the effect of the different stages of the transduction filtering process on a frequency-modulated (FM) hyperbolic chirp with two harmonics, a start-frequency of 100 kHz, and an end-frequency of 25 kHz [Fig. 2(a)]. Figure 2(b) is the output of a bandpass filter with a center frequency of 40 kHz and $Q=10$. The output is then half-wave rectified [Fig. 2(c)] and low-pass filtered [Fig. 2(d)].

The effect of the entire filter bank is to encode the waveforms in the spectrogramlike time/frequency format characteristic of the mammalian auditory system. Figure 3 shows the output of the entire filter bank for three different signals. Each horizontal slice in the bottom plots corresponds to the output of 1 of the 81 cochlear filters after half-wave rectifi-

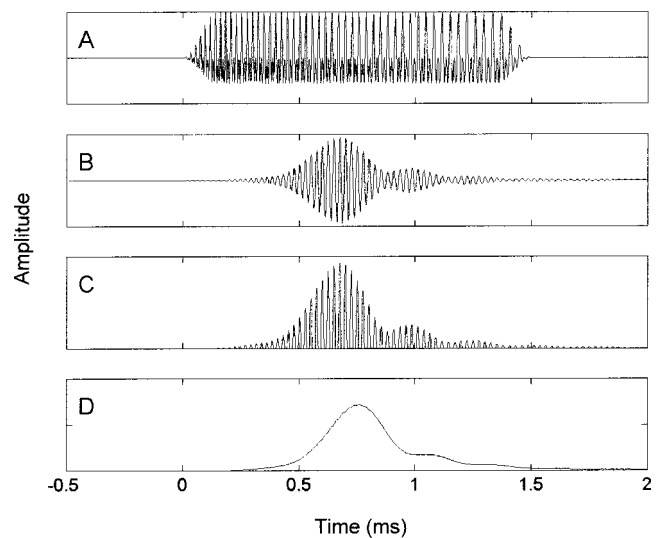


FIG. 2. (a) Hyperbolic chirp with two harmonics. (b) Output of a bandpass filter with center frequency of 40 kHz and $Q=10$. (c) Output after halfwave rectification. (d) Output of a low-pass frequency with COF=3 kHz.

cation and low-pass filtering. The signals in this example are based on a hyperbolic chirp with one harmonic. Since the center frequencies of the filters are hyperbolically spaced as well, the peaks in the response curves that trace the FM sweep fall on a straight line (the chirp's sweep is linearized).

If the echo is composed of distinct overlapping reflections from two closely spaced glints, then interference occurs. In particular, if the delay between the two glints is larger than the integration time of the filters ($\sim 300 \mu\text{s}$ in our model), then the echoes from the two glints generate two separate response ridges in the time-frequency representation (Fig. 3, right), and the distance from each glint to the emitted pulse can be computed independently entirely from the locations of the ridges in time. However, for smaller two-glint separations, it is no longer possible to distinguish between the two glints in the time domain because together they produce only one ridge. Interference gives rise to notches in the ridge whose number and location in frequency depend on the

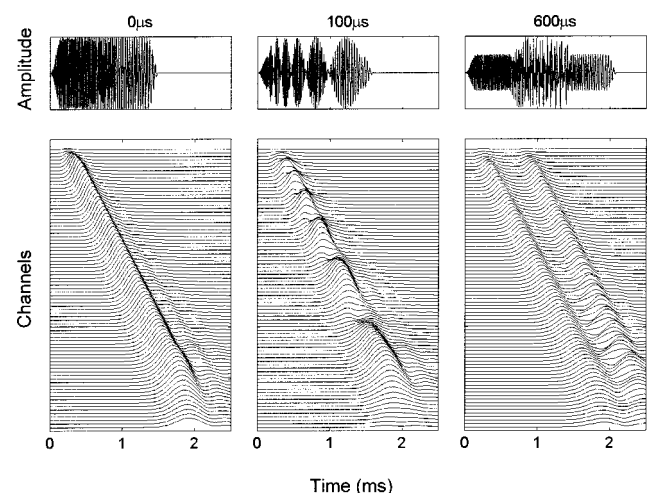


FIG. 3. FM hyperbolic echo with one harmonic and different two-glint separations (top row). Output of the filter-bank (bottom row). The bandpass filters have center frequencies hyperbolically spaced between 25 and 100 kHz and have constant $Q=10$. The low-pass filter has a COF=3 kHz.

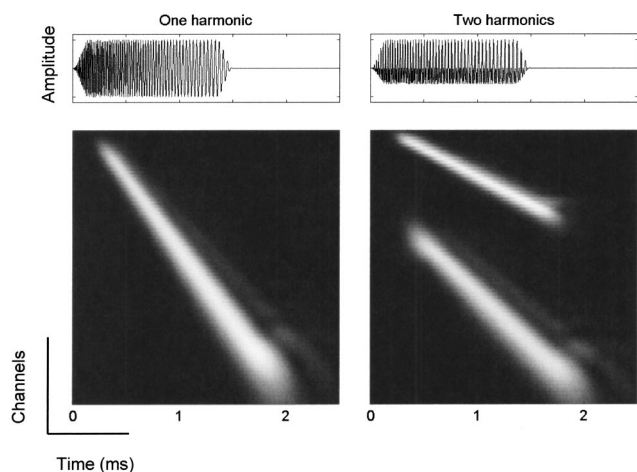


FIG. 4. One harmonic versus two harmonics.

two-glnt separation in time (Fig. 3, center). Figure 3 shows these effects with a single harmonic to simplify the illustration, but the same effects occur for multiple harmonic signals. In our simulations we used a more biologically realistic chirp with two harmonics, the first harmonic spanning the interval between 25 to 50 kHz. The difference between one and two harmonics covering the same 25–100-kHz band is shown in Fig. 4.

In order to study the effects of background noise on the two-glnt discrimination task, we generated band-limited white noise in the frequency range of the pulse—i.e., from 25 to 100 kHz—and added this noise to the echoes used in the simulations. The signal-to-noise ratio (SNR) was computed according to (Menne and Hackbarth, 1986)

$$\text{SNR(dB)} = 20 \log_{10} \sqrt{\frac{2E}{N_0}}, \quad (1)$$

where E is the total energy of the returning signal and N_0 is the spectral density of the noise. Figure 5 shows the effect of

different noise levels on the amplitude of a two-glnt echo [Fig. 5(a)] and on its time frequency representation [Fig. 5(b)]. Notice that, as the noise level increases (i.e., the SNR decreases), the location of the notches becomes increasingly more difficult to determine [Fig. 5(b), center]. For very high noise levels the notches are almost completely masked by the noise [Fig. 5(b)].

To build the model of echo-processing, a set of time-frequency templates was generated by applying the bandpass filter bank to a set of FM signals with different two-glnt separations but no noise. These templates were then used as a set of parallel matched filters on the time-frequency representation of echoes mixed with noise created using the same bandpass filters. Figure 6 shows a collection of time-frequency templates used in the simulation. Each template corresponds to a specific two-glnt separation, from 0 to 48 μs , with steps of 2 μs . The resolution of this set of templates thus is 2 μs . To explore higher resolution levels we generated a different set of templates in the same way, but using 0.2- μs intervals from 0 to 5 μs in the two-glnt separations. We will refer to the 2- μs interval simulation as *coarse* and to the 0.2- μs interval simulation as *fine*.

For a given two-glnt separation Δt in the echo at a fixed SNR, we generated its time-frequency representation $X(\Delta t)$ using the 81 cochlear filters described earlier. We then compared this representation to each time-frequency template $Y(n)$ according to the following similarity measure:

$$M_{\Delta t}(n) = \frac{\sum_{ij} X_{ij}(\Delta t) \cdot Y_{ij}(n)}{\|X(\Delta t)\| \cdot \|Y(n)\|}, \quad (2)$$

where i is the time index and $j=1, \dots, 81$ is the frequency index. The curve $M_{\Delta t} = M_{\Delta t}(n)$ represents the responses of the collection of templates to the particular echo. The two-glnt delay separation corresponding to the template with the highest response was used as our estimate $\hat{\Delta t}$ for the two-glnt separation Δt . This procedure was used for a series of

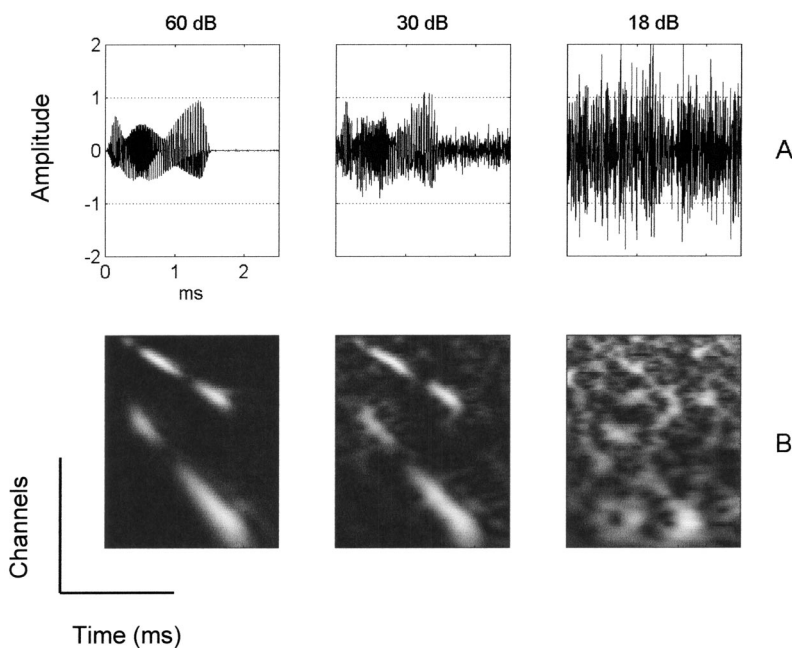


FIG. 5. (a) Echo with two glints separated by 40 μs embedded in band-limited noise for different signal-to-noise ratios. The frequency band of the noise is set to be equal to that of the pulse. (b) Time-frequency representation corresponding to the signals in (a).

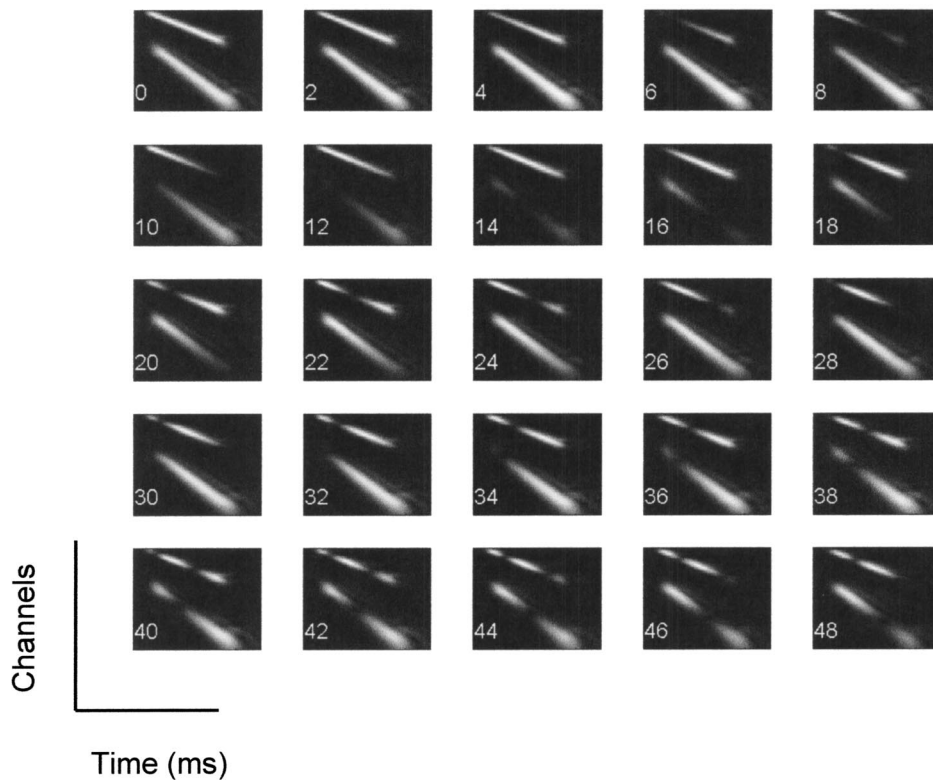


FIG. 6. Time-frequency templates used in the simulation. Each template corresponds to a specific two-glint separation, from 0 to 48 μs with a step of 2 μs .

two-glint echoes with different Δt 's. In particular, we used $\Delta t = 0 \mu\text{s}, 2 \mu\text{s}, 4 \mu\text{s}, \dots, 50 \mu\text{s}$ for the coarse resolution simulations and $\Delta t = 0 \mu\text{s}, 0.2 \mu\text{s}, 0.4 \mu\text{s}, \dots, 5 \mu\text{s}$ for the fine resolution ones. For any given SNR, the family of response curves corresponding to different Δt 's can be combined together to form a surface, which we will refer to as the *response surface*. Since the delays in both the filters and the echoes are the same, i.e., $\Delta t = 0 \mu\text{s}, 2 \mu\text{s}, 4 \mu\text{s}, \dots, 50 \mu\text{s}$ for the coarse resolution simulation, the correct estimates lay on the diagonal of the base plane of the surface. Figure 7 shows the response surface for single simulations of the coarse resolution experiment at three different noise levels. As the noise level increases (from top to bottom) the maximum response in some regions shifts away from this diagonal. For very low signal-to-noise levels the responses of each template to different two-glint separation in the echo are approximately equal (bottom figure), indicating a reduced discrimination power. Figure 8 shows the same surfaces for fine resolution simulations.

III. RESULTS

Each of our simulated experiments with the model is performed with a specific SNR and two-glint separation Δt in the echo. The outcome of each experiment is the estimate $\hat{\Delta t}$ of the two-glint separation. We performed a Monte Carlo simulation for each experiment, generating 20 realizations of the noise for each given SNR, and determining a new $\hat{\Delta t}$ for each realization. The variability in the estimate of the two-glint separation can be visually assessed by looking at the histogram of $\hat{\Delta t}$. Each box in Figs. 9 and 10 shows the histogram of $\hat{\Delta t}$ for a given Monte Carlo simulation at different two-glint separations (columns) and noise levels (rows). In

each box a thin vertical line marks the position of the correct response. For large signal-to-noise ratios, the template with the highest response corresponds to the correct value of Δt , and all the points in the histogram fall in this correct bin. For

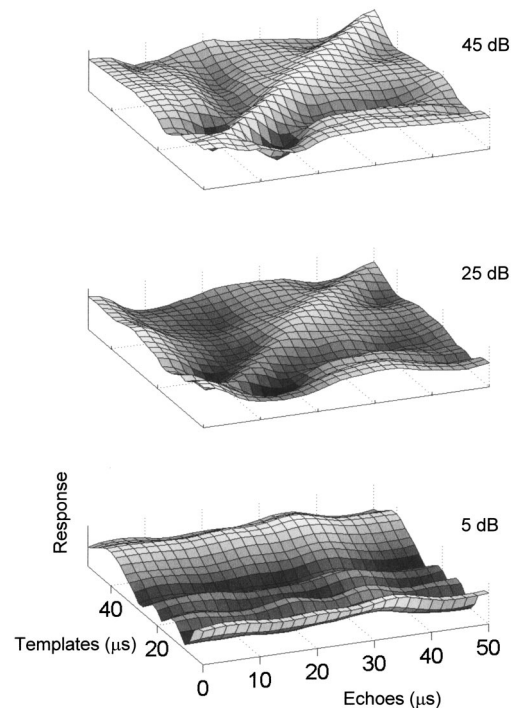


FIG. 7. Response surface of the templates for different two-glint separations in the echo. The main diagonal represents the correct response. As the noise level increases (from top to bottom) the maximum response moves away from the diagonal. For very low signal-to-noise levels the responses of each template to different two-glint separation in the echo are approximately equal (bottom figure), indicating a reduced discrimination power. The above figures correspond to a COF=7 kHz.

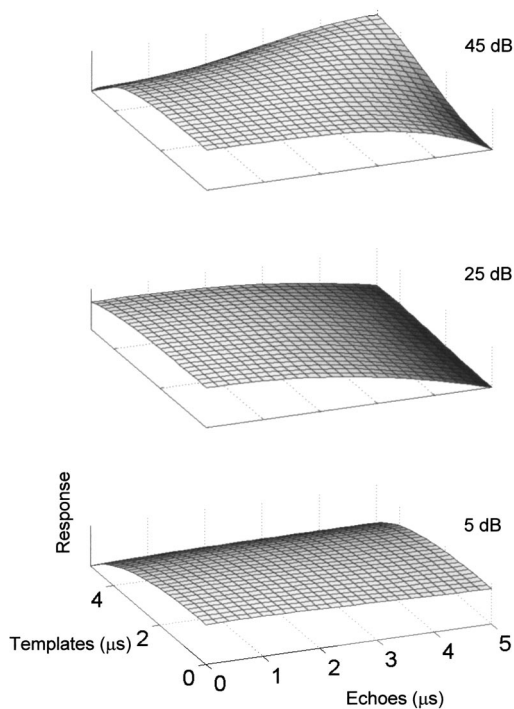


FIG. 8. Response surface of the templates in the case of $0.2\text{-}\mu\text{s}$ increments in the two-glint separations.

lower SNRs, the model starts making mistakes, so that some points in the histograms appear in the wrong bins (e.g., SNR = 30, 25 in Fig. 9, SNR = 45 in Fig. 10). For very low SNRs, most of the responses fall in the $0\text{-}\mu\text{s}$ bin, for both the fine

and the coarse resolution experiments. In addition, in the coarse experiment some of the responses cluster around the values 22 and $46\text{ }\mu\text{s}$.

The bottom rows of Figs. 9 and 10 show the combined histogram of the errors for all the Monte Carlo simulations with the same SNR, and provide a summary of the model's performance for a specific noise level. These summary histograms are centered on $0\text{ }\mu\text{s}$ as the nominally correct delay separation for all the different actual delay separations in the columns above. As the SNR decreases, the combined histogram becomes flatter and its spread around the $0\text{-}\mu\text{s}$ center increases. Notice that for very low SNR there are only errors on the negative side of $0\text{ }\mu\text{s}$. This is a consequence of the bias towards the $0\text{-}\mu\text{s}$ template for all the high noise levels mentioned earlier.

A major reason for doing these experiments was to evaluate the effects of changing a physiologically relevant component of the model, the smoothing filter applied to the half-wave rectified output of each cochlear bandpass filter. Previous work has identified the frequency cutoff of this smoothing filter as the critical design parameter for echo transduction (Simmons, 1980). We compared the performance of various models that differ in the smoothing frequency (COF) of the low-pass filter in each of the 81 parallel cochlear bandpass channels. As expected, high COFs are more resilient to noise than low ones. Figures 11 and 12 show error histograms for different COFs (rows) and SNRs (columns). The bandpass filter case (BP only) corresponds to a $\text{COF} = \infty$. A summary plot for all the simulations discussed so far is shown in Fig. 13, which shows the root mean square error (RMSE) in μs versus signal-to-noise level in dB. Different curves correspond to different cutoff frequencies in the

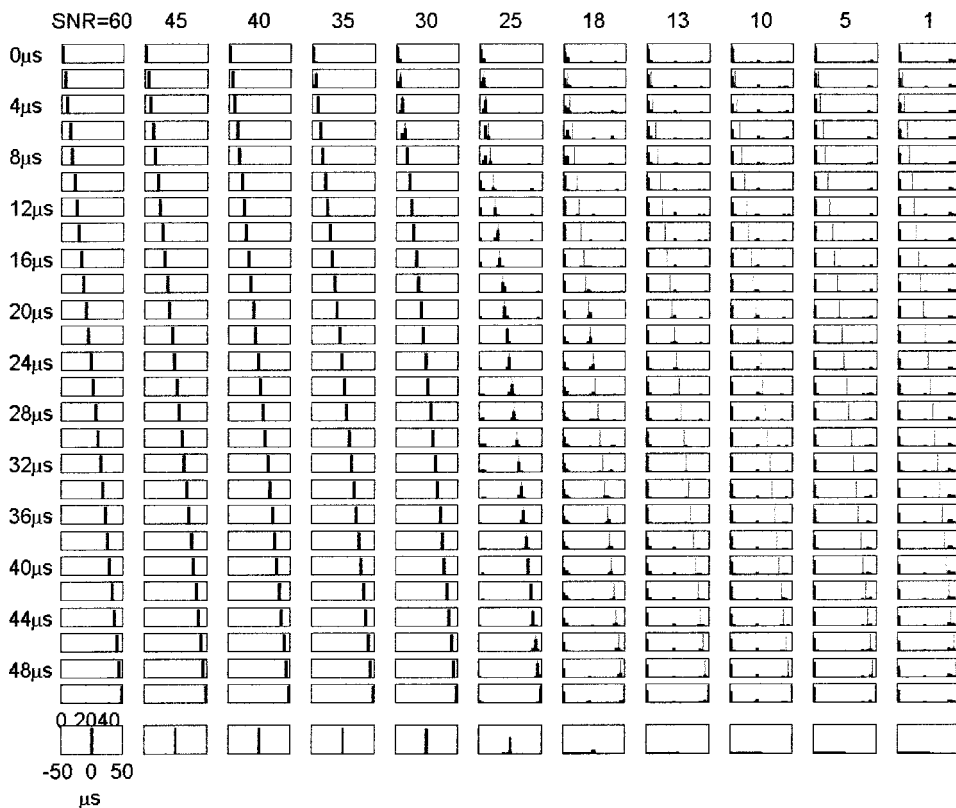


FIG. 9. Histogram of $\hat{\Delta t}$ estimates from the maximum responses of the templates for different two-glint separations (columns) and signal-to-noise ratios (rows). The bottom row shows the combined histogram of these estimates of delay separation for the entire collection of trials in the corresponding column (with Δt values centered on nominal values as zero). Results are given for a low-pass COF = 7 kHz, and $2\text{-}\mu\text{s}$ increments in the two-glint separations of the templates, and for 20 realizations of the noise for each given SNR.

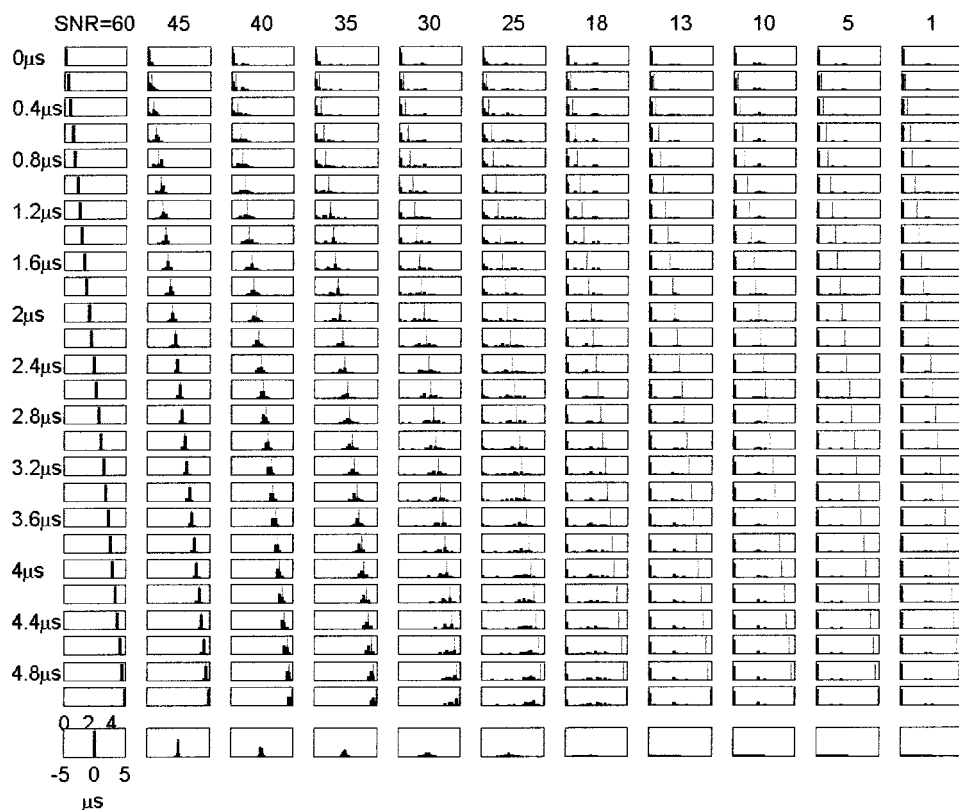


FIG. 10. Histogram of $\hat{\Delta t}$ estimates from the maximum responses of the templates for different two-glint separations (columns) and signal-to-noise ratios (rows). The bottom row shows the combined histogram of these estimates of delay separation for the entire collection of trials in the corresponding column (with Δt values centered on nominal values as zero). Results are given for a low-pass COF=7 kHz, and 0.2- μ s increments in the two-glint separations of the templates, and for 20 realizations of the noise for each given SNR.

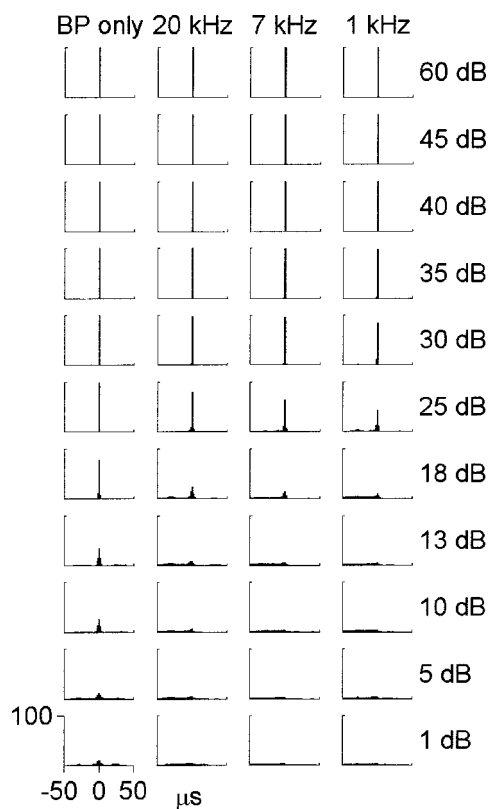


FIG. 11. Plots of summary histograms for errors in coarse-resolution (2- μ s steps) simulations across the entire collection of trials (bottom rows in Fig. 9) for different COFs and SNRs.

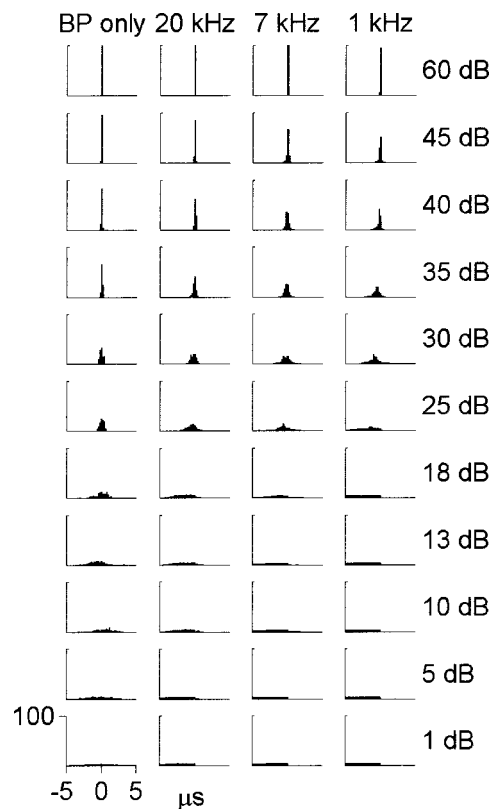


FIG. 12. Plots of summary histograms for errors in fine-resolution (0.2- μ s steps) simulations across the entire collection of trials (bottom rows in Fig. 10) for different COFs and SNRs.

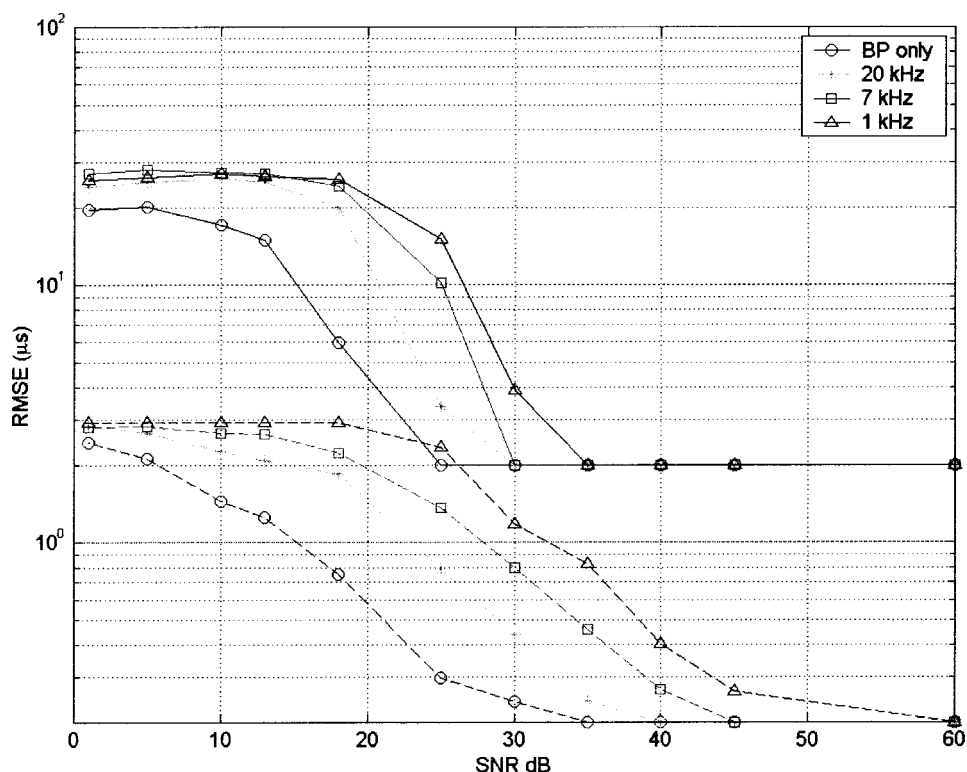


FIG. 13. Root mean square error (RMSE) in μs versus signal-to-noise level in dB. Different curves correspond to different cutoff frequencies in the low-pass filter; circles correspond to no halfway rectification and no low-pass filtering, i.e., the bandpass filters alone are used to create the time-frequency representation. The solid lines in the top part of the graph show the results for a collection of templates generated with coarse increments of $2 \mu\text{s}$ in the two-glitter separation. The dashed lines in the lower part of the graph show the case of fine $0.2\text{-}\mu\text{s}$ increments.

low-pass filter; circles correspond to no halfway rectification and no low-pass filtering, i.e., the bandpass filters alone are used to create the time-frequency representation. The top figure shows the results for a collection of templates generated with increments of $2 \mu\text{s}$ in the two-glitter separation. The dashed lines correspond to the case of $0.2\text{-}\mu\text{s}$ increments. Notice that for high SNRs, the coarse resolution lines all start at $2 \mu\text{s}$, and the fine resolutions ones all start at $0.2 \mu\text{s}$, since those are the maximal resolutions of the coarse and fine template sets. The fine resolution model is more sensitive to noise than the coarse one. In fact, the differences between the templates in the coarse resolution experiment are more substantial. The templates for the fine resolution experiments are more similar to each other, and lower levels of noise can easily mask those differences. In particular, the break points in SNR appear to be shifted between the two models by approximately 10 dB for high COFs, to 15–20 dB for lower ones.

IV. DISCUSSION

The mammalian auditory system segments the frequencies in sounds into parallel frequency bands at the inner ear, smoothes the envelopes of the filtered signals, and transmits them in volleys of neural action potentials to higher brain centers for processing of acoustic information into auditory images. The inner ear of the little brown bat (*Myotis lucifugus*), a close relative of the big brown bat, contains about 900 inner hair cells (each comprising a single frequency-tuned channel), and the auditory nerve contains about 55 000 afferent fibers for transforming receptor activity into neural spikes (Ramprasad *et al.*, 1978). These spikes are transmitted to a cascade of higher auditory centers (brainstem to midbrain to cortex) for processing the acoustic information

they convey into images of objects. Nuclei at the first several stages of this cascade (brainstem) contain neurons numbered in thousands and tens of thousands with temporal and spectral response properties similar to the frequency tuning and time constants of the bandpass filters themselves. These neurons have to perform early auditory processing on the initial time-frequency representation so that it can be efficiently relayed to higher auditory areas. In contrast, neurons in the inferior colliculus (midbrain) and auditory cortex are numbered in millions and have poor temporal response properties. Only their frequency tuning resembles that of the bandpass filters, and even this is modified by inhibitory responses at frequencies adjacent to excitatory frequencies (Casseday and Covey, 1995; Pollak and Casseday, 1989). These areas are candidates for performing the 3D scene analysis from acoustic data delivered in a time/frequency format from the brainstem. It seems likely that one function of the very large numbers of higher-level neurons is to support a representation of the spectrum of echoes with different degrees of frequency resolution to accommodate coarse to fine local spectral shape around each frequency. It is also likely that these auditory areas perform the binding of time/frequency information into perceived objects, so that the fine delay separations extracted from the interference pattern in frequency come to be associated with each target's correct absolute range extracted in the time-domain from the overall delay of echoes. Echoes that return from different objects at different distances and directions on successive pings have to be integrated into objects that are stable across multiple pings. The time scale for this kind of integration across pings necessarily is much longer than the $300\text{-}\mu\text{s}$ integration time of the bandpass filters because it has to encompass the entire time interval from one ping to the next, which typically is 1 to 50

ms in bats. The temporal response properties of midbrain and cortical neurons are aligned with this longer epoch, implicating them in the interpretation and registration of stable acoustic scenes. The efficient relay of data from lower centers (brainstem) should convey information about the fine temporal structure of echoes while using low bandwidth channels to match the ascending information about each echo to the time scale of target scene analysis.

The echolocation of bats at ultrasonic frequencies requires perception of aspects of the detailed temporal structure of the waveform of echoes measured in microseconds, even though the neural channels for conveying this information upward to the auditory midbrain and cortex have much lower bandwidth and poor temporal response properties measured in milliseconds. Closely spaced echoes fall inside the approximately 300- μ s integration time of the inner ear's bandpass filters, so their time separation is transposed into an interference pattern of amplitudes (notches and flanking peaks) at different frequencies. Transmission of spectral shape in parallel frequency channels can be achieved at low data rates compared to direct transmission of the equivalent time separations, and the auditory system may have evolved to exploit this feature of time/frequency coding so that higher auditory centers can have the long time constants required to assemble acoustic scenes across successive pings and still be able to receive and use information extracted from each echo. We have found that a time/frequency representation can be used to break the high-frequency, broad bandwidth signals received by the bat into parallel signals of greatly reduced frequency and bandwidth to register the interference pattern of overlapping echoes in a time-frequency representation so that channels with biologically realistic low temporal resolution can still account for the bat's fine delay resolution. We have specifically found that adequate resolution is achievable with parallel channels limited by input smoothing to frequencies no higher than 7 kHz. This representation is very efficient, being suitable for higher cortical areas to perform additional calculations on the fine temporal structure using slow neurons, and as such offers a valuable approach to high-resolution sonar receiver design.

ACKNOWLEDGMENTS

We thank Leon N. Cooper and Quyen Huynh for many fruitful discussions. We also thank Daniele Gazzola who ran early simulations with the model. This work was supported by ONR Grant Nos. N00014-01-1-0306 and N00014-99-1-0350, as well as by a grant from the Burroughs-Wellcome Fund to the Brown University Brain Sciences Program.

Casseday, J. H., and Covey, E. (1995). "Mechanisms for analysis of auditory temporal patterns in the brainstem of echolocating bats," in *Neural Representation of Temporal Patterns*, edited by E. Covey, H. L. Hawkins, and R. F. Port (Plenum, New York), pp. 25–51.

Dalland, J. (1965). "Hearing Sensitivity in Bats," *Science* **150**, 1185–1186.

Dear, S. P., Fritz, J., Haresign, T., Ferragamo, M. J., and Simmons, J. A. (1993). "Tonotopic and functional organization in the auditory cortex of the big brown bat, *Eptesicus fuscus*," *J. Neurophysiol.* **70**, 1988–2009.

Griffin, D. R., Friend, J. H., and Webster, F. A. (1965). "Target discrimination by the echolocation of bats," *J. Exp. Zool.* **158**, 155–168.

Grinnell, A. D. (1995). "Hearing in bats: An overview," in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer-Verlag, New York), pp. 1–36.

Haplea, S., Covey, E., and Casseday, J. H. (1994). "Frequency tuning and response latencies at three levels in the brainstem of the echolocating bat, *Eptesicus fuscus*," *J. Comp. Physiol., A* **174**, 671–683.

Kalko, E. K. V., and Schnitzler, H.-U. (1998). "How echolocating bats approach and acquire food," in *Bat Biology and Conservation*, edited by T. H. Kunz and P. A. Racey (Smithsonian Institution, Washington, DC), pp. 197–204.

Kick, S. A. (1982). "Target detection by the echolocating bat, *Eptesicus fuscus*," *J. Comp. Physiol., A* **145**, 431–435.

Koay, G., Heffner, H. E., and Heffner, R. S. (1997). "Audiogram of the big brown bat (*Eptesicus fuscus*)," *Hear. Res.* **105**, 202–210.

Kössl, M., and Vater, M. (1995). "Cochlear structure and function in bats," in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York), pp. 191–234.

Matsuo, I., Tani, J., and Yano, M. (2001). "A model of echolocation of multiple targets in 3D space from a single emission," *J. Acoust. Soc. Am.* **110**, 607–624.

Menne, D., and Hackbarth, H. (1986). "Accuracy of distance measurement in the bat, *Eptesicus fuscus*: Theoretical aspects and computer simulations," *J. Acoust. Soc. Am.* **79**, 386–397.

Neuweiler, G. (2000). *The Biology of Bats* (Oxford U.P., New York).

O'Neill, W. E. (1995). "The bat auditory cortex," in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York).

Peremans, H., and Hallam, J. (1998). "The spectrogram correlation and transformation receiver, revisited," *J. Acoust. Soc. Am.* **104**, 1101–1110.

Pollak, G. D., and Casseday, J. H. (1989). *The Neural Basis of Echolocation in Bats* (Springer, New York).

Popper, A. N., and Fay, R. R. (1995). *Hearing by Bats* (Springer, New York).

Ramprasad, F., Money, K. E., Landholt, J. P., and Laufer, J. (1978). "A neuroanatomical study of the little brown bat (*Myotis lucifugus*)," *J. Comp. Neurol.* **178**, 347–363.

Ruggero, M. A. (1992). "Physiology and coding of sound in the auditory nerve," in *The Mammalian Auditory Pathway: Neurophysiology*, edited by A. N. Popper and R. R. Fay (Springer, New York), pp. 34–93.

Saillant, P. A., Simmons, J. A., Dear, S. P., and McMullen, T. A. (1993). "A computational model of echo processing and acoustic imaging in FM echolocating bats: The SCAT receiver," *J. Acoust. Soc. Am.* **94**, 2691–2712.

Schnitzler, H. U., and Henson, O. W. (1980). "Performance of airborne animal sonar systems: Microchiroptera," in *Animal Sonar Systems*, edited by R.-G. Busnel and J. F. Fish (Plenum, New York), pp. 109–181.

Simmons, J. A. (1980). "The processing of sonar echoes by bats," in *Animal Sonar Systems*, edited by R.-G. Busnel and J. F. Fish (Plenum Press, New York), pp. 695–714.

Simmons, J. A., Ferragamo, M. J., and Moss, C. F. (1998). "Echo-delay resolution in sonar images of the big brown bat, *Eptesicus fuscus*," *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12647–12652.

Simmons, J. A., Ferragamo, M. J., Saillant, P. A., Haresign, T., Wotton, J. M., Dear, S. P., and Lee, D. N. (1995). "Auditory dimensions of acoustic images in echolocation," in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York), pp. 146–190.

Simmons, J. A., Saillant, P. A., Ferragamo, M. J., Haresign, T., Dear, S. P., Fritz, J., and McMullen, T. A. (1996). "Auditory computations for biosonar target imaging in bats," in *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (Springer, New York), pp. 401–468.

Investigations of the precedence effect in budgerigars: Effects of stimulus type, intensity, duration, and location

Micheal L. Dent^{a)} and Robert J. Dooling^{b)}

Department of Psychology, University of Maryland, College Park, Maryland 20742

(Received 22 March 2002; revised 12 January 2003; accepted 13 January 2003)

Auditory experiments on the localization of sounds in the presence of reflections, or echoes, that arrive later and from different directions are important to understanding hearing in natural environments. The perceived location of the auditory image can change with the time delay between the presentations of a leading and lagging sound. These changes in perceived location, encompassing the precedence effect, have been examined behaviorally or physiologically in humans and a number of animals. Here, these results are extended to include budgerigars. Behavioral methods were used to measure the discrimination performance between a stimulus presented at + and -90° azimuth with a delay (left-right), from the same two stimuli presented with the opposite delay (right-left). At short delays, where humans experience summing localization, budgerigars have difficulty discriminating between the two presentation types. With increasing delays, where humans experience localization dominance, budgerigars show improved discrimination performance. At even longer delays, where echo thresholds are found in humans, discrimination performance worsens again. The shapes of the discrimination functions are affected by the intensity, locations, and durations of the stimuli, and are subject to a buildup effect. These results show that budgerigars exhibit the phases of the precedence effect similar to humans and other animals. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1558391]

PACS numbers: 43.80.Lb, 43.66.Gf, 43.66.Qp [WWA]

I. INTRODUCTION

The precedence effect is an important aspect of binaural hearing and has a long history in both psychoacoustic and neurophysiological experiments (e.g., Blauert, 1997; Litovsky *et al.*, 1999; Wallach *et al.*, 1949). Previous work has shown that the precedence effect has three critical time periods, based on judgments of the perceived location of the stimulus as a function of the interstimulus delay between two sounds and these are called summing localization, localization dominance, and echo threshold [see review in Litovsky *et al.* (1999)]. The echo threshold is defined as the ISI (interstimulus interval between a lead and a lag) past which an observer hears both the lead and lag stimuli at separate locations (Blauert, 1997; Freyman *et al.*, 1991). The echo threshold is the longest temporal interval relevant to the precedence effect, and, above this delay, the precedence effect is no longer evident. Within certain limits below the echo threshold, stimuli emanating from two different speakers produce the perception of a single stimulus coming solely from the position occupied by the leading speaker. The range of ISIs over which this perception occurs defines localization dominance. Finally, summing localization occurs when the ISI is even shorter than in localization dominance. Work with humans shows that during this time period, a phantom source is perceived as being somewhere between the two sound sources. The exact values of the time ranges described in the above three pieces of the precedence effect depend on the

types of stimuli used (e.g., Schubert and Wernick, 1969), the instructions given to the listener (Blauert, 1997; Zurek, 1987), the location of the lead and lag stimuli (Litovsky and Shinn-Cunningham, 2001), and the immediately preceding history of the auditory information presented to the listener (Clifton, 1987). Taken together, these results suggest that the precedence effect is not a simple auditory echo suppression mechanism, but a more general auditory mechanism that improves hearing in natural situations.

Comparative studies of the precedence effect are much more difficult to conduct than those in humans because animals cannot be directly queried about the location of real or phantom targets in space. Behavioral and physiological studies, however, have both experienced some degree of success by using a wide variety of creative techniques to answer the question of underlying mechanisms of the auditory system. Litovsky, Yin, and colleagues (e.g., Litovsky and Yin, 1998a, b; Yin, 1994), for example, have conducted an extensive set of experiments on aspects of the precedence effect in cats. They have demonstrated some interesting underlying neural correlates to localization dominance, summing localization, and echo thresholds in the inferior colliculus of kittens and adult cats (*Felis catus*) and have compared these results with some of the behavioral findings in humans.

The behavior and underlying neural correlates of the precedence effect have also been demonstrated in several experiments with albino rats (*Rattus norvegicus*). Kelly (1974), for example, showed through behavioral procedures that discrimination between background click pairs from the right then left (R-L) and target click pairs from the left then right (L-R) with the same-but-opposite ISI occurred when one click in the pair was delayed or reduced in intensity over

^{a)}Current address: Department of Physiology, University of Wisconsin-Madison, Madison, WI 53706. Electronic mail: dent@physiology.wisc.edu

^{b)}Author to whom correspondence should be addressed. Electronic mail: dooling@psyc.umd.edu

a certain range. Discrimination was worse at very short and very long delays, as in the cats. Generally, results from independent laboratories using different methodologies suggest that mammals do indeed perceive summing localization and localization dominance stimuli in a manner similar to humans [see review in Litovsky *et al.* (1999)].

Correlates of the precedence effect have been found in one bird species to date, the barn owl (*Tyto alba*). Keller and Takahashi (1996; see also Takahashi and Keller, 1994) measured responses to simulated echoes by neurons in the external nucleus of the inferior colliculus of the barn owl and correlated those responses to behavioral data. Keller and Takahashi (1996) found these birds perceived sounds presented simultaneously from two lateral speakers as coming from a central location in space in behavioral experiments. In these experiments, the barn owl's echo thresholds were found to be about 5 ms (Keller and Takahashi, 1996; Takahashi and Keller, 1994), similar to those found by Tollin *et al.* (2000) in the cat, and those found in humans (e.g., Blauert, 1997). While cats and barn owls have similar head sizes (and therefore similar interaural sound localization cues), they have extremely different evolutionary histories, life styles, and auditory neuroanatomy (Carr, 1992). Thus, it is interesting that their echo thresholds are so similar.

Measuring the precedence effect in barn owls is also interesting because, like cats, they are nocturnal predators with highly specialized auditory systems designed for tracking and capturing prey. It was not known whether other smaller birds without obvious external ear or central nervous system adaptations for sound localization would exhibit the precedence effect or whether the timecourses of summing localization, localization dominance, and echo thresholds would be similar across different bird species. This is a problem of some intrigue because small birds have small heads with closely spaced ears. These anatomical characteristics result in small interaural time and intensity differences, which would seem to work against a robust precedence effect.

The present study sought to investigate the precedence effect in the budgerigar (*Melopsittacus undulatus*), a small parrot and popular cagebird. It was unclear whether these birds would exhibit a robust precedence effect or how it would differ from that demonstrated in humans, other animals, and specialized birds such as barn owls. Budgerigars were chosen, in part, because they are well studied in the psychoacoustics field. More is known about hearing and auditory discriminations in this species, in both quiet and noisy backgrounds, than in any other bird (see, for example, Dooling *et al.*, 2000).

The capabilities of the budgerigar's binaural auditory system have also been investigated and show some intriguing results. For instance, absolute localization thresholds for these birds are not remarkable [about 20° for the budgerigar compared to 2° for the barn owl; review in Klump (2000)]. This is slightly, but not dramatically, better than what one would expect from their small heads and closely spaced ears. However, budgerigars do exhibit the phenomenon of binaural release from masking [the improvement in threshold when separating the signal from the background noise in

space (Dent *et al.*, 1997)]. Masked thresholds for a tone embedded in noise emitted from one speaker on one side (e.g., right side tone, right side noise) of the birds were much worse than thresholds for a tone embedded in noise emitted from two speakers on two sides (e.g., right side tone, right and left side noise) of the birds. Moreover, the amount of binaural release from masking was similar to the amount seen in humans with much larger heads (Dent *et al.*, 1997).

One possible mechanism for this enhancement in free-field binaural masking release is the existence of the interaural canal in these and other birds [review in Klump (2000)]. Using laser vibrometry, it has been shown that the phase response of one tympanum is affected by sound traveling through the interaural pathway from the opposite ear, providing a potential mechanism for improving directional hearing, especially at low frequencies (Larsen *et al.*, 1997). The pressure-difference receiver can only aid in sound localization if the sound pressure level of the sound to the inner tympanum is not attenuated too much relative to the sound pressure level of the sound to the external surface of the tympanum [see review in Klump (2000)]. In budgerigars, especially at low frequencies, the attenuation of sounds through the interaural canal range from 2 dB at 0.5 kHz to 17 dB at 2 kHz (Larsen *et al.*, 1997). The experiments here attempted to determine whether budgerigars could enhance interaural timing information of primary sounds versus echoes using the interaural canal (or another) mechanism.

The results in humans suggesting that binaural time or intensity difference cues are not necessary for localization dominance or summing localization (Litovsky *et al.*, 1997; Rakerd and Hartmann, 1994; Rakerd *et al.*, 2000) are interesting with respect to budgerigars as well. It is known that (in humans at least) each position in space is associated with characteristic peaks and troughs in the spectrum of sounds and that human listeners can use this information for sound localization (Hebrank and Wright, 1974; Searle *et al.*, 1976). The head, pinnae, and body interacting with the sound at each position in space create these spectral peaks and troughs. It was not known how many cues these small birds with no pinnae might generate or if they would be able to exhibit the precedence effect with minimal binaural interaural time and intensity cues, so the precedence effect was also determined along the median sagittal plane. The results from these experiments should provide interesting insights as to the importance of the above-mentioned binaural cues for locating an object in space. The following experiments define the timecourses for echo thresholds, localization dominance, and summing localization in budgerigars and compares these values to those found in humans and other animals.

II. GENERAL METHODS

A. Subjects

Four adult budgerigars (three males, one female) were used in these experiments. The birds were either bred from commercial stock in a vivarium at the University of Maryland or purchased from a local breeder. The birds were housed in individual cages and kept on a normal day/night cycle correlated with the season at approximately 90% of

their free-feeding weights. All animal experimentation was conducted under the auspices of an approved protocol from the Animal Care and Use Committee at the University of Maryland, College Park.

B. Testing apparatus

The psychoacoustic experiments took place in a wire test cage ($25 \times 18 \times 14 \text{ cm}^3$) mounted 115 cm from the ground in a sound-attenuated chamber ($2.8 \times 2.5 \times 2.0 \text{ m}^3$). The test cage consisted of a perch, an automatic feeder on the floor of the cage (food hopper), and two vertical response keys extending from the floor in front of the bird. The response keys were two sensitive microswitches with 8-mm light emitting diodes (LEDs) located just above the food hopper. The bird pecking the LED tripped the microswitch. The left microswitch and LED served as the observation key, and the right microswitch and LED served as the report key. During a session, a small light at the top of the test cage illuminated the chamber. The behavior of the animals during test sessions was monitored at all times by an overhead video camera system.

The experiment was controlled by an IBM Pentium III microcomputer operating Tucker-Davis Technologies (TDT, Gainesville, FL) modules. Stimuli were generated in advance of testing, stored in digital form, and output at a sampling rate of 50 kHz via a timing generator (TDT, Model TG6) to a four-channel D/A converter (TDT, Model DA3-4). Each signal was then output from a separate channel of the D/A converter to a separate digital attenuator (TDT, Model PA4) and a separate amplifier (TDT, Model HB6) to separate speakers located in the testing chamber (Realistic 3" midrange tweeter). Two speakers were placed at the same height as the cage, 60 cm from the position of the bird's head during testing. Stimulus calibration was performed with a Larson-Davis sound level meter (Model 825, Provo, UT) and 20-ft extension cable. For all measurements, a $\frac{1}{2}$ -in. microphone was placed in the position normally occupied by the bird's head during testing. Additional measurements of the spectral and temporal characteristics of the stimuli were made by passing the output of the sound level meter to a Stanford Research Systems FFT spectrum analyzer (Model SR760, Sunnyvale, CA) and a Tektronix digital oscilloscope (Model TDS3014B, Portland, OR).

C. Training and testing procedures

The birds were trained by a standard operant auto-shaping program to peck at the microswitch keys for food reinforcement. First, they pecked repeatedly at the left key (observation key) during repeating presentations (rate of 2/s) of a stimulus emitted from single source (background; e.g., a left speaker only). After a random interval of 1–6 s, the same stimulus was played from a different source (target; e.g., from the right speaker). The bird was trained to peck the right key (report key) when it detected this change in order to obtain food. Percent correct hit values on trials involving a change from background to target were recorded for further analysis.

Incorrect report key pecks were punished with a time-out during which all of the room lights were extinguished. A miss was recorded if the bird failed to peck the report key within 2 s of the change in stimulus presentation. Sham trials, where there was no change of sound presentation type from background to target, consisted of 30% of all trials. Pecks to the report key during these trials were counted as false alarms, and the birds were again punished with a lights-out period. Sessions with a false alarm rate of 15% or higher were discarded. Approximately 6% of all sessions were discarded for this reason.

An experimental session consisted of 100 trials. The birds were tested at least two sessions a day, and between 5 and 7 days a week. Each bird was run on at least 200 trials at each experimental condition, and the last 100 trials after percent correct values stabilized were analyzed. To minimize any response biases, two of the birds were trained to repeating background sounds from the left speaker, and two were trained to repeating background sounds from the right speaker.

III. EXPERIMENT 1: THE EFFECTS OF STIMULUS TYPE

The exact values of the time ranges of summing localization, localization dominance, and echo thresholds depend on the types of stimuli used (e.g., Schubert and Wernick, 1969). In humans, different stimuli give very different echo thresholds [see reviews in Blauert (1997) and Litovsky *et al.* (1999)]. The onset durations, overall durations, and carrier frequencies of the stimuli can all have an influence on the timecourses of the precedence effect (Houtgast and Aoki, 1994; Rakerd and Hartmann, 1986; Yang and Grantham, 1997). In this first experiment, we tested the birds' discrimination of pairs of broadband noises (0.1, 1, and 50 ms in duration), as well as recorded natural contact calls of the budgerigar, at different ISIs.

The birds were also tested on discrimination of stimulus pairs with different lead locations. In humans, echo thresholds are higher (later in time) when the lead originates from the right and the lag from the left than the reverse condition. One of the arguments for this asymmetry is that overall hearing abilities are asymmetrical in humans (e.g., Ward, 1957; Emmerich *et al.*, 1988), so the suppression of echoes is more efficient when the direct source is presented to the better ear (Grantham, 1996). Asymmetrical hearing abilities have never been found in birds, which made them interesting subjects for a study of asymmetries in localization dominance and echo thresholds.

The budgerigar experiments here are modeled after earlier discrimination experiments on rats. Kelly (1974) tested rats on their ability to discriminate between paired-background sounds presented from a left then a right speaker (L-R) from paired-target presentations of sounds from a right then left speaker (R-L) with the same ISI between sound presentations. The rats had difficulty distinguishing between the two presentation types at extremely short delays (background versus target; 31–62 μs), where summing localization is operating. Presumably, the rats perceived only one

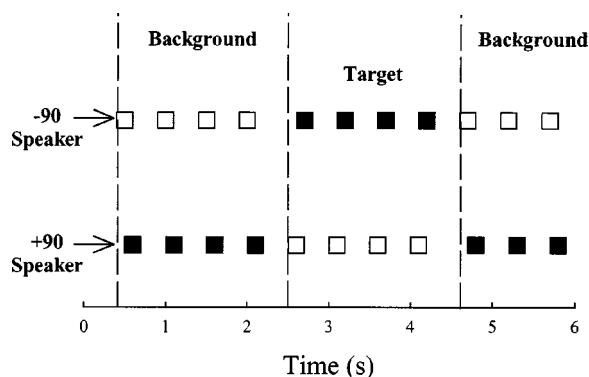


FIG. 1. Schematic of trial events. Stimuli are presented at a rate of 2/s. Stimuli are presented from the -90° speaker first, followed, after a short delay, from the $+90^\circ$ speaker. Background stimuli are presented a random number of times (6–12) before targets are presented. Targets have the same delay between lead and lag speakers, they are merely reversed in position. Birds have four presentations of the target to respond before the repeating background continues again for the next trial.

sound as coming from somewhere in the middle for both background and target presentation types (the phantom sound image) so they could not distinguish between the background and targets. The rats could easily discriminate the L-R pair from the R-L pair at presentation delays from 0.25 to 16.0 ms, where localization dominance is operating. The rats could distinguish the left-first from the right-first stimuli because they only perceived the sounds as coming from one speaker location in the background (the leading speaker position). Therefore, the perceived left-only background sound was easy to discriminate from the target. The rats became unable to distinguish between the two presentation types again with much longer delays (20–32 ms). This is above the echo threshold for the rats; they heard both sounds from separate locations during both the background and the target presentations. The sounds were too close together for rats to be able to tell which sound came first, however, so they could not discriminate between the background and target presentations. The first experiment was designed to measure whether budgerigars would exhibit the precedence effect using similar methods, what the timecourses of the components of the precedence effect were, and the effect of using different stimuli.

A. Methods

1. Procedures

Once the birds were trained to peck repeatedly during presentation of a single stimulus emitted from the left speaker (-90° , repeating background) and to report a change in the location of this repeating background ($+90^\circ$ target), paired stimuli were introduced into the experiments. Here, the repeating background involved presenting the stimuli from two speakers as in Kelly's (1974) experiment with rats (see Fig. 1). In other words, the background was a repeating stimulus involving the left then the right speakers, with very short delays between the two sound presentations. In each session, the target stimulus pairs had the same (but opposite)

ISI between presentations as the repeating background stimulus pairs. The repeating background (L-R) was presented a random number of times (between 6 and 12) at a rate of 2/s before the targets (R-L) were presented. Target pairs were presented four times total (2 s), or until the birds responded. No sounds were played during the reinforcement period. Only one ISI was tested in each experimental session. Testing of the different ISIs was varied randomly between experimental sessions. These ISIs ranged from 0.1 to 40 ms (although some different ISIs were used in the different stimulus conditions), and a total of at least 13 ISI values were tested for each stimulus type.

All four subjects were tested on L-R versus R-L discriminations at different ISIs with 1-ms broadband noise pulses (clicks) to determine the nature of the precedence effect in budgerigars. To examine the effect of stimulus duration, two subjects (Penny and Will) were also tested on 0.1-ms clicks and 50-ms noise bursts. To assess the precedence effect using the natural vocalizations of these birds, three of the subjects (Penny, Will, and Spike) were also tested on three recorded contact calls from three different budgerigars.

Finally, two subjects (Cirrus and Will) were tested across all ISIs on both the left-first and the right-first click conditions, in a random order. They were also tested on the left-first clicks where the testing cage was rotated 180° within the chamber. This experimental condition measured the effect of the location of the lead stimulus on the precedence effect. In humans there is an asymmetry in the buildup of localization dominance depending on which side the leading sound is emitted (Clifton and Freyman, 1989; Grantham, 1996), and we could test for a similar phenomenon in birds. Results should also reveal whether any spectral artifacts or cues used for the precedence effect were due to the experimental setup.

2. Stimuli

All of the stimuli were presented at peak sound pressure level of 60 dB(A) SPL measured at the bird's head. The clicks had a 0-ms rise-decay time and the 50-ms noise bursts had a 5-ms rise-decay duration. The longer stimuli also generally contained more energy at higher frequencies than the shorter stimuli. This is unlikely to be a significant factor since the hearing abilities of these decreases dramatically above about 5 kHz. The contact calls were typical of budgerigar calls (see Farabaugh and Dooling, 1996): frequency modulated patterns with most of their energy between 2 and 4 kHz, durations of between 120 and 160 ms, and with onset times of approximately 5–10 ms.

As a control for potential monaural intensity fluctuations in the stimuli that the budgerigars may have used as cues for discrimination between background and target click pairs, the 1-ms click pairs were also tested with the intensity roved ± 4 dB (around 60 dB) from presentation to presentation. Two of the budgerigars (Penny and Spike) were tested at all ISIs in this control experiment, and results were compared to their results from the 1-ms click pair discrimination experiment where the intensities were not roved.

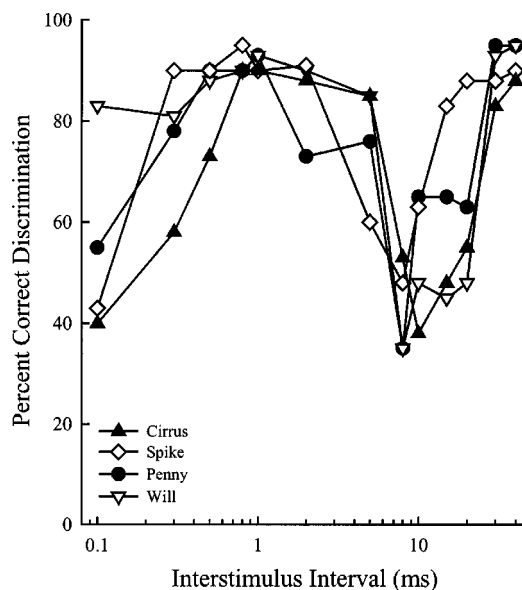


FIG. 2. Results from four individuals tested on the discrimination of 1-ms click pairs. Percent correct discrimination values are shown as a function of ISI.

B. Results

Discrimination for the single clicks used in training (left only background versus right only targets) was close to 100% for all four birds. Percent correct values for paired stimuli tested at different ISIs are shown in Fig. 2 for the four individual budgerigars tested on 1-ms clicks. Here, the task was to discriminate a L-R repeating background from a R-L target, at different ISIs ranging from 0.1 to 40 ms. Two of the birds were tested in the above-mentioned condition, with the L-R background and R-L targets (Cirrus and Spike), and two of the birds were tested in the opposite condition, with the R-L background and L-R targets (Will and Penny). For all four subjects, percent correct discrimination increased, decreased, and then increased again as ISIs increased. The four birds deviated by 3% to as much as 43% from each other in their discrimination values at the different ISIs, but shapes of the curves were similar across birds. As one check on whether the variation across birds was due to subject differences in perceptual processes or to subject differences in the adaptation to the testing procedures, a d' analysis was used. The average d' value across the four birds at all ISIs for this experiment was 2.32 and the range of d' values across birds was 2.05 to 2.53. The interindividual variation in the data shown in this experiment is fairly small compared to similar experiments in humans [see, for example, Litovsky and Shinn-Cunningham (2001)].

A one-way repeated measures analysis of variance (ANOVA) showed that there were significant differences across ISIs for all subjects [$F(12,36)=9.15$, $p<0.001$]. At intermediate delays (0.5 to 5 ms), all birds performed well—they easily discriminated the R-L target from the L-R background. *Posthoc* Bonferroni t -tests showed that the intermediate delays of 0.8, 1.0, and 2.0 ms were all significantly different from the shorter delay of 0.1 ms and the longer delays of 8, 10, and 15 ms ($p<0.05$). The percent correct

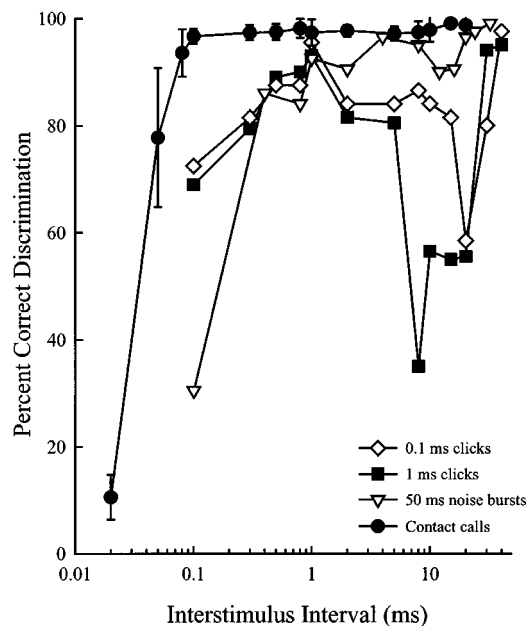


FIG. 3. Results from two individuals tested on the discrimination of three types of broadband stimulus pairs, and from three individuals tested on discrimination of budgerigar contact calls. Percent correct discrimination values are shown as a function of ISI.

values at these intermediate ISIs were significantly higher than those at the longer and shorter delays.

At shorter delays, percent correct values in the budgerigars for the 0.1 ms ISI were significantly lower than those from the 0.5-, 0.8-, 1.0-, 2.0-, 30.0-, and 40.0-ms intervals ($p<0.05$), and the 0.3- and 0.5-ms ISIs were significantly different from the 8.0-ms ISI ($p<0.05$). At the longer delays, the lowest percent correct values (at 8 ms) were significantly different from those at shorter ISIs (0.3, 0.5, 0.8, 1.0, 2.0, and 5.0 ms), and both the 8- and 10 ms-ISIs were significantly different from some longer ISIs (30 and 40 ms; $p<0.05$). These results, as a whole, establish that budgerigars do exhibit the phenomena of the precedence effect, including summing localization, localization dominance, and echo thresholds.

Two subjects were tested with the intensities of the stimuli roved by ± 4 dB from presentation to presentation. The discrimination functions were not significantly different across ISI between the roved and unroved conditions for either bird as shown by a paired t -test [Penny: $t(12)=0.46$, $p>0.05$; Spike: $t(12)=0.65$, $p>0.05$]. The birds could not reliably use intensity difference cues to discriminate background from target in the roved condition, suggesting other mechanisms were shaping the discrimination functions.

To test the effects of stimulus duration on discrimination abilities, two budgerigars were additionally tested on stimulus pairs where the stimuli were 0.1 and 50 ms in duration (Fig. 3). The results for all three broadband stimuli are similar. Discrimination was low at the shortest ISIs, high at intermediate ISIs, and (somewhat) lower again at longer ISIs. At the shortest ISIs, the 0.1-ms click pairs were the easiest to discriminate, followed by the 1-ms clicks, and then the 50-ms noise bursts. The peak of responding (i.e., where localization dominance might be at a maximum) was at a simi-

lar I SI across stimulus types: at 1 ms for the 0.1-ms clicks, at 1 ms for the 1-ms clicks, and at 4 ms for the 50-ms noise bursts. Echo thresholds, or the lowest point of responding (lowest percent correct discrimination values), were at 20 ms for the 0.1-ms stimuli, at 8 ms for the 1-ms stimuli, and at 12-ms for the 50 ms stimuli.

The ISIs used for the three click conditions were not all exactly the same, so statistics were completed on the six ISIs that were tested with all three stimulus types (representative of short through long ISIs): 0.1, 0.8, 1.0, 2.0, 8.0, and 20.0 ms. A two-way repeated measures ANOVA (stimulus type \times ISI) showed that there were significant differences across stimulus type ($F(2,17)=3.86, p<0.05$), across ISIs ($F(5,17)=9.54, p<0.001$), and a significant interaction between stimulus type and ISI ($F(10,17)=6.79, p<0.001$). *Posthoc* Bonferroni *t*-tests showed that at the shortest ISIs of 0.1 ms, the 50-ms stimulus was significantly different from the other two shorter stimuli ($p<0.05$), but the two shorter stimuli were not significantly different from each other.

Average results for three individuals tested on three different contact calls (calls from three different budgerigars) across ISIs are also shown in Fig. 3. Discrimination between these call pairs was high from 0.1 to 100 ms, so ISIs down to 0.02 ms were tested. The three birds were very similar in their discrimination abilities across ISIs and across contact calls. The contact calls varied in peak frequency, amounts of amplitude modulation, and duration, yet performance was similar for all three calls. These results show that contact calls can also elicit at least some of the components of the precedence effect.

The average percent correct discrimination functions across ISIs were compared for the left-first condition, the right-first condition, and with the cage rotated 180° within the testing chamber. Discrimination was almost identical regardless of which side the leading stimulus was located. A two-way (speaker location \times ISI) repeated measures ANOVA showed that there was a significant effect of ISI [$F(10,32)=6.09, p<0.001$], but there were no significant differences between the speaker locations [$F(2,32)=0.86, p>0.05$] or an interaction effect between the two variables [$F(20,32)=0.19, p>0.05$]. Two important conclusions can be made from this experimental condition. First, the testing chamber is acoustically symmetrical, as seen when comparing the almost-identical right first and cage rotated 180° discrimination curves. Second, the location of the leading stimulus does not affect the timecourses of the precedence effect in budgerigars.

Average data from the four budgerigars using 1-ms clicks are plotted along with data from cats, rats, and humans in Fig. 4. The methodology across studies varies widely, but the overall shapes of the percent correct functions across ISIs can be compared in a very general sense. The rats were required to discriminate R-L from L-R clicks (Kelly, 1974), similar to the budgerigar task. The other rat data were from Hoeffding and Harrison (1979), where rats were required to identify the location of a leading source. Percent correct data for the cat gives the proportion of trials the cats correctly identified the location of the leading source in a pair of clicks (Cranford, 1982). Percent correct values for the humans give

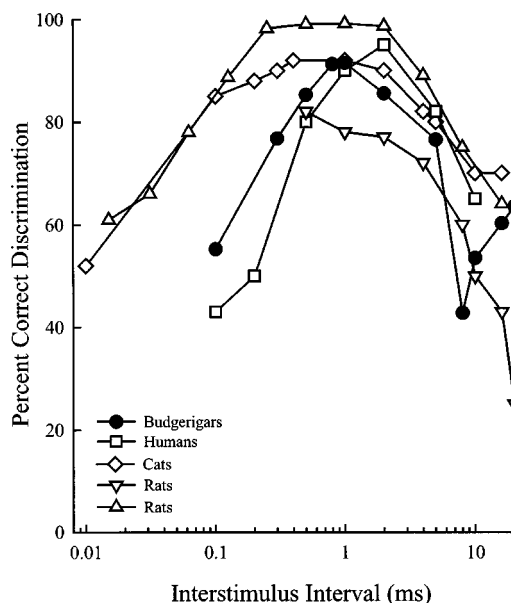


FIG. 4. Discrimination functions for budgerigars (1-ms clicks, this study), humans [0.025 ms clicks (Litovsky *et al.*, 1997)], cats [0.05-ms clicks (Cranford, 1982)], rats [triangles, 0.05-ms clicks (Kelly, 1974)], and rats [upside-down triangles, 0.05-ms clicks (Hoeffding and Harrison, 1979)].

the proportion of trials that a sound image was identified at the position of the leading source (Litovsky *et al.*, 1997). For all experiments, a high percent correct value indicates that localization dominance was operating; in other words, the location of the lead was easily discriminated or identified. For all species, localization dominance was at a maximum at delays of 0.5–5.0 ms and at a minimum at longer and shorter ISIs.

IV. EXPERIMENT 2: THE EFFECTS OF FRONT VERSUS REAR STIMULUS LOCATION

This experiment tested whether the precedence effect occurred on the midline. In humans, the precedence effect occurs on the median sagittal plane at similar timecourses as are found in the azimuthal plane (Litovsky *et al.*, 1997; Rakerd and Hartmann, 1994; Rakerd *et al.*, 2000). To determine if the binaural properties of the sound were important or necessary for the precedence effect in budgerigars, similar procedures as described above were utilized in the median plane.

A. Methods

This experiment used same apparatus, 1-ms click stimuli, and procedures as in Experiment 1. Three of the budgerigars from experiment 1 (Penny, Spike, and Will) were used in this experiment. In this experiment, the lead and lag speakers were placed at 0° and 180° (directly in front of and directly behind the budgerigars at eye/ear level). The sounds were played from each speaker in quick succession during the repeating background (front then back with a short delay between presentations), and the order of presentation was reversed during the target trials (to back then front). Click delays ranging from 0.1 to 40 ms were tested in a random order, and results were compared to the results from these subjects tested in the left-right speaker condition

in experiment 1. Although the budgerigars were free to move their heads during the course of these experiments, the requirement that the observation key be pressed down to initiate the presentation of the target stimulus pairs inhibited head rotation dramatically (± 1 cm to the right or left).

B. Results

The variance between the three subjects in this experiment was much larger for the front back discrimination than for the left right discrimination at most ISIs, especially at the longest ones (above 2 ms). A two-way (speaker position \times ISI) repeated measures ANOVA showed that there were significant effects of speaker position [$F(12,52)=3.06$, $p<0.01$] and ISI [$F(1,52)=8.29$, $p<0.005$], but no significant interaction between the two variables [$F(12,52)=0.63$, $p>0.05$]. Overall, percent correct values were significantly higher for the left-right speaker condition than for the front-back speaker condition. For both conditions, there was a significant effect of ISI (as in previous experiments). A lack of a significant interaction between the two variables suggested that the shapes of the two functions were not significantly different. Discrimination of front-back stimuli was equivalent to that of left-right discrimination.

V. EXPERIMENT 3: THE EFFECTS OF STIMULUS SEPARATION DISTANCE

In humans, there are no differences in echo thresholds with changes in spatial separation of a lead and lag source (Litovsky and Shinn-Cunningham, 2001). However, localization dominance is stronger in stimuli that have smaller interaural time differences (closer together in space) than those with larger interaural time differences [further apart in space (Shinn-Cunningham *et al.*, 1993; Litovsky and Shinn-Cunningham, 2001)]. This experiment tested whether the timecourses of the components of the precedence effect changed with decreased speaker separation distance in budgerigars.

A. Methods

This experiment used same apparatus, 1-ms click stimuli, and procedures as in experiment 1. Three of the budgerigars from experiment 1 (Cirrus, Will, and Penny) were used in this experiment. In this experiment, the lead and lag speakers were at $+30^\circ$ and -30° or at $+60^\circ$ and -60° . Click delays ranging from 0.1 to 20 ms were tested in a random order under both speaker separation conditions, and results were compared to the results from these subjects tested in the $\pm 90^\circ$ speaker condition in experiment 1.

B. Results

The results for the three budgerigars tested on three stimulus separations are shown in Fig. 5. Discrimination performance was generally highest in the condition with the largest separation of lead-lag stimulus pairs. A two-way (stimulus separation \times ISI) repeated measures ANOVA showed that there was a significant effect of ISI [$F(10,53)=7.14$, $p<0.0001$], stimulus separation [$F(3,53)=17.57$, $p<0.0001$], and a significant interaction between the two

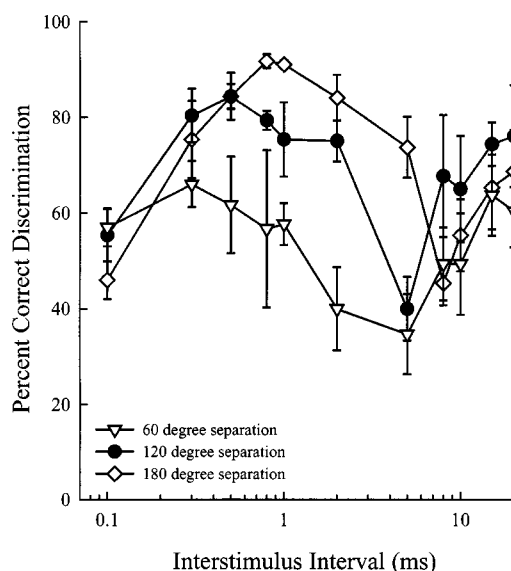


FIG. 5. Average discrimination functions for three budgerigars tested on discrimination of 1-ms click pairs as a function of ISI with the stimuli presented from $\pm 30^\circ$, 60° , and 90° . Error bars represent between-subject standard errors.

variables [$F(30,53)=2.02$, $p<0.05$]. *Posthoc* Bonferroni *t*-tests found some differences in percent correct discrimination values across stimulus separations. The 30° and 90° speaker locations were significantly different at the intermediate ISIs of 0.8, 1, 2, and 5 ms ($p<0.05$). The 30° and 60° speaker locations were significantly different at the 2-ms ISI only ($p<0.05$). The 60° and 90° speaker locations were significantly different at the 5-ms ISI only ($p<0.05$). The echo thresholds across the three speaker conditions ranged from 5 ms in the 60° and 120° separation conditions to 8 ms in the 180° separation condition. Generally, at the timecourses where localization dominance was operating, discrimination performance was higher when the speakers were further apart than when they were closer together.

VI. EXPERIMENT 4: THE EFFECTS OF STIMULUS INTENSITY

In humans, the precedence effect decreases at very low sensation levels (45 to 10 SL; Goverts *et al.*, 2000). However, Shinn-Cunningham *et al.* (1993) found only a small effect on localization dominance when increasing the stimulus levels from 80 to 110 dB. In two early studies with headphones in humans, an increase in click sensation level resulted in a decrease in echo thresholds (Babkoff and Sutton, 1966; Schubert and Wernick, 1969). The effect of stimulus level on the timecourses of localization dominance and echo thresholds have not been well examined in behavioral studies with animals. However, in cats, suppression decreased in most neurons in the inferior colliculus with increasing stimulus level (Litovsky and Yin, 1998a, b). Here, the birds were tested on 1-ms click stimuli at three different intensities to see if there was an effect of the intensity of the stimuli on the precedence effect.

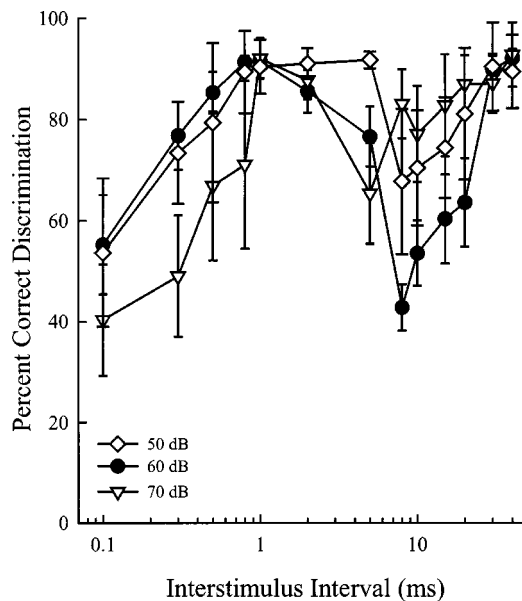


FIG. 6. Average discrimination functions for three budgerigars tested on discrimination of 1-ms click pairs as a function of ISI with the stimuli presented at three stimulus intensities. Error bars represent between-subject standard errors.

A. Methods

This experiment used the same apparatus, 1-ms click stimuli, and procedures as in experiment 1. Three of the budgerigars from experiment 1 (Cirrus, Penny, and Will) were also used in this experiment. Click-pair discriminations across ISIs were measured with the stimuli presented a level of 50 and 70 dB(A) SPL and compared to the results with 1-ms clicks measured at 60 dB(A) SPL in experiment 1.

B. Results

The average discrimination functions across ISIs for three different stimulus intensities are shown in Fig. 6 for three subjects. A two-way (stimulus intensity \times ISI) repeated measures ANOVA showed that there was a significant effect of ISI [$F(12,76)=6.84, p<0.001$], stimulus intensity [$F(2,76)=4.67, p<0.05$], and a significant interaction between the two variables [$F(24,76)=2.05, p<0.01$]. *Posthoc* Bonferroni *t*-tests found some differences in percent correct discrimination values across stimulus intensities at both the long and short ISIs. At 0.1 ms, the 70 dB percent correct values were significantly different from those at 60 dB ($p<0.05$). At 0.3 ms, both the 50 and 60 dB percent correct values were significantly different from those at 70 dB ($p<0.05$). At these short ISIs, where summing localization is operating in humans, the stimuli presented at 70 dB SPL were more difficult to discriminate than those at the lower intensities.

Differences arose at the longer ISIs as well. At the 5 ms ISI, the 50-dB response levels were significantly higher than those at 70 dB ($p<0.05$). At the 8–10-ms ISIs, the discrimination between the 60- and 70-dB stimuli were also significantly different ($p<0.05$). Performance for the 70-dB stimuli dropped to its lowest point at 5 ms while per-

formance did not drop to its lowest point for the 50- and 60-dB condition until 8 ms. These results show that echo thresholds occur later for less intense stimuli.

VII. EXPERIMENT 5: BUILDUP OF THE PRECEDENCE EFFECT

In humans, several repetitions of the pairs of lead-lag stimuli cause a “fading out” of the perceived location of the echo (Clifton and Freyman, 1989). This buildup of suppression in humans also occurs over a longer time period for longer ISIs, increasing the echo threshold with each presentation (Clifton and Freyman, 1989). The buildup and breakdown of the suppression of echoes has only been behaviorally found in one animal, the cat (Kalmykova, 1993). Kalmykova found that echo thresholds in cats increased from 12 to 17 ms with repeated lead-lag presentations. This is an interesting problem, but the neural substrates for this buildup effect have not yet been found (Litovsky and Yin, 1998a). This experiment tested whether localization dominance builds up for budgerigars in a manner similar to the way it builds up in humans.

A. Methods

This experiment used the same apparatus, 1-ms click stimuli, and procedures as in experiment 1. Three subjects from experiment 1 (Cirrus, Penny, and Will) were also used in this experiment. The birds’ percent correct discrimination values following different numbers of background stimulus-pair presentations (ranging from 1 to 12 presentations at the 2/s rate) and at different ISIs (1, 2, and 5 ms) were measured. Because the number of background presentations had to be kept random from trial to trial to ensure validity in the psychophysical data from the birds, at least 600 total trials were collected from each bird at each ISI. Then, hits and misses for each number of background presentations at each ISI were counted. This resulted in approximately 60 trials (± 10 trials) for each data point for each bird. As a control, the birds were tested on their discrimination of a single click from the left as the background and a single click on the right as the target after different numbers of background presentations (i.e., a simple localization task). This test determined whether the birds had a general tendency to respond correctly after an increased number of background presentations or whether the effect was unique to the precedence effect.

B. Results

Average percent correct values for three budgerigars across different numbers of background presentations are shown in Fig. 7. Each bird was tested on three ISIs of paired stimuli, with between 1 and 12 backgrounds presented before the targets were presented, and the control condition. In the control condition, percent correct discriminations were high across all number of background presentations. This was not the case for the paired click stimuli presented at different ISIs, however. For all three birds at all three ISIs, discrimination increased with increasing number of background pre-

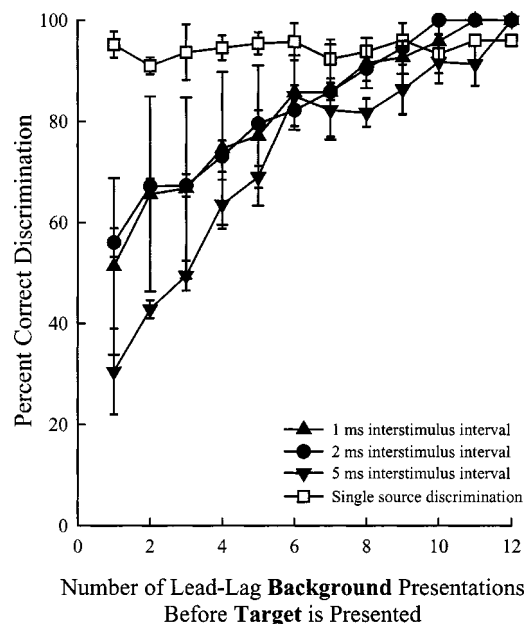


FIG. 7. Average of three individuals tested on discrimination of R-L background clicks from L-R target clicks as a function of the number of background presentations before the target is presented, at three ISIs. Open squares represent the control condition, where the background was a single click from the left and the targets were single clicks from the right. Error bars represent between-subject standard errors.

sentations. A two-way (number of backgrounds \times ISI) repeated measures ANOVA showed that there was a significant effect of number of background repetitions [$F(11,94) = 16.06$, $p < 0.001$], ISI [$F(3,94) = 22.12$, $p < 0.001$], and a significant interaction between the two variables [$F(33,94) = 2.00$, $p < 0.001$].

Posthoc Bonferroni *t*-tests showed that within the 1-ms ISI condition, the 1–3 background presentations were significantly different from the 8–12 background presentations, and the 1 background presentation was also significantly different from the 5–7 number of backgrounds ($p < 0.05$). Within the 2-ms ISI condition, the 1–4 background presentations were significantly different from the 9–12 background presentations, and the 1 background presentation was also significantly different from the 6–8 background presentations ($p < 0.05$). Within the 5-ms ISI condition, the 1–4 background presentations were significantly different from the 8–12 background presentations, and the 1 background presentation was also significantly different from the 4–7 background presentations ($p < 0.05$). Within the control condition, however, there were no differences between any of the background presentation numbers ($p > 0.05$). Generally, within the paired click conditions, the higher the ISI, the more background presentations were necessary to increase the percent correct discriminations to the high rate of responding seen in the control condition.

Posthoc Bonferroni *t*-tests were also used to analyze whether the ISIs were significantly different from the control at each of the background presentation numbers. Within the 1–4 background presentations, the 5-ms ISI was significantly different from the control condition ($p < 0.05$). Within the 1–2 background presentations, the 2-ms ISI was signifi-

cantly different from the control condition ($p < 0.05$). Within only the 1 background presentation, the 1-ms ISI was significantly different from the control condition ($p < 0.05$). Overall, as ISI decreased, the functions were increasingly more similar to the control function. The buildup of localization dominance was unique to the paired click stimuli and was not seen in the simple control discrimination task.

VIII. DISCUSSION

A. Summing localization, localization dominance, and echo thresholds

Using a discrimination paradigm, these experiments show that budgerigars exhibit summing localization, localization dominance, and echo thresholds in a manner similar to humans and other animals tested. The delays resulting in excellent discrimination by the birds in all experiments correspond to the delays where localization dominance is operating in humans. During the repeating lead-lag background, the lag's spatial location was suppressed. When the lead-lag stimulus locations changed during the targets, the suppression of the lag's position was released (the breakdown of localization dominance), and the birds easily heard the change from the background condition to the target condition, and discrimination values were high.

At shorter ISIs, discrimination performance was worse for the budgerigars. In humans, localization dominance is not yet complete at these short ISIs. The results from budgerigars lend support to a similar perceptual illusion as the phantom image location that humans report. If the birds perceived a "phantom" image somewhere in-between the positions of the lead and lag location due to summing localization, discrimination between the phantom image background and the phantom image target would have been very difficult due to the poor sound localization abilities of these birds. Park and Dooling (1991) found that single-source broadband noises needed to be separated by almost 30° for discrimination in budgerigars. Discrimination suppression experiments measuring minimum audible angles in humans have found that localization of paired sound sources is higher for leads and lags than for single sound sources [see review in Litovsky *et al.* (1999)]. If this is also true for budgerigars, it is not surprising that discrimination between the paired sound sources at short ISIs was poor.

At longer delays, where echo thresholds are found in humans, discrimination was also more difficult for the budgerigars. In humans, both sounds are heard at separate locations, but they often cannot distinguish which came first. If fusion also breaks down in budgerigars past 8 ms, this may be the reason that discrimination of these stimulus pairs became so difficult. If the repeating backgrounds were heard as two separate stimuli presented from two separate locations, and the targets were heard as two separate stimuli presented from two separate locations, the temporal resolution (pattern perception) of these birds is not good enough to resolve which came first in the pair during the background to notice that the sequence had changed during the targets. This could account for the drop in discrimination rates.

B. Effects of stimulus duration

The duration of the stimulus has an effect on discrimination in budgerigars during the ISIs where summing localization is operating in humans (i.e., discrimination was worse for the shorter stimulus types). In most studies of summing localization where the stimuli overlap significantly (i.e., the 50-ms stimuli), the perceived image is more than the “phantom” average of the positions of the lead and lag stimuli (Litovsky *et al.*, 1999; Tollin and Henning, 1998, 1999). Here, the combined amplitudes and phases of the sounds interact to create the perceived location of the sound source. This may explain why the discrimination of the shorter stimuli at the short ISIs was worse than for the longer stimuli. However, since the shorter stimuli had a slightly narrower bandwidths than the longer stimuli, spectral differences might also contribute to the differential discrimination of shorter and longer stimuli. Although the birds were not identifying the actual perceived location of the stimuli in these experiments, human listeners presented with pairs of broadband noises actually perceive images towards the location of the lag at very short ISIs (Tollin and Henning, 1998, 1999). This is due to the interaural spectral cues that result from delaying and adding identical stimuli to the two ears. This effect was not seen at longer ISIs. It was also not seen with shorter stimuli (such as the 1-ms clicks). Again, this may be due to the fact that shorter stimuli contained less energy at higher frequencies than did the longer stimuli. The results from these experiments suggest that summing localization inhibited discrimination between stimulus pairs at the shortest ISIs in these birds as it did in humans.

The peak of responding, or where localization dominance might be at a maximum, occurred at similar ISIs across all stimulus types: at 1 ms for the 0.1- and 1-ms clicks and at 4 ms for the 50-ms noise bursts. Generally, performance was above 80% from 0.5 to 5 ms for all three stimulus types. Stimulus duration does not seem to have a large effect on localization dominance in budgerigars.

Echo thresholds, or the lowest percent correct discriminations, did not increase from the 0.1- to 1-ms stimuli, but did increase with increasing stimulus durations from 1 to 50 ms. Echo thresholds were at 20 ms for the 0.1-ms stimuli, at 8 ms for the 1-ms stimuli, and at 12 ms for the 50-ms stimuli. In studies with humans, barn owls, and cats [see review in Litovsky *et al.* (1999); also Keller and Takahashi (1996) and Litovsky and Yin (1998a, b)], later echo thresholds are generally seen with longer stimuli. The average function for budgerigars did not follow this trend when increasing stimulus length from 0.1 to 1 ms. However, differences were found when examining individual results. In the discrimination functions for the 0.1-ms stimuli, one of the birds had an echo threshold at 2 ms while the other bird's echo threshold was at 20 ms. Discrepancies between the birds are similar to those found in humans, however, where differences in echo thresholds can range from 1 to 50 ms across subjects (Litovsky and Shinn-Cunningham, 2001).

The echo thresholds for the 50-ms stimuli are also not as pronounced as in the shorter stimulus conditions. The birds slightly decreased their responses at the very long ISIs (~10 ms), the delay of the presumed echo threshold. It is possible

that the echo thresholds coincided with the time interval where the birds were able to determine which came first in the pair of noise bursts (a pattern perception task) and were thus able to discriminate the background from the target. This could account for the decrease in discrimination compared to the other two stimulus conditions.

The echo thresholds in the contact call discrimination functions were not apparent at all. In humans, echo thresholds are later in short stimuli with abrupt onsets (e.g., Rakerd and Hartmann, 1986). These contact calls were more speech-like in their appearance than the broadband stimuli, lasting between 120 and 160 ms and with more gradual onsets than the broadband stimuli. This, along with the increased pattern perception abilities described above, may have masked the echo thresholds for these conditions.

Overall, the findings of increasing echo thresholds with increasing stimulus duration agree with the results from humans. Unfortunately, only one systematic study of the effect of stimulus duration (where the same procedures and setup were used) has been psychophysically conducted on humans. Schubert and Wernick (1969) found that the point where the lead and lag became “equally loud” increased significantly as stimulus duration increased from 20 to 100 ms. Physiological studies of the effect of stimulus duration in cats also showed increased timecourses of neural echo suppression in the central nucleus of the inferior colliculus with increased stimulus duration (Litovsky and Yin, 1998a, b). A review of the literature by Litovsky *et al.* (1999) shows that, across studies, echo thresholds are generally higher for speech than they are for noise bursts and higher for noise bursts than for clicks. Those findings are generally supported here in budgerigars as well.

C. Effects of stimulus location and intensity

No differences were found between the discrimination functions for budgerigars where the leading stimulus was located on the left and those where the leading stimulus were located on the right. In humans, there is an asymmetry in the buildup of localization dominance depending on which side the leading sound is emitted from (Clifton and Freyman, 1989; Grantham, 1996). One of the arguments for the left-right asymmetry of the precedence effect in humans is that overall hearing abilities are asymmetrical in humans (e.g., Ward, 1957; Emmerich *et al.*, 1988) and that suppression of echoes is better (due to different amounts of buildup) when the lead is presented to the better ear and the lag to the worse ear (Grantham, 1996). As far as we know, there is no ear advantage in hearing abilities in budgerigars, or any other bird, and the existence of such things as the interaural pathway would seem to work against such a phenomenon. Further research is needed to examine these interesting differences between budgerigars and humans, such as testing other animals with hearing asymmetries on the precedence effect to see if the correlation holds true for more than these two species.

Discrimination functions of budgerigars were also similar to each other when the lead and lag stimuli were presented from the front and back as opposed to the left and right. Their results are in accord with results from humans.

Several studies have found similar amounts of localization dominance on the front-rear and left-right planes in humans (Blauert, 1971; Litovsky *et al.*, 1997; Rakerd and Hartmann, 1994; Rakerd *et al.*, 2000). In those studies, localization dominance was found in humans with only monaural spectral cues and no binaural cues. While the birds' heads were not held in a fixed position in this experiment, the binaural difference cues the birds would have received were below the threshold for azimuthal cues found in earlier experiments (Park and Dooling, 1991). Yet, although overall discrimination performance was slightly lower, the precedence effect was still found because the shapes of the functions remained similar.

Head-related transfer functions have not yet been measured in budgerigars, and they have no external pinnae, but it is still possible that spectral changes in the stimuli caused by transformations of the sounds as they traveled through the external ear canal were sufficient to allow for discrimination between these stimulus pairs. In fact, in a study of azimuthal sound localization in budgerigars, when one ear was plugged (dramatically reducing most binaural cues) localization accuracy for broadband noise bursts only increased from 27° to 30° (Park and Dooling, 1991). Monaurally deafened birds, however, were unable to localize sounds separated in the azimuthal plane by as much as 180°. This suggests that even small binaural cues, possibly enhanced by tympanic coupling via the interaural canal, might have been sufficient for localization dominance on the median sagittal plane. Without holding the bird's head fixed, it is still unknown as to whether budgerigars can use spectral cues alone for the precedence effect, as in humans.

In humans, the suppression of the lag is greater when the lead and lag come from similar locations in space (Litovsky and Shinn-Cunningham, 2001). The data reported here for budgerigars do not support that idea; discrimination performance was actually highest in the condition with the largest separation of lead-lag stimulus pairs. At the timecourses where localization dominance was operating, discrimination performance was higher when the speakers were further apart than when they were closer together. These results are different from those in humans (Litovsky and Shinn-Cunningham, 2001). Here, the poor sound localization abilities of budgerigars may have affected discrimination performance for the stimuli that were closer together in space. Taken together, these experiments demonstrate that two speakers separated by a large spatial distance are adequate to produce the precedence effect whether they are located at the spatial position corresponding to the maximal ITDs and ILDs or not, but that interaural difference cues may, in fact, be related to some aspects of the precedence effect in budgerigars.

The results from the intensity experiment showed that more intense stimuli were more difficult to discriminate than the less intense stimuli at low and intermediate ISIs, and that echo thresholds were earlier for more intense stimuli. These results, as a whole, agree with those from humans that echo thresholds are later for less intense stimuli (Babkoff and Sutton, 1966; Schubert and Wernick, 1969). The effects of stimulus level on the timecourses of localization dominance

and echo thresholds have not been well examined in behavioral studies with animals, but the results here support the findings in humans.

D. Buildup of the precedence effect

Experiment 5 showed that budgerigars exhibit the buildup and breakdown of localization dominance like humans. The precedence effect stimuli, but not the single source stimuli, required repeated background presentations to increase discrimination performance. Moreover, that discrimination performance increased at a slower rate for the 1- and 2-ms ISIs than for the 5-ms ISI. These results were consistent with the human results—more background repetitions were necessary for the precedence effect to build up at longer ISIs (Clifton and Freyman, 1989).

There were some interesting differences between the three subjects in this experiment. Penny and Will both showed higher responses as ISI decreased from 5 to 2 ms and from 2 to 1 ms. Cirrus, on the other hand, had an increase in responses from the 5- to the 2-ms conditions, but his lowest rate of responding was for the 1 ms ISI condition. In humans, one of the requirements for this buildup of echo suppression is that it only works at ISIs where localization dominance is strong. Figure 2 shows that Cirrus (in the original precedence effect experiment) had lower discrimination rates than the other two birds until the ISI reached 1 ms, suggesting that his localization dominance did not build up until later ISIs compared to the other birds. Localization dominance for the other two birds was seen at shorter ISIs, but not for Cirrus. It could be that localization dominance was not yet complete for Cirrus at the 1-ms ISI (he was still in summing localization), so his buildup and breakdown were not as strong as in the other two birds or in the other two ISIs.

Clifton *et al.* (1994) postulated that ongoing echoes provide the listener with some information about the room's acoustics and that making unlikely changes in the lead and/or the echo result in a release from suppression. The suggestion that the buildup process comes from expectations that human listeners have about room acoustics was strengthened in an experiment where the intensity and frequency spectra of the stimuli were changed from presentation to presentation, and the buildup still occurred (Clifton *et al.*, 1994). Changes in frequency and intensity are not “unlikely” events that would naturally occur in a room, so the buildup of echo suppression continued. The underlying neural correlates for buildup have yet to be discovered in any animal (Litovsky and Yin, 1998a), and buildup has previously been found behaviorally only once in nonhuman mammals (Kalmykova, 1993). The results here are the first that show that the buildup of echo suppression occurs in a bird species.

E. Conclusions

The present results extend what is known about hearing in budgerigars to a phenomenon known as the precedence effect. Results from these experiments move beyond studies of basic hearing and the limits of complex auditory temporal processing in budgerigars to a more complicated phenomenon—the precedence effect (Dent *et al.*, 1997;

Dooling *et al.*, 2000; Park and Dooling, 1991). Previous experiments examining simple sound localization abilities in simple acoustic environments are important for determining the basic capabilities of animals, but they tell us nothing about how an animal perceives the auditory world outside of the experimental situation. The present detailed examination of an auditory illusory phenomenon in an animal species provides further information about the complex auditory world of these small birds. It also suggests that studies that treat animals as simple auditory receivers may be missing some important complicated variables.

The purpose of these experiments was to define the timecourse of summing localization, localization dominance, and echo thresholds in budgerigars and to compare the results with those found in humans and other animals. Despite differences in experimental procedures and setups, as well as differences between species in their anatomical pathways for audition, ecological pressures of predation, and physical cues available for directional hearing, the comparative results are remarkably similar across all animals tested to date. These results support the notion that the precedence effect operates in budgerigars and thus provide evidence that the precedence effect is a general auditory mechanism that improves hearing in complex environmental conditions. As a whole, these results provide robust baseline data establishing this phenomenon in birds and also provide a foundation for future investigations of the anatomical and physiological substrates that can support this behavior.

ACKNOWLEDGMENTS

This work was supported by NIH Grants No. DC-00198 to RJD and NRSAs from NIH (DC-00046 and MH-12698) to MLD. We are very grateful to B. Brittan-Powell, F. Kubke, M. Leek, D. Yager, C. Carr, C. Moss, R. Litovsky, T. Takahashi, and D. Tollin for their comments on earlier versions of this manuscript.

- Babkoff, H., and Sutton, S. (1966). "End point of lateralization for dichotic clicks," *J. Acoust. Soc. Am.* **39**, 87–102.
- Blauert, J. (1971). "Localization and the law of the first wavefront in the median plane," *J. Acoust. Soc. Am.* **50**, 466–470.
- Blauert, J. (1997). *Spatial Hearing* (MIT, Cambridge, MA).
- Carr, C. E. (1992). "Evolution of the central auditory system in reptiles and birds," in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 511–543.
- Clifton, R. K. (1987). "Breakdown of echo suppression in the precedence effect," *J. Acoust. Soc. Am.* **82**, 1834–1835.
- Clifton, R. K., and Freyman, R. L. (1989). "Effect of click rate and delay on breakdown of the precedence effect," *Percept. Psychophys.* **46**, 139–145.
- Clifton, R. K., Freyman, R. L., Litovsky, R. Y., and McCall, D. (1994). "Listeners' expectations about echoes can raise or lower echo thresholds," *J. Acoust. Soc. Am.* **95**, 1525–1533.
- Cranford, J. L. (1982). "Localization of paired sound sources in cats: Effects of variable arrival times," *J. Acoust. Soc. Am.* **72**, 1834–1835.
- Dent, M. L., Larsen, O. N., and Dooling, R. J. (1997). "Free-field binaural unmasking in budgerigars (*Melopsittacus undulatus*)," *Behav. Neurosci.* **111**, 590–598.
- Dooling, R. J., Lohr, B., and Dent, M. L. (2000). "Hearing in birds and reptiles," in *Comparative Hearing: Birds and Reptiles*, edited by R. J. Dooling, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 308–359.
- Emmerich, D. S., Harris, J., Brown, W. S., and Springer, S. P. (1988). "The relationship between auditory sensitivity and ear asymmetry on a dichotic listening task," *Neuropsychologia* **26**, 133–143.
- Farabaugh, S. M., and Dooling, R. J. (1996). "Ecology and evolution of acoustic communication in parrots: Laboratory and field studies in parrots," in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Cornell U.P., Ithaca), pp. 97–117.
- Freyman, R. L., Clifton, R. K., and Litovsky, R. Y. (1991). "Dynamic processes in the precedence effect," *J. Acoust. Soc. Am.* **90**, 874–884.
- Goverts, S. T., Houtgast, T., and van Beek, H. H. (2000). "The precedence effect for lateralization at low sensation level," *Hear. Res.* **148**, 88–94.
- Grantham, D. W. (1996). "Left-right asymmetry in the buildup of echo suppression in normal-hearing adults," *J. Acoust. Soc. Am.* **99**, 1118–1123.
- Hebrank, J., and Wright, D. (1974). "Spectral cues in the localization of sound sources on the median plane," *J. Acoust. Soc. Am.* **56**, 1829–1834.
- Hoeffding, V., and Harrison, J. M. (1979). "Auditory discrimination: Role of time and intensity in the precedence effect," *J. Exp. Anal. Behav.* **32**, 157–166.
- Houtgast, T., and Aoki, S. (1994). "Stimulus-onset dominance in the perception of binaural information," *Hear. Res.* **72**, 29–36.
- Kalmykova, I. V. (1993). "Investigation of the precedence effect in the cat auditory system," *Sens. Syst.* **7**, 208–211.
- Keller, C. H., and Takahashi, T. T. (1996). "Responses to simulated echoes by neurons in the barn owl's auditory space map," *J. Comp. Physiol., A* **178**, 499–512.
- Kelly, J. B. (1974). "Localization of paired sound sources in the rat: Small time differences," *J. Acoust. Soc. Am.* **55**, 1277–1284.
- Klump, G. M. (2000). "Sound localization in birds," in *Comparative Hearing: Birds and Reptiles*, edited by R. J. Dooling, R. R. Fay, and A. N. Popper (Springer, Berlin), pp. 249–307.
- Larsen, O. N., Dooling, R. J., and Ryals, B. M. (1997). "Roles of intracranial air pressure on hearing in birds," in *Diversity in Auditory Mechanics* (Bechtel Engineering Center, University of California, Berkeley), pp. 253–259.
- Litovsky, R. Y., and Shinn-Cunningham, B. G. (2001). "Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression," *J. Acoust. Soc. Am.* **109**, 346–358.
- Litovsky, R. Y., and Yin, T. C. T. (1998a). "Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics," *J. Neurophysiol.* **80**, 1285–1301.
- Litovsky, R. Y., and Yin, T. C. T. (1998b). "Physiological studies of the precedence effect in the inferior colliculus of the cat. II. Neural mechanisms," *J. Neurophysiol.* **80**, 1302–1316.
- Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (1999). "The precedence effect," *J. Acoust. Soc. Am.* **106**, 1633–1654.
- Litovsky, R. Y., Rakerd, B., Yin, T. C. T., and Hartmann, W. M. (1997). "Psychophysical and physiological evidence for a precedence effect in the median sagittal plane," *J. Neurophysiol.* **77**, 2223–2226.
- Park, T. J., and Dooling, R. J. (1991). "Sound localization in small birds: Absolute localization in azimuth," *J. Comp. Physiol.* **105**, 125–133.
- Rakerd, B., and Hartmann, W. M. (1986). "Localization of sound in rooms, III: Onset and duration effects," *J. Acoust. Soc. Am.* **80**, 1695–1706.
- Rakerd, B., and Hartmann, W. M. (1994). "More on the precedence effect in three planes," *J. Acoust. Soc. Am.* **95**, 2917.
- Rakerd, B., Hartmann, W. M., and Hsu, J. (2000). "Echo suppression in the horizontal and median sagittal planes," *J. Acoust. Soc. Am.* **107**, 1061–1064.
- Schubert, E. D., and Wernick, J. (1969). "Envelope versus microstructure in the fusion of dichotic signals," *J. Acoust. Soc. Am.* **45**, 1525–1531.
- Searle, C. L., Braid, L. D., Davis, M. F., and Colburn, H. S. (1976). "Model for auditory localization," *J. Acoust. Soc. Am.* **60**, 1164–1175.
- Shinn-Cunningham, B. G., Zurek, P. M., and Durlach, N. I. (1993). "Adjustment and discrimination measurements of the precedence effect," *J. Acoust. Soc. Am.* **93**, 2923–2932.
- Takahashi, T. T., and Keller, C. H. (1994). "Representation of multiple sound sources in the owl's auditory space map," *J. Neurosci.* **14**, 4780–4793.
- Tollin, D. J., and Henning, G. B. (1998). "Some aspects of the lateralization of echoed sound in man. I. The classical interaural-delay based precedence effect," *J. Acoust. Soc. Am.* **104**, 3030–3038.
- Tollin, D. J., and Henning, G. B. (1999). "Some aspects of the lateralization of echoed sound in man. II. The role of the stimulus spectrum," *J. Acoust. Soc. Am.* **105**, 838–849.
- Tollin, D. J., Populin, L. C., and Yin, T. C. T. (2000). "Neural correlates of

- the precedence effect and echo threshold in the inferior colliculus of the behaving cat," paper presented at the 30th Annual Meeting of the Society for Neuroscience, New Orleans, LA.
- Wallach, H., Newman, E. B., and Rosenzweig, M. R. (1949). "The precedence effect in sound localization," *Am. J. Psychol.* **62**, 315–336.
- Ward, W. D. (1957). "Hearing of naval aircraft maintenance personnel," *J. Acoust. Soc. Am.* **29**, 1289–1301.
- Yang, X., and Grantham, D. W. (1997). "Cross-spectral and temporal factors in the precedence effect: Discrimination suppression of the lag sound in free-field," *J. Acoust. Soc. Am.* **102**, 2973–2983.
- Yin, T. C. T. (1994). "Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat," *J. Neurosci.* **14**, 5170–5186.
- Zurek, P. M. (1987). "The precedence effect," in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York), pp. 85–105.

Investigations of the precedence effect in budgerigars: The perceived location of auditory images

Micheal L. Dent^{a)} and Robert J. Dooling^{b)}

Department of Psychology, University of Maryland, College Park, Maryland 20742

(Received 15 May 2002; revised 10 January 2003; accepted 24 January 2003)

The perceived location of auditory images has been recently studied in budgerigars [Dent and Dooling, *J. Acoust. Soc. Am.* **113**, 2146–2158 (2003)]. Those results suggested that budgerigars (*Melopsittacus undulatus*) perceive precedence effect stimuli in a manner similar to humans and other animals. Here we extend those experiments to include the effects of intensity on the perceived location of auditory images and the perceived location of paired stimuli from multiple locations in space. We measured the abilities of budgerigars to discriminate between paired stimuli separated in time, intensity, and/or location. Increasing the intensity of a lag stimulus disrupted localization dominance. Budgerigars also perceived simultaneously presented (away from the midline) stimuli as very similar to a single sound presented from the midline, much like the phantom image reported in humans. The perception of paired stimuli from one side of the head versus two sides of the head was also examined and showed that the spatial cues available in these stimuli are important and that echoes are not perceptually inaccessible during localization dominance conditions. The results from these experiments add further data showing the precedence effect in budgerigars is similar to that found in humans and other animals. © 2003 Acoustical Society of America.

[DOI: 10.1121/1.1560161]

PACS numbers: 43.80.Lb, 43.66.Gf, 43.66.Qp [WA]

I. INTRODUCTION

A variety of techniques have been used to measure aspects of the precedence effect in animals. For instance, behavioral studies tracking eye movements of cats previously trained to look at the position of a single sound source show that two sounds played simultaneously caused the cats to look directly ahead, at a “phantom” image location, much like summing localization in humans (Populin and Yin, 1998; Tollin *et al.*, 2000). At increasing delays, the cats shifted their eyes towards the location of the lead sound, suggesting that the cats perceive a single sound at the location of the lead during that timecourse, again in a manner similar to humans. At even longer delays, cats often looked towards the location of the lead and then the position of the lag, or vice versa. This suggests that at those timecourses, they perceive both the lead and lag at their respective locations, similar to what happens once past the echo threshold in humans (Tollin *et al.*, 2000).

In another study, Cranford (1982) trained cats to release either a right or left foot pedal, whichever corresponded to the side of the leading sound source location. At intermediate interstimulus delays, the position of leading source was identified on close to 100% of the trials. At shorter and longer delays, performance was at chance levels. Here, the cats were unable to reliably locate the position of the leading sound source. Taken together, the results from these independent laboratories using different methodologies suggest that cats do indeed perceive summing localization and localiza-

tion dominance stimuli in a manner similar to humans.

The precedence effect has also recently been found in budgerigars (*Melopsittacus undulatus*). Rather than the identification methods used above, discrimination experiments utilizing the buildup and breakdown of localization dominance were used as an assay of the precedence effect. The results obtained here are consistent with the three phases of the precedence effect (summing localization, localization dominance, and echo thresholds) and the timecourses were similar to those found in humans and other animals (Dent and Dooling, 2003). To review these data, budgerigars had difficulty discriminating the L-R (left-right) background pairs from R-L (right-left) target pairs with the same interstimulus interval at the timecourses where summing localization is operating in humans (below 1 ms). At these timecourses, the perceived location of the auditory image is in-between the actual locations of the sound sources. Budgerigars had little difficulty discriminating the L-R background from the R-L target at the timecourses where localization dominance operates in humans (1–5 ms). Here, the perceived location of the auditory image is at the location of the lead stimulus only. These results suggest that in budgerigars, as in humans, the spatial attributes of the lag are perceptually inaccessible. At longer timecourses, where the echo thresholds are found in humans (5–10 ms), budgerigars again had difficulty discriminating background stimuli from targets.

The present experiments test the validity of these earlier results by measuring whether an intensity increase in a lag stimulus can decrease performance and whether simultaneously presented sounds are perceived from a phantom location in space. In humans, it is known that intensity differences can elicit localization dominance and summing

^{a)}Current address: Department of Physiology, University of Wisconsin–Madison, Madison, WI 53706.

^{b)}Author to whom correspondence should be addressed. Electronic mail: dooling@psyc.umd.edu

localization, and that there is a time-intensity trading effect when the intensity of the lead stimulus is decreased (Haas, 1951; Leahey and Cherry, 1957). Less is known about the effects of intensity on the precedence effect in animals, however. Hoeffding and Harrison (1979) and Kelly (1974) found thresholds of 3–5 dB for discriminating pairs of stimuli separated by intensity differences only, although the latter authors could not disrupt performance by decreasing the intensity of the lead when time differences were introduced, suggesting that the time-intensity trading ratio is not found in all animals. Here, we measure the intensity differences required for discrimination between simultaneously presented stimuli and test whether a time-intensity trading can disrupt performance in these birds.

In a second experiment, the nature of summing localization in budgerigars is investigated. Cats look towards the midline when presented with simultaneously presented sounds (Populin and Yin, 1998), behaving as if they perceive a phantom image in a similar manner as humans. It is not known, however, whether the quality of that phantom image perceived by cats is indistinguishable from a single sound presented from that central location. Using a discrimination paradigm, we measured the ability of birds to discriminate between a centrally located single stimulus and laterally located paired stimuli presented simultaneously.

In a third series of experiments, we measure the discriminability of “echoes.” During normal localization dominance conditions, animals behave appropriately towards a primary sound source while ignoring echoes that may be only a few decibels below and a few ms after the primary sound. It is important to note that although the echoes are ignored, they do affect perception—listeners can easily discriminate between sounds with and without echoes. Perrott *et al.* (1987, 1989) found that spatial resolution for stimuli under localization dominance conditions was only mildly affected. Litovsky and MacMillan (1994) found that minimum audible angles for lead stimuli were better than those for lag stimuli, but they were still measurable in the lag. The delay between presentation of the lead and lag sounds also has an effect on the extent that the lag affects the perceived location of the sound image. During summing localization, both the lead and lag contribute to the perceived location, while during localization dominance, the lag has little to no effect on the perceived location of the auditory image.

Some researchers have stressed the positive perceptual influences of having echoes present and only their directional cues rendered inaccessible. Echoes can have an effect on speech communication and can enhance the overall quality of complex sounds (Haas, 1972; Freyman *et al.*, 1998; Mershon *et al.*, 1989). In humans, perceived distance judgments of a sound source in a room became much more accurate as the number of reflections is increased up to a certain point (Bronkhorst and Houtgast, 1999). Recent field work suggests that birds may be able to use reverberations of calls as cues for auditory distance perception (Naguib, 1995; Nelson and Stoddard, 1998), and of course bats and some species of birds are able to use returning echolocation signals for maneuvering in their environment (Griffin, 1958; Griffin and Suthers, 1970; Konishi and Knudsen, 1979). So while the

precedence effect allows humans and other animals to disregard echoes as true separate sound sources, the preservation of the information in those echoes may also be providing listeners with important cues about their auditory environment. For this reason, a detailed examination of the potential uses and perceptual accessibility of these cues is important for understanding hearing mechanisms as well as acoustic communication strategies. In three experiments, we measure the discrimination of paired stimuli presented unilaterally or bilaterally to determine the nature of echo perception and discrimination in these small birds. As a whole, these experiments will lend further support to the hypothesis that the precedence effect, its component phenomena, and the general perception of echoes are similar across all animals.

II. GENERAL METHODS

A. Subjects

Four adult budgerigars (three males, one female) were used in these experiments. The number of subjects used within each individual experiment is noted within that experiment. The birds were either bred from commercial stock in a vivarium at the University of Maryland or purchased from a local breeder. The birds were housed in individual cages and kept on a normal day/night cycle correlated with the season at approximately 90% of their free-feeding weights. All animal experimentation was conducted under the auspices of an approved protocol from the Animal Care and Use Committee at the University of Maryland, College Park.

B. Testing apparatus and stimuli

The psychoacoustic experiments took place in a wire test cage ($25 \times 18 \times 14 \text{ cm}^3$) mounted 115 cm from the ground in a sound-attenuating chamber ($2.8 \times 2.5 \times 2.0 \text{ m}^3$). The test cage consisted of a perch, an automatic feeder on the floor of the cage (food hopper), and two vertical response keys extending from the floor in front of the bird. The response keys were two sensitive microswitches with 8-mm light emitting diodes (LEDs) located just above the food hopper. The bird pecking the LED tripped the microswitch. The left microswitch and LED served as the observation key, and the right microswitch and LED served as the report key. During a session, a small light at the top of the test cage illuminated the chamber. The behavior of the animals during test sessions was monitored at all times by an overhead video camera system.

The experiment was controlled by an IBM Pentium III microcomputer operating Tucker-Davis Technologies (TDT, Gainesville, FL) modules. The stimuli were 1-ms broadband clicks presented at 70 dB SPL (unless otherwise noted). Stimuli were generated in advance of testing, stored in digital form, and output at a sampling rate of 50 kHz via a timing generator (TDT, Model TG6) to a four-channel D/A converter (TDT, Model DA3-4). Each signal was then output from a separate channel of the D/A converter to a separate digital attenuator (TDT, Model PA4) and a separate amplifier (TDT, Model HB6) to separate speakers located in the testing chamber (Realistic 3" midrange tweeter). The speakers were

placed at the same height as the cage, 60 cm from the position of the bird's head during testing. Stimulus calibration was performed with a Larson-Davis sound level meter (Model 825, Provo, UT) with a 20-ft extension cable and $\frac{1}{2}$ -in. microphone and measured at the position normally occupied by the bird's head during testing.

C. Training and testing procedures

The training and testing procedures have been described in detail in Dent and Dooling (2003). The birds were trained by a standard operant autoshaping program to peck at the microswitch keys for food reinforcement. First, they pecked repeatedly at the left key (observation key) during repeating presentations (rate of 2/s) of a stimulus emitted from single source (background; e.g., a left speaker only). After a random interval of 1–6 s, the same stimulus was played from a different source (target; e.g., from the right speaker). The bird was trained to peck the right key (report key) when it detected this change in order to obtain food. The birds were rewarded on 70% of the trials with a 2-s free access to seed if they detected the change. Percent correct hit values on trials involving a change from background to target were recorded for further analysis.

Incorrect report key pecks were punished with a timeout during which all of the room lights were extinguished. A miss was recorded if the bird failed to peck the report key within 2 s of the change in stimulus presentation. Sham trials, where there was no change of sound presentation type from background to target, consisted of 30% of all trials. Pecks to the report key during these trials were counted as false alarms, and the birds were again punished with a lights-out period. Sessions with a false alarm rate of 15% or higher were discarded. Approximately 6% of all sessions were discarded for this reason. An experimental session consisted of 100 trials. The birds were tested at least two sessions a day, and between 5 and 7 days a week. Each bird was run on at least 200 trials at each experimental condition, and the last 100 trials after discrimination performance stabilized were analyzed.

III. EXPERIMENT 1: TIME-INTENSITY TRADING RATIO

The first experiment tested whether intensity differences would elicit localization dominance in budgerigars, and if there was a time-intensity trading ratio for localization dominance, in other words, if intensity differences between stimuli could abolish the buildup of localization dominance that was due to time differences between stimuli. In humans, increasing the intensity of the echo sound relative to the primary sound can disrupt localization dominance on stimuli separated by a short delay (Leakey and Cherry, 1957; Wallach *et al.*, 1949). Snow (1954) investigated the time differences between two stimuli that could be compensated for by level differences between the stimuli. Placing a primary speaker at -45° azimuth and an echo speaker (delayed for a few ms) at $+45^\circ$ azimuth (where normally the perceived sound would be at -45°), Snow (1954) showed that decreasing the intensity of the primary speaker by 5–8 dB actually

centered the perceived auditory image at 0° . The birds were tested on their discrimination of stimuli with a similar manipulation.

A. Methods

Once the birds were trained to peck the observation key repeatedly during presentation of a single stimulus emitted from the left speaker (-90° , repeating background) and to report a change in the location of this repeating background ($+90^\circ$ target) by pecking the report key, paired stimuli were substituted for the single stimuli. Here the repeating background involved presenting the stimuli from two speakers [for a detailed explanation, see Dent and Dooling (2003)]. The birds were required to discriminate a L-R (left first, then right) paired background from a R-L paired target. The inter-stimulus interval (ISI) between the right and left stimuli was the same from background to target; the lead-lag click delay was merely reversed in location. The lead-lag clicks constituting the background were played at a rate of 2/s a random number of times (between 6 and 12) before the targets were presented. The lead-lag clicks constituting the targets were presented four times (also at a rate of 2/s) or until the birds responded.

Three budgerigars were tested on discrimination of pairs of 1-ms broadband clicks at three different ISIs (0, 1, and 5 ms). Here, the task was to discriminate a repeating click pair background with the left speaker leading the right by some ISI from a repeating click pair target with the right speaker leading the left by the same ISI. Baseline (time differences between lead and lag only) discrimination performance for all L-R stimulus discriminations at different ISIs had been determined earlier [Fig. 2 of Dent and Dooling (2003)]. The subjects were then retested with the “echo” (second click in the pair of clicks) at 3, 5, 7, and 10 dB more intense than the lead sound at each of the three ISIs.

In the intensity difference *only* condition, the sounds were presented simultaneously from the left and right speakers. One of the clicks was less intense during the repeating background, and the other one was less intense during the target. In the intensity *and* time difference condition, the leading sound was less intense than the lagging sound during the repeating background. When the targets were presented, the ISIs and intensities of the lead and lag were switched. For instance, if the less-intense lead was on the left and the more-intense lag was presented from the right in the repeating background, when the targets were presented the less-intense lead was switched to the right speaker and the more-intense lag was switched to the left speaker. The leads and lags retained their timing and intensity separation from background to target; they merely changed locations. Overall discrimination performance was determined for each intensity and time difference combination in each separate session for each bird.

B. Results

Discrimination performance as a function of stimulus intensity difference for the three subjects is shown in Fig. 1 for three different ISIs. The mean false alarm rate for all experiments was 4% (mean d' value = 2.32). The interindi-

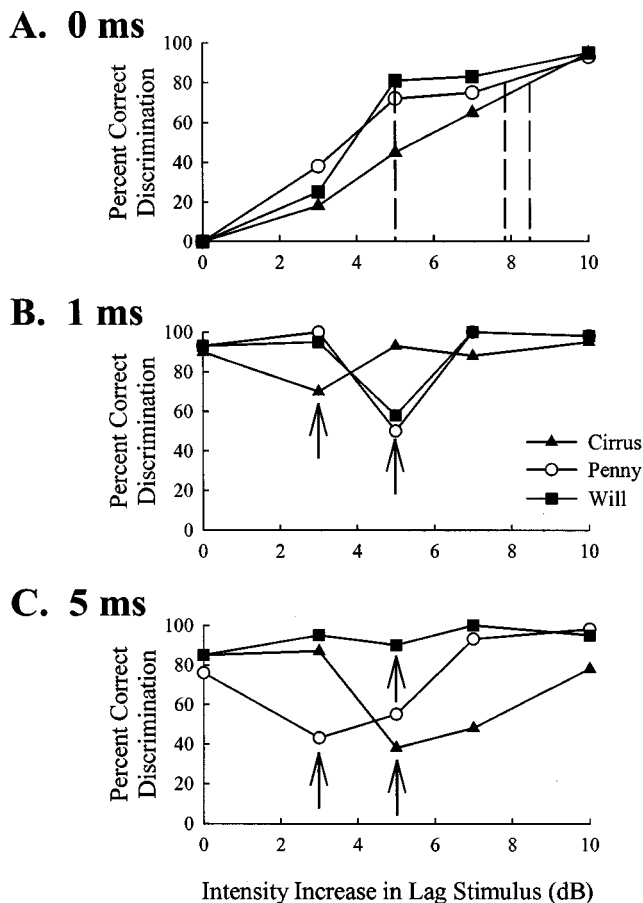


FIG. 1. Discrimination performance for three individuals (Cirrus=black triangles, Penny=white circles, Will=black squares) tested at three ISIs [(a)=0 ms, (b)=1 ms, (c)=5 ms], with the lag stimulus equal in intensity to (0 on the x axis) or more intense than the lead (>0 on the x axis).

vidual variation in the data shown in this experiment is fairly conservative compared to similar experiments in humans (see, for example, Litovsky and Shinn-Cunningham, 2001). With a 0-ms ISI [clicks presented simultaneously from the L and R speakers; Fig. 1(a)], it took between 5 and 8 dB to increase the discrimination performance to an arbitrary 80% criterion level (dotted lines). For the 1- [Fig. 1(b)] and 5-ms [Fig. 1(c)] ISIs, it took between 3 and 5 dB to decrease responding (arrows). Here, although the click on the left was presented first (either 1 or 5 ms earlier than the lag), the later click on the right was louder. This stimulus condition disrupted discrimination abilities for all three birds, although to varying degrees for each subject. At the 0-ms ISI, as intensity differences between the two stimuli increased, discrimination performance increased. At the 1-ms ISI, performance fell from close to 100% correct to under 60% correct for two birds (Penny and Will) and to 70% for one bird (Cirrus). At 5 ms, performance fell to almost 40% correct for two birds (Cirrus and Penny) but barely decreased for the third bird (Will). After this intensity increase in the lag caused a decrease in discrimination, performance returned to high levels for all three birds with further intensity increases in the lag. This experiment shows that intensity differences as well as time differences can have an effect on discrimination of paired sounds during the timecourse of localization dominance, although this effect is variable across subjects.

IV. EXPERIMENT 2: THE DISCRIMINATION OF MULTIPLE SOUND SOURCES

In humans, two simultaneously presented sounds from opposite sides of the listener result in an identification of that source at the midline (Wallach *et al.*, 1949). Cats, too, look towards the midline at simultaneously presented sounds (Populin and Yin, 1998). Although the budgerigars do not identify the location of the auditory image in our discrimination paradigm, we test the discrimination of two simultaneously presented sounds from the sides from one sound played from the midline. Difficulty in discriminating between the two stimulus conditions would suggest that budgerigars, like humans and cats, perceive a phantom midline image, and that the image has similar spectral qualities as well.

A. Methods

In this experiment, three budgerigars were tested on their discrimination of single or multiple source stimuli from different locations. The background 1-ms click stimuli were presented from one of four location conditions during each session: -60° , $+60^\circ$, 0° , or simultaneously from $+60^\circ$ and -60° . The targets in a session consisted of stimuli presented from the other three background conditions. All background-target combinations were tested. When the simultaneously presented $+60^\circ$ and -60° sounds were played, they were attenuated so that the signal received at the ears of the birds equaled that of a single sound source reaching the ears of the birds.

B. Results

When the task was a simple discrimination (e.g., left background versus right target), discrimination performance was close to 100% (Fig. 2). This was true when discriminations involved a large spatial separation (e.g., left versus right) or a smaller separation (e.g., left versus center). When the discriminations involved simultaneously presented backgrounds or targets from $+60^\circ$ and -60° , discrimination of a single sound source from the side was high (80%–90% correct), while discrimination of a single sound source from the center was very low (25%–40% correct). These birds could not easily discriminate a 0° background from a simultaneously presented -60° and $+60^\circ$ target, and vice versa, while a 0° background was easily discriminated from either the -60° or $+60^\circ$ targets when they were presented alone. Although these results were well above chance levels of performance, these discriminations were 65% lower than in the simpler discrimination tasks. These results suggest that budgerigars perceive simultaneously presented clicks at -60° and $+60^\circ$ as qualitatively similar to a single click presented alone at 0° .

V. EXPERIMENT 3: THE IMPORTANCE OF SPATIAL CUES

This experiment was a test of whether the spatial cues of the paired stimuli were salient to budgerigars, especially in light of the fact that they have generally poor sound localization abilities (Park and Dooling, 1991). In earlier prece-

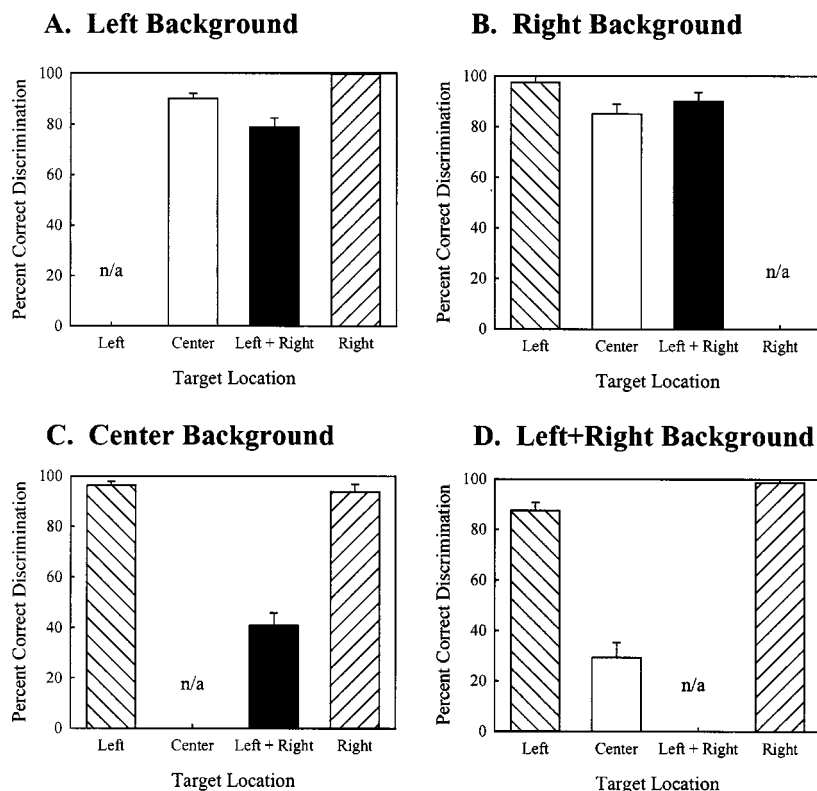


FIG. 2. Means from three individuals tested on discrimination of one of four background location conditions (A–D) and the corresponding three target location conditions.

dence effect experiments with budgerigars (Dent and Dooling, 2003), stimuli were presented at different times as well as from different locations. Here, a single click from one location was the background, and paired clicks from either one or two locations was the target. If the birds are able to use spatial cues in this task, discriminating between pairs of clicks from a single location should be more difficult than discriminating among pairs of clicks from multiple locations because of the added location cues.

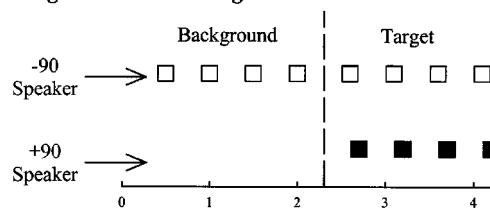
A. Methods

In this experiment, a single 1-ms click located at -90° was the background and the four subjects were required to discriminate targets where the lead was at -90° and the lag was at $+90^\circ$ [Fig. 3(a)] or where the targets had the same temporal properties of the previously mentioned lead-lag targets, but no spatial cues—because the targets were both at -90° and played from the same speaker [Fig. 3(b)]. The birds were tested on six lead-lag delays ranging from 0.1 to 20.0 ms. The single-side (-90°) targets ranged from clicks with a longer duration (0.1- to 1-ms ISI) to two clicks separated by a delay (5- to 20-ms ISI). This is not a precedence effect experiment because the repeating backgrounds were only single clicks presented from a single side. If the spatial cues are important, the task where the lag was on the opposite side of the lead [Fig. 3(b)] should be easier than the task where the lead and lag were on the same side [Fig. 3(a)] at each ISI. If *only* temporal factors play a role in discriminating these stimuli, performance should be equal for the two conditions and would increase with increasing ISI for both conditions.

As a control to ensure that the target speaker's location was the only cue, and not the temporal onset of the stimuli

(because at the shorter ISIs, the single speaker was simply emitting one long click where the onset of the second click was smeared into the first click), the condition where the target leads and lags were presented from one speaker on one side was compared for two of the subjects to the condition where two speakers were placed on top of each other at -90° (putting two speakers on top of each other ensures that the onsets of both clicks were present). In this configuration, the centers of the speakers were 11 cm apart. A second control experiment was conducted to measure potential monaural intensity cues that may have been created when presenting the two sounds from two speakers. Here, the background

A. Background = -90° / Targets = $+90^\circ$ and -90°



B. Background = -90° / Targets = -90°

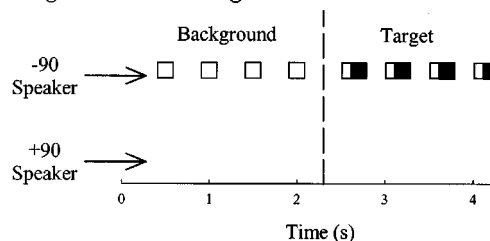


FIG. 3. Schematic of stimulus presentation. Background stimuli were presented from a -90° speaker. Targets were presented either from (a) -90° and $+90^\circ$ or from (b) -90° only.

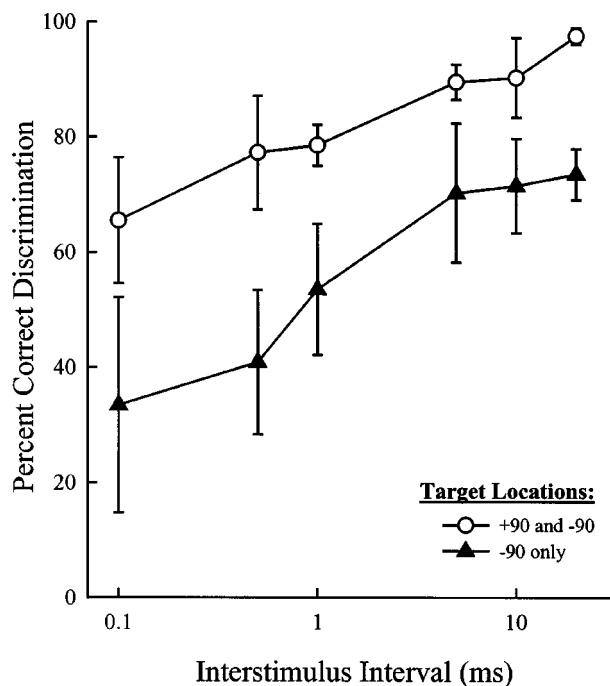


FIG. 4. Mean of four individuals tested on discrimination of a background from -90° versus targets from -90° and $+90^\circ$ (white circles) or from -90° only (black triangles). Error bars represent standard errors.

and target stimuli were roved by ± 4 dB from presentation to presentation and discrimination was compared for two subjects in the roved versus unroved conditions.

B. Results

For all four budgerigars, as ISI increased, discrimination performance increased for both target location conditions (Fig. 4). Additionally, the condition where the targets were emitted from two locations in space (white circles) yielded higher performance levels than the condition where the targets were only emitted from one location (black triangles). A two-way (ISI \times target speaker condition) repeated measures ANOVA showed that there was a significant effect of ISI [$F(5,33)=5.91$, $p<0.001$] and target speaker location condition [$F(1,33)=29.36$, $p<0.001$], but no significant interaction between the two variables [$F(5,33)=0.36$, $p>0.05$]. A *posthoc* Bonferroni *t*-test showed that the performance at 0.1-ms ISI was significantly lower than the 5–20-ms ISIs ($p<0.05$). The increase in ISI and the addition of a second speaker location (black versus white symbols) both increased discrimination performance.

The control experiment where the lead and lag on one side came from one speaker was not significantly different from the condition where the lead and lag on one side came from two speakers for either subject tested in this condition, as shown by paired *t*-tests [Cirrur: $t(5)=0.85$, $p>0.05$; Will: $t(5)=0.29$, $p>0.05$]. There were also no differences in the control roved versus unroved conditions, again shown by paired *t*-tests for each subject [Cirrur one-side target: $t(5)=0.39$, $p>0.05$; Cirrur two-sides target: $t(5)=1.57$, $p>0.05$; Will one-side target: $t(5)=1.74$, $p>0.05$; Will two-sides target: $t(5)=2.26$, $p>0.05$]. The re-

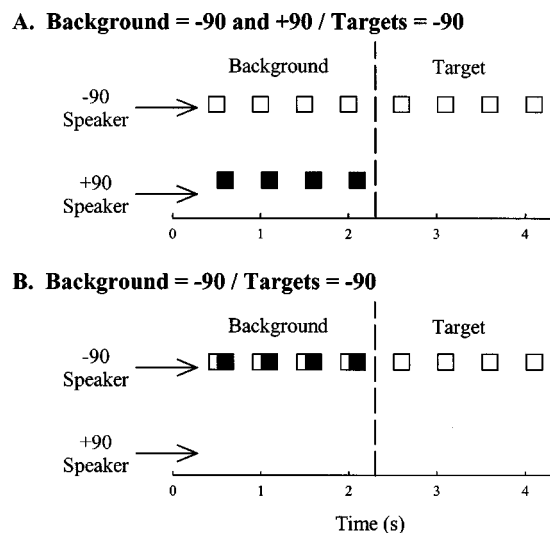


FIG. 5. Schematic of stimulus presentation. Target stimuli were always presented from -90° . Backgrounds were presented either from (a) -90° and $+90^\circ$ or from (b) -90° only.

sults from this experiment showed that the spatial properties of the stimuli in all of these experiments were salient cues for discrimination by the birds.

VI. EXPERIMENT 4: THE DISCRIMINATION OF ECHOES

In humans, under localization dominance conditions, the lag stimulus adds fullness to perceived images, even though its spatial properties are not perceived (e.g., Blauert, 1997). Humans can also distinguish between trials where a lag is presented and when a lag is not presented [see review in Litovsky *et al.* (1999)]. In this experiment, we tested whether the lag clicks are perceived, and whether the spatial properties of the lag clicks have an influence on discrimination performance. Paired stimuli from either one or two locations were the background conditions in this experiment, and birds were required to discriminate a single target located at the position of the background's lead click location. If the lag stimuli are perceptually inaccessible, discrimination should be very difficult. If only the spatial attributes of the stimuli are inaccessible, discrimination performance should not differ across lag location conditions. Further, performance should change across lead-lag interstimulus intervals in conjunction with the timecourses of summing localization, localization dominance, and echo thresholds.

A. Methods

In this experiment, the four budgerigars were required to discriminate between a repeating background pair of clicks from a single-click target located at -90° (opposite background-target conditions from experiment 3). The lead-lag background clicks were presented from opposite sides of the cage at $\pm 90^\circ$ [Fig. 5(a)] or on the same side (at -90° only) but with the same temporal properties as in the bilateral speaker condition [Fig. 5(b)]. The birds were tested on six lead-lag ISI backgrounds ranging from 0.1 to 20.0 ms.

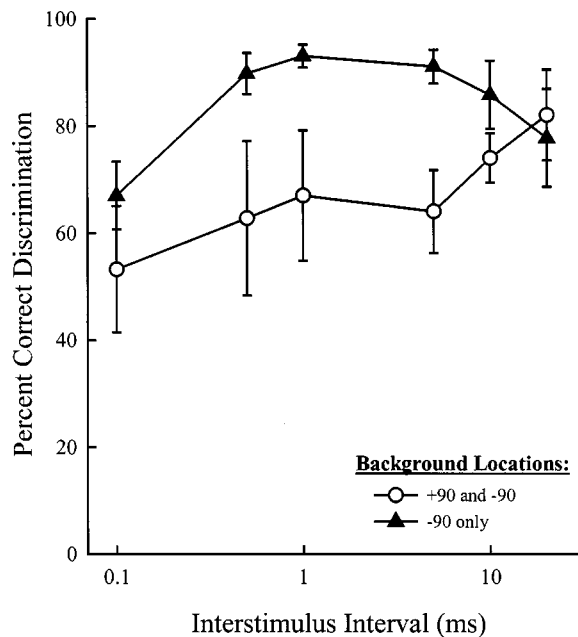


FIG. 6. Mean of four individuals tested on discrimination of a target from -90° versus backgrounds from -90° and $+90^\circ$ (white circles) or from -90° only (black triangles). Error bars represent standard errors.

As a control that the background speaker location was the only cue and not the temporal onset of the stimuli, two birds were tested on one additional speaker location condition. The condition in which the lead and lag were on the same side and presented from one speaker was compared to a second condition in which two speakers were placed on top of each other at -90° . As another control for potential intensity cues resulting from the addition of two sounds, two of the birds were retested with the stimuli roved by ± 4 dB from presentation to presentation.

B. Results

Unlike the results for experiment 3, in this experiment there was not a simple monotonic increase in discrimination performance with increasing ISI in either condition (Fig. 6). The condition where the background was emitted from two locations [shown in Fig. 5(a)] was more difficult than the condition where the background was emitted from only one location [shown in Fig. 5(b)], especially at the intermediate ISIs (0.5–10 ms). A two-way (ISI \times background speaker condition) repeated measures ANOVA showed that there were significant effects of ISI [$F(5,33)=2.79$, $p<0.05$] and target speaker condition [$F(1,33)=19.88$, $p<0.001$], but no interaction between the two variables [$F(5,33)=1.79$, $p>0.05$]. A posthoc Bonferroni t -test showed that within the 0.5-, 1-, and 5-ms ISIs, there were differences in discrimination between the two background speaker conditions ($p<0.05$), the same ISIs where localization dominance is operating in budgerigars (Dent and Dooling, 2003).

The control condition where the lead and lag from one side were emitted from one speaker was not significantly different from the condition where the lead and lag from one side were emitted from two speakers for either bird [Cirrus: $t(5)=1.52$, $p>0.05$; Spike: $t(5)=1.10$, $p>0.05$]. The

temporal onsets of these stimuli were not the important factor in discrimination between the backgrounds and targets. In the second control experiment, the condition where the intensities were roved was not significantly different from the condition where the intensities were held constant [Cirrus one-side background: $t(5)=0.97$, $p>0.05$; Cirrus two-sides background: $t(5)=2.08$, $p>0.05$; Spike one-side background: $t(5)=2.40$, $p>0.05$; Spike two-sides background: $t(5)=0.56$, $p>0.05$]. Making the intensity cues unpredictable did not change discrimination in this experiment either.

This experiment demonstrates that the discrimination of paired stimuli from two locations [shown in Fig. 5(a)] from a single target was more difficult than in a single source stimulus condition [paired one-location background shown in Fig. 5(b)]. Further, the differences between the two background conditions are largest during the timecourses of localization dominance and smaller during the timecourses of summing localization and echo thresholds [where the spatial attributes of the lag are known to have an influence on the perception of the auditory image(s)]. Overall, these results support previous findings in humans that the spatial properties of the lag stimulus are inaccessible during localization dominance conditions.

VII. EXPERIMENT 5: THE DISCRIMINATION OF PAIRED STIMULI

Experiment 4 demonstrated that during localization dominance timecourses, the spatial properties of the lag click are inaccessible as localization cues, not the lags themselves. Paired background stimuli (bilateral or unilateral) were always perceived as being different from a single target stimulus, however, because discrimination was greater than 0% at all ISIs. In this experiment, we attempted to decrease discrimination performance further by making the targets even more similar to the backgrounds. Here, the task was to discriminate pairs of background stimuli from pairs of target stimuli where only one of the pairs contained stimuli from two locations (e.g., we tested whether a pair of background clicks located at -90° and $+90^\circ$ were perceived as being the same as a pair of target clicks both located at -90° and vice versa).

A. Methods

A pair of clicks located at $+90^\circ$ and -90° separated by an ISI was the repeating background in the first task [Fig. 7(a)]. The four birds were required to discriminate targets with the same temporal properties of the previously mentioned lead-lag background but with both targets located at -90° . In another task, they were tested on the reversed background/target conditions [Fig. 7(b)]. In both tasks, the birds were tested on six lead-lag delays ranging from 0.1 to 20.0 ms.

Experiment 3 showed that under nonprecedence effect conditions (discriminating a single click from a pair of clicks), the spatial and temporal properties of the stimuli could both be used as discrimination cues. Experiment 4, however, showed that under precedence effect conditions and only during localization dominance timecourses, the spatial properties of the stimuli are inaccessible and the birds are not

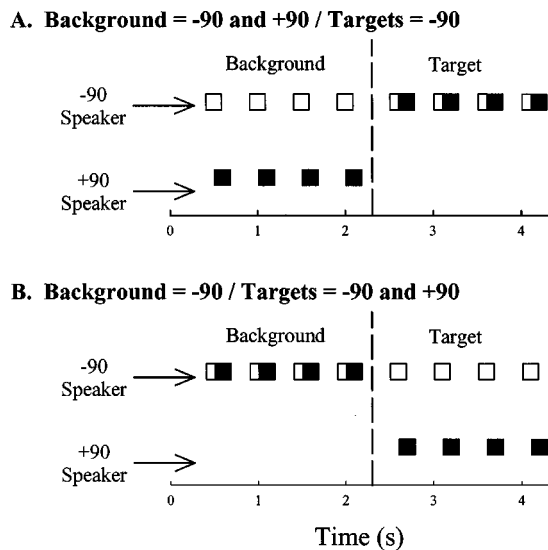


FIG. 7. Schematic of stimulus presentation. (a) Background stimuli were presented from -90 and $+90$. Targets were presented from -90 only. (b) Background stimuli were presented from -90 only. Targets were presented from $+90$ and -90 .

able to use them as cues for discrimination. If the spatial cues of the lag are completely inaccessible, then discrimination between a bilateral pair of background clicks and a unilateral pair of target clicks should be extremely difficult during the timecourses of localization dominance.

There should be differences in discrimination performance between the two tasks, however, because of the buildup and breakdown of localization dominance. In humans, the lag is reported as “fading out” after repeated presentations of lead-lag pairs (e.g., Clifton and Freyman, 1989). In budgerigars, discrimination of paired stimuli during localization dominance timecourses improves with repeated presentations of lead-lag pairs (Dent and Dooling, 2003). In this experiment, the task where the background stimuli are emitted from two locations [Fig. 7(a)] should be more difficult than the task where the background stimuli are emitted from only one location [Fig. 7(b)]. In the former condition, repeated presentations of the background are concurrent with decreased availability of spatial information from the lag. That would make the background and target very similar when the targets are finally presented.

In the condition where the background is unilateral and the bilateral targets are presented suddenly, it is possible that the spatial information from the lag will be immediately available (no fading out of the echo yet), and discrimination will be easier. If the buildup of localization dominance is important for the budgerigars, then there should be significant differences between the two conditions in this experiment.

B. Results

Like the results from experiment 4, there is not a monotonic increase of discrimination performance with increasing ISIs in either condition of this experiment (Fig. 8). The condition where the backgrounds were coming from two locations and the targets from one location (white circles) was

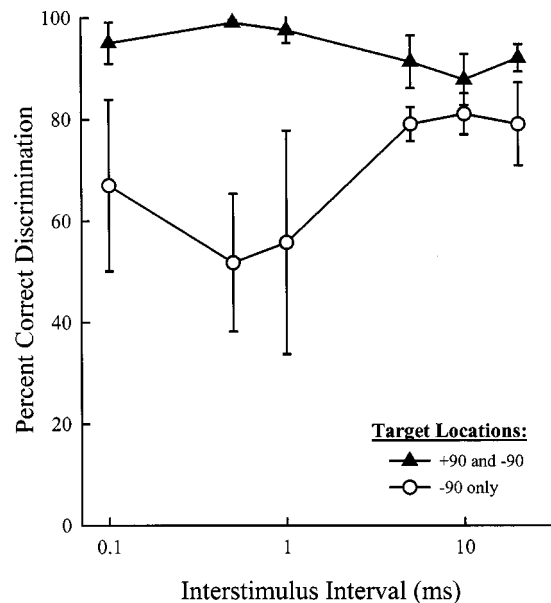


FIG. 8. Mean of four individuals tested on discrimination of a background from -90 versus targets from -90 and $+90$ (black triangles) or the reverse background/target condition (white circles). Error bars represent standard errors.

much more difficult than the reverse condition (black triangles), especially at the low and middle ISIs (0.1–5 ms), despite the only difference between the two data points at each ISI being a reversed background/target presentation condition. A two-way (ISI \times speaker conditions) repeated measures ANOVA showed that there was no significant effect of ISI [$F(5,33) = 0.51$, $p > 0.05$], there was an effect of speaker condition [$F(1,33) = 23.76$, $p < 0.001$], but there was not an interaction between the two variables [$F(5,33) = 1.83$, $p > 0.05$]. A posthoc Bonferroni t -test showed that within each ISI, there were differences in discrimination between the two background speaker conditions for the 0.1-, 0.5-, 1-, and 5-ms ISIs ($p < 0.05$), the same ISIs where localization dominance and summing localization were found to be operating in budgerigars (Dent and Dooling, 2003). There were no differences between the speaker conditions at the 10- and 20-ms ISIs ($p > 0.05$). This experiment showed that in the bilateral background stimulus condition only [Fig. 7(a)], the spatial properties of the lag sound were inaccessible and discrimination was difficult due to the buildup and breakdown of localization dominance.

VIII. DISCUSSION

A. Time-intensity trading

The present findings show how much of a change in intensity between the lead and lag stimuli was required to elicit localization dominance (0-ms ISI) or to abolish localization dominance (1- and 5-ms ISIs). In budgerigars, when simultaneous sounds were presented, the “echo” (not later in time, just less intense) needed to be 5–8 dB lower in intensity for high levels of discrimination. Hoeffding and Harrison (1979) found similar values in rats (~ 5 dB) although Kelly (1974) had somewhat lower values (~ 2.8 – 4.2 dB) for rats.

The echo (later sound source) being louder than the direct (first) sound source is not a natural occurrence in nature, but it is of interest that the results presented here are similar to those found for humans (e.g., Babkoff and Sutton, 1966; Blodgett *et al.*, 1956). Babkoff and Sutton (1966) found that in humans, echo thresholds changed from 3.8 to 2.4 ms when the lag was changed from 0 to 8 dB more intense than the lead stimulus. Increasing the intensity of the lag made the lag more detectable at shorter ISIs. Unfortunately, Kelly (1974) could not replicate this effect in rats, even when the lead click was attenuated by up to 20 dB. Leakey and Cherry (1957) found that, in humans, two clicks separated by a delay (so that the sound image was perceived to be coming from one side) could be perceptually recentered if the lag click was made louder than the lead click. With an ISI of 0.9 ms, that required intensity difference was found to be 8.6 dB, an ISI of 1.8 ms required an intensity difference of 10.6 dB, and an ISI of 2.2 ms required an intensity difference of 10.4 dB. At longer ISIs, more and more of an intensity increase in the lag was needed to recenter the image in humans. Mickey and Middlebrooks (2001) recently found that ISIs of 0.6 to 0.8 ms needed only a concurrent 5-dB increase in the lag to shift the perceived judgments back to the midline.

Although there are significant task differences between the experiments on humans and budgerigars, and the birds were only tested at two ISIs, it seems for both species that increasing the intensity of the lag does change the perceived location of the image. In humans, larger intensity differences are needed for larger time differences between lead and lag. It is difficult to tell from the experiments reported here whether the same holds true for budgerigars due to large between-subject variability and the testing of only two interstimulus intervals. In fact, it did not seem to hold for the 1- and 5-ms interstimulus intervals for the birds. Although in both cases the intensity differences could override the time differences, the intensity differences were similar for both time difference conditions. Generally, though, this experiment showed that in budgerigars, as in humans, intensity differences between lead and lag stimuli could elicit localization dominance—and they could even override localization dominance in certain conditions.

B. Discrimination of phantom sound images

The results here are among the first to show that animals perceive phantom sound images at a location that is different from the location of the presented sound sources. Although the task did not require budgerigars to identify the location of the auditory image, we tested the birds' ability to discriminate two simultaneously presented sounds from the sides from one sound presented from the midline. Failure to discriminate would indicate that budgerigars perceived a phantom midline image. The results here were intermediate, suggesting that, like humans, the birds are perceiving a difference in the fullness or richness quality of the auditory image when stimuli are emitted from two locations instead of one (e.g., Perrott *et al.*, 1989). However, discrimination performance was poor overall; the budgerigars could not easily distinguish between the two stimulus types. It was not a function of poor sound localization abilities, however, be-

cause discrimination of the midline stimulus from either of the side stimuli alone was good. The perception of a phantom image has also been described by Populin and Yin (1998) in cats and Keller and Takahashi (1996) in barn owls using different techniques. All of the studies so far suggest that animals perceive summing localization stimuli in a similar manner as humans.

C. Discrimination of paired stimuli

The present results show that the spatial cues of the stimuli are important in these experiments, along with the temporal cues. If the temporal onset of the stimuli were the critical cue, the results from experiment 3 with the bilateral target location would have been exactly the same as the results with the unilateral target location. This was not the case, however. The temporal onsets of these stimuli do not seem to be the only important factor in the discrimination between the backgrounds and targets. The control roved intensity condition showed that making potential intensity cues extremely unreliable also did not affect discrimination in this experiment. Taken together, the results show that both spatial and temporal cues are used by birds in these experiments, despite their poor sound localization abilities (Park and Dooling, 1991).

Experiment 4 used the reverse background/target conditions from experiment 3. Results here showed that increasing the temporal interval between paired clicks was enough to support discrimination of a pair of clicks from a single click. Also, under non-localization dominance conditions, both temporal and spatial cues could be used to improve discrimination. Interestingly, in humans, localization dominance has been reported to be stronger when the lead and lag are spatially coincident than when they are spatially distant (Litovsky and Shinn-Cunningham, 2001). This was not true for the budgerigars, however. Budgerigars were worse at distinguishing a pair of stimuli from different locations from a single target [experiment shown in Fig. 5(a)] than they were at distinguishing a pair of stimuli from the same location from a single target [experiment shown in Fig. 5(b)]. One interpretation is that there is less accessibility of the information from the echoes in the experiment shown in Fig. 5(a) than in the experiment shown in Fig. 5(b), counter to the results from humans. The differences between the budgerigars and humans may be related to the psychophysical tasks. Recall that humans were asked to identify the location of the lead, but budgerigars were asked only to discriminate between a paired background lead and single source lag. However, the results from Dent and Dooling (2003), where localization dominance was stronger and echo thresholds were later with larger speaker separation distances, further support the results from this experiment.

There were several potential outcomes of this experiment. First, if the echoes in this experiment were accessible, then performance for the experiment shown in Fig. 5(a) should have been better than performance for the experiment shown in Fig. 5(b). Second, if the echoes were inaccessible, then performance where the lag was located on the opposite side of the lead [shown in Fig. 5(a)] should have been very poor compared to the condition where the lag was at the

same location as the lead [shown in Fig. 5(b)] because the perception would have been of a single 1-ms click located at -90° in the second scenario and a very long or double click in the first scenario. Third, if only the spatial properties of the lag were inaccessible, then the performance between the two conditions should have been identical across ISIs. The temporal properties of the background would have been identical in the experiments shown in Figs. 5(a) and (b). The results from experiment 4 lie somewhere in-between the latter two possible outcomes. The addition of a second speaker location did hinder performance, suggesting that the echoes were at least partially inaccessible under localization dominance conditions. The echoes were not completely inaccessible, however, because discrimination performance was above chance level at all ISIs. The difference between these two conditions at certain ISIs strongly suggests that more than the spatial properties of the lag stimuli were inaccessible for budgerigars.

Of course there are other possible cues to discrimination that the birds could have been using in these experiments. The overlapping stimuli from multiple locations could have created signals with complex acoustic features such as rippled spectra. Budgerigars are very good at discriminating between sounds with flat- and rippled-spectra (Amagai *et al.*, 1999). As a hedge against the birds using other potential acoustic cues, the stimuli were roved in intensity from presentation to presentation by ± 4 dB. This makes other potential acoustic cues unreliable as a basis for discrimination. The overall duration of the stimuli are another cue to discrimination that the birds could have used. The duration discrimination abilities of birds are generally not well known, especially for very short stimuli. The evidence to date suggests that birds probably need about a 10%–20% change in the overall duration of a stimulus to discriminate a change (Dooling and Haskell, 1978). The non-monotonic discrimination functions (not increasing with increasing separation delays) in the third and fourth experiments suggest that the budgerigars are using neither of these cues.

Finally, the results from experiment 5 strongly support the hypothesis that localization dominance is operating in these birds. The task of discriminating background from target in the first speaker condition [Fig. 7(a)] was identical and opposite to the discrimination of target from background in the second speaker condition [Fig. 7(b)]. The two discrimination functions should have been identical across all ISIs. Yet, there were differences in response levels AND these differences were large at the low and intermediate ISIs where summing localization and localization dominance operate in humans. The more difficult discrimination, where the background lead was at $+90^\circ$ and the background lag was at -90° [shown in Fig. 7(a)] was difficult because only the spatial location of the lag was inaccessible, and the birds could not discriminate the target trials where the lead and lags had the same temporal properties but different spatial properties. In the reverse condition, and because of the requirement by these birds for a buildup of localization dominance [found in the companion paper, Dent and Dooling (2003)], the backgrounds coming from only one side being easy to discriminate from the targets in two locations [shown

in Fig. 7(b)] was not surprising. Here, localization dominance was not operating on the background (as shown also in experiment 4), and the targets had additional spatial cues (shown in experiment 3) to aid in discrimination. This made discrimination easier in this condition [Fig. 7(b)] than in the reverse condition [Fig. 7(a)]. The results showing that discrimination between the two conditions was not different at the long ISIs, past the echo thresholds, lends further support for this line of reasoning.

IX. CONCLUSIONS

Overall, the results from these experiments lend further support that the precedence effect exists in all animals as a mechanism to resolve potential confusions that may arise from multiple, competing sound sources. These experiments showed that intensity and location are two more potential sources of confusion a bird may overcome when localizing a sound source. Taken together, the results from these experiments, as well as those from Dent and Dooling (2003), extend the database of behavioral data describing the many facets of the precedence effect and their similarities and differences across species. The experiments here further suggest that the precedence effect acts in a similar manner across animals regardless of general hearing or sound localization abilities, habitats, or evolutionary histories.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. DC-00198 to RJD and NRSAs from NIH (DC-00046 and MH-12698) to MLD. We are very grateful to B. Brittan-Powell, F. Kubke, M. Leek, D. Yager, C. Carr, C. Moss, R. Litovsky, T. Takahashi, and D. Tollin for their comments on earlier versions of this manuscript.

- Amagai, S., Dooling, R. J., Shamma, S., Kidd, T. L., and Lohr, B. (1999). "Detection of modulation in spectral envelopes and linear-rippled noises by budgerigars (*Melopsittacus undulatus*)," *J. Acoust. Soc. Am.* **105**, 2029–2035.
- Babkoff, H., and Sutton, S. (1966). "End point of lateralization for dichotic clicks," *J. Acoust. Soc. Am.* **39**, 87–102.
- Blauert, J. (1997). *Spatial Hearing* (MIT, Cambridge, MA).
- Blodgett, H. C., Wilbanks, W. A., and Jeffress, L. A. (1956). "Effect of large interaural time differences upon the judgment of sidedness," *J. Acoust. Soc. Am.* **28**, 639–643.
- Bronkhorst, A. W., and Houtgast, T. (1999). "Auditory distance perception in rooms," *Nature* **397**, 517–520.
- Clifton, R. K., and Freyman, R. L. (1989). "Effect of click rate and delay on breakdown of the precedence effect," *Percept. Psychophys.* **46**, 139–145.
- Cranford, J. L. (1982). "Localization of paired sound sources in cats: Effects of variable arrival times," *J. Acoust. Soc. Am.* **72**, 1834–1835.
- Dent, M. L., and Dooling, R. J. (2003). "Investigations of the precedence effect in budgerigars: Effects of stimulus type, intensity, duration, and location," *J. Acoust. Soc. Am.* **113**, 2146–2158.
- Dooling, R. J., and Haskell, R. J. (1978). "Auditory duration discrimination in the parakeet (*Melopsittacus undulatus*)," *J. Acoust. Soc. Am.* **63**, 1640–1643.
- Freyman, R. L., McCall, D. M., and Clifton, R. K. (1998). "Intensity discrimination for precedence effect stimuli," *J. Acoust. Soc. Am.* **103**, 2031–2041.
- Griffin, D. R. (1958). *Listening in the Dark* (Yale U. P., New Haven).
- Griffin, D. R., and Suthers, R. A. (1970). "Sensitivity of echolocation in cave swiftlets," *Biol. Bull.* **139**, 495–501.
- Haas, H. (1951). "On the influence of a single echo on the intelligibility of speech," *Acustica* **1**, 49–58.

- Haas, H. (1972). "The influence of a single echo on the audibility of speech," *J. Audio Eng. Soc.* **20**, 146–159.
- Hoeffding, V., and Harrison, J. M. (1979). "Auditory discrimination: Role of time and intensity in the precedence effect," *J. Exp. Anal. Behav.* **32**, 157–166.
- Keller, C. H., and Takahashi, T. T. (1996). "Responses to simulated echoes by neurons in the barn owl's auditory space map," *J. Comp. Physiol., A* **178**, 499–512.
- Kelly, J. B. (1974). "Localization of paired sound sources in the rat: Small time differences," *J. Acoust. Soc. Am.* **55**, 1277–1284.
- Konishi, M., and Knudsen, E. I. (1979). "The oilbird: hearing and echolocation," *Science* **204**, 425–427.
- Leakey, D. M., and Cherry, E. C. (1957). "Influence of noise upon the equivalence of intensity differences and small time delays in two-loudspeaker systems," *J. Acoust. Soc. Am.* **29**, 284–286.
- Litovsky, R. Y., and MacMillan, N. A. (1994). "Sound localization precision under conditions of the precedence effect: Effects of azimuth and standard stimuli," *J. Acoust. Soc. Am.* **96**, 752–758.
- Litovsky, R. Y., and Shinn-Cunningham, B. G. (2001). "Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression," *J. Acoust. Soc. Am.* **109**, 346–358.
- Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (1999). "The precedence effect," *J. Acoust. Soc. Am.* **106**, 1633–1654.
- Mershon, D. H., Ballenger, W. L., Little, A. D., McMurtry, P. L., and Buchanan, J. L. (1989). "Effects of room reflectance and background noise on perceived auditory distance," *Perception* **18**, 403–416.
- Mickey, B. J., and Middlebrooks, J. C. (2001). "Responses of auditory cortical neurons to pairs of sounds: correlates of fusion and localization," *J. Neurophysiol.* **86**, 1333–1350.
- Naguib, M. (1995). "Auditory distance assessment of singing conspecifics in Carolina wrens: The role of reverberation and frequency attenuation," *Anim. Behav.* **50**, 1297–1307.
- Nelson, B. S., and Stoddard, P. K. (1998). "Accuracy of auditory distance and azimuth perception by a passerine bird in natural habitat," *Anim. Behav.* **56**, 467–477.
- Park, T. J., and Dooling, R. J. (1991). "Sound localization in small birds: Absolute localization in azimuth," *J. Comp. Psychol.* **105**, 125–133.
- Perrott, D. R., Strybel, T. Z., and Manligas, C. L. (1987). "Conditions under which the Haas precedence effect may or may not occur," *J. Aud. Res.* **27**, 59–72.
- Perrott, D. R., Marlborough, K., Merrill, P., and Strybel, T. Z. (1989). "Minimum audible angle thresholds obtained under conditions in which the precedence effect is assumed to operate," *J. Acoust. Soc. Am.* **85**, 282–288.
- Populin, L. C., and Yin, T. C. T. (1998). "Behavioral studies of sound localization in the cat," *J. Neurosci.* **18**, 2147–2160.
- Snow, W. B. (1954). "Effect of arrival time on stereophonic localization," *J. Acoust. Soc. Am.* **26**, 1071–1074.
- Tollin, D. J., Populin, L. C., and Yin, T. C. T. (2000). "Neural correlates of the precedence effect and echo threshold in the inferior colliculus of the behaving cat," paper presented at the 30th Annual Meeting of the Society for Neuroscience, New Orleans, LA.
- Wallach, H., Newman, E. B., and Rosenzweig, M. R. (1949). "The precedence effect in sound localization," *Am. J. Psychol.* **62**, 315–336.

Diversity in noise-induced temporary hearing loss in otophysine fishes

Sonja Amoser and Friedrich Ladich^{a)}

Institute of Zoology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria

(Received 12 July 2002; revised 13 December 2002; accepted 13 January 2003)

The effects of intense white noise (158 dB *re* 1 μ Pa for 12 and 24 h) on the hearing abilities of two otophysine fish species—the nonvocal goldfish *Carassius auratus* and the vocalizing catfish *Pimelodus pictus*—were investigated in relation to noise exposure duration. Hearing sensitivity was determined utilizing the auditory brainstem response (ABR) recording technique. Measurements in the frequency range between 0.2 and 4.0 kHz were conducted prior and directly after noise exposure as well as after 3, 7, and 14 days of recovery. Both species showed a significant loss of sensitivity (up to 26 dB in *C. auratus* and 32 dB in *P. pictus*) immediately after noise exposure, with the greatest hearing loss in the range of their most sensitive frequencies. Hearing loss differed between both species, and was more pronounced in the catfish. Exposure duration had no influence on hearing loss. Hearing thresholds of *C. auratus* recovered within three days, whereas those of *P. pictus* only returned to their initial values within 14 days after exposure in all but one frequency. The results indicate that hearing specialists are affected differently by noise exposure and that acoustic communication might be restricted in noisy habitats. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1557212]

PACS numbers: 43.80.Nd, 43.64.Ri, 43.64.Wn [WA]

I. INTRODUCTION

In the aquatic environment, sound is one of the most important signal carriers, for it is transported five times faster than in air, is not attenuated as fast as light or chemical substances, and is propagated over large distances due to existing sound channels (Hawkins and Myrberg, 1983). There is a natural background noise due to currents, surf, rain, seismic events and sounds of biological origin (e.g., snapping shrimps, vocalizing animals). In the last decades, however, this has been augmented by an ever-increasing amount of anthropogenic noise, such as noise from shipping, hydroelectric power plants, drilling, or seismic surveys. For example, Andrew *et al.* (2002) compared ocean ambient sound data off the Californian coast from the 1960s with data from the 1990s and found an increase of up to 10 dB between 20 and 400 Hz.

There is growing concern that anthropogenic noise affects aquatic animals, in particular whales. Certain whale species change their vocalizing or migration behavior when vessels approach (Richardson *et al.*, 1986, 1995; Lesage *et al.*, 1999), during the presence of low-frequency sonar (Miller *et al.*, 2000) or during acoustic thermometry of ocean climate (ATOC) transmissions (Frankel and Clark, 2000).

Intense noise exposure can impair auditory abilities by causing elevated auditory thresholds or damage to ear organs. In the last decades, the study of sound- or noise-induced hearing loss (NIHL) has concentrated on humans and mammals (humans: e.g., Hamernik *et al.*, 1982; Bauer *et al.*, 1991; mammals: e.g., Clark, 1991; Griffiths *et al.*, 1994). Only a few studies have dealt with other taxa such as

birds (Saunders and Dooling, 1975; Dooling *et al.*, 1997).

Fish, as other aquatic animals, are highly dependent on the auditory system. By simply listening to the sounds coming from different sources, they obtain the information relevant for survival, finding mates and prey, or avoiding predators (Hawkins and Myrberg, 1983; Popper and Fay, 1993; Ladich, 1999; Fay and Popper, 2000). Based on current knowledge, all fish are able to perceive low-frequency sounds (“hearing generalists”), whereas some groups have developed accessory hearing structures to enable them to detect high frequencies up to 4 kHz (“hearing specialists”). Hearing enhancement is based on a connection between an air-filled cavity within the body and the inner ear. Within otophysans, vibrations of the swimbladder are transmitted via 1–4 bony ossicles (Weberian apparatus) to the inner ear.

Noise can lead to behavioral changes in fish. This was investigated mainly in commercially important marine fish species (Blaxter and Hoss, 1981; Blaxter *et al.*, 1981; Schwarz and Greer, 1984). Pearson *et al.* (1992) observed alarm and startle responses of rockfish, *Sebastes* spp., in the presence of air-gun sounds.

The effects of intense sound on the auditory epithelia have rarely been investigated. Enger (1981) observed damages to the sensory epithelia of the inner ear of the cod *Gadus morhua* after exposure to intense noise. Hastings *et al.* (1996) exposed the hearing generalist *Astronotus ocellatus* to 60- and 300-Hz pure tones at three different sound pressure levels (SPLs) to determine the effects on the sensory epithelia of the inner ear and the lateral line.

The detection of a signal is frequently impaired by the presence of another. Several studies described a decrease in the ability to detect a signal in the presence of a second (masking) sound for the goldfish *C. auratus* and the cod *G. morhua*. In both species, the masking effect of a sound (pure

^{a)} Author to whom correspondence should be addressed. Electronic mail: friedrich.ladich@univie.ac.at

tone or noise band) is confined to the frequency region of the signal to be detected (Buerkle, 1968, 1969; Fay, 1974; Hawkins and Chapman, 1975; Fay *et al.*, 1978); threshold shifts occurred during nonsimultaneous masking, as well (Popper and Clarke, 1979).

Another approach is to assess the influence of intense noise on auditory thresholds after exposure. In a behavioral study, Popper and Clarke (1976) stimulated goldfish with intense pure tones (0.3, 0.5, 0.8 and 1.0 kHz) for 4 h and observed temporary threshold shifts (TTS) and complete recovery within 24 h. Scholik and Yan (2001, 2002a) exposed the fathead minnow *Pimephales promelas* to intense white noise or boat noise for 2 to 24 hours. White noise bands (0.3–4 kHz) increased the auditory thresholds at five out of eight frequencies, whereas boat noise only affected three frequencies. Recovery was dependent on the frequency and exposure duration. In a subsequent study, Scholik and Yan (2002b) could not observe any significant changes in the auditory sensitivity of the bluegill sunfish *Lepomis macrochirus*, a hearing generalist.

The goal of our present study was threefold: (1) to investigate the immediate effects of unfiltered white noise on the hearing thresholds of two hearing specialists, (2) to assess the course of recovery and (3) to clarify any exposure duration effects. We chose two otophysan species, the goldfish *C. auratus* and the catfish *P. pictus*. The latter produces two types of sound: low-frequency drumming sounds by muscular vibrations of the swimbladder and broad-band pulsed sounds by stridulating the pectoral spines (Ladich 1999, 2001). *C. auratus* and *P. pictus* were chosen to compare a nonvocal and a vocal fish species and to assess the influence of the hearing loss on acoustic communication.

II. MATERIAL AND METHODS

A. Animals

Twelve specimens of *C. auratus* (78–92 mm standard length; 14.5–20.9 g body mass) and 19 of *P. pictus* (43–79 mm; 2.8–7.1 g) were obtained from local aquarium fish suppliers. Animals were maintained in externally filtered aquaria and a 12:12 h light:dark cycle was maintained. All aquaria were planted, equipped with half flowerpots as hiding places, and the bottom was covered with sand. Fish were fed live *Tubifex* and commercially prepared food (TetraMin®). Efforts were made to provide a quiet environment (125.4 ± 1.4 dB, no submerged pumps or air stones). The animal-use protocol used in this study was approved by the Austrian Commission on Experiments in Animals (GZ 68.210/30-Pr/4/2001).

B. Noise exposure

The animals were exposed singly to unfiltered white noise at 157.8 ± 1.6 dB ($x \pm S.D.$) *re* 1 μ Pa for 12 and 24 h in a plastic bucket (20 cm height, 24 cm diameter, 15 cm water depth). Fish could move freely within the bucket. White noise was generated by a noise generator (IVIE Electronics IE 20B), sent to a 24-band equalizer (Alesis MEQ 230) to obtain a flat noise spectrum, and fed to a power amplifier

(Brüel & Kjaer 2713) that drove an underwater loudspeaker (University Sound UW 30) situated on the bottom of the bucket.

C. Auditory sensitivity measurements

The auditory sensitivity was measured utilizing the auditory brainstem response (ABR) recording technique, following the ABR recording protocol recently described in Ladich (1999) and Wysocki and Ladich (2001). Therefore, only a brief summary of the basic technique will be given here.

All fish were measured at least 6 days prior to noise exposure to obtain a baseline audiogram. Control tests keeping goldfish ($N=6$) in the noise exposure bucket without presenting noise revealed that manipulations *per se* did not lower auditory sensitivity at any frequency. After the end of the 12- or 24-h noise exposure period, hearing thresholds were tested immediately (day 0). This measurement was completed within 3 h. Both species were remeasured 3 and 7 days following the exposure to obtain the recovery hearing thresholds (day 3, day 7). In addition *P. pictus* was measured after 14 days of recovery (day 14).

1. Experimental setup

Test subjects were secured in a half-bowl shaped 11-l plastic tub (33 cm diameter, 13 cm height, 1 cm layer of fine sand) filled with water and adjusted so that the head was approximately 1 mm above the water surface; a respiration pipette was inserted into the subject's mouth. Respiration was achieved through a simple temperature-controlled (24 °C), gravity-fed water circulation system. In order to immobilize animals and to reduce the myogenic noise level they were injected with a curariform agent (galamine triethiodide—Flaxedil), with a required dosage of $1\text{--}2 \mu\text{g g}^{-1}$ for *C. auratus* and $7\text{--}10 \mu\text{g g}^{-1}$ for *P. pictus*. The plastic tub was positioned on an air table (TMC Micro g-63-450), which rested on a vibration-isolated concrete plate. The entire setup was enclosed in a walk-in soundproof room, which was constructed as a Faraday cage (interior dimensions: $3.2 \times 3.2 \times 2.4$ m).

The recording of ABRs was performed using silver wire electrodes (0.25 mm diameter), which were pressed firmly against the skin. The portion of the skin above the water surface was covered by a small piece of Kimwipes tissue paper to keep it moist in order to ensure proper contact between the skin and electrodes during the experiment. The recording electrode was placed on the midline of the skull over the region of the medulla. The reference electrode was placed cranially between the nares.

2. ABR-recording apparatus and stimulus presentation

Both sound stimuli presentation and ABR waveform recording were accomplished using a Tucker-Davis Technologies (Gainesville, FL) modular rack-mount system controlled by an optically-linked Pentium PC containing a TDT digital processing board and running TDT Bio-Sig 2.2 Software.

Sound stimuli consisted of tone bursts. The number of cycles in a tone burst (2–8) was adjusted according to the

frequency in order to obtain the best combination of stimulus rise time (shorter rise time=greater efficacy at generating ABRs) and peak frequency bandwidth (longer duration=sharper spectral peak) (Silman and Silverman, 1991). Sound stimuli waveforms were constructed using TDT Sig-Gen software and fed through a DA1 digital-analog converter, a PA4 programmable attenuator, and a power amplifier (Denon PMA 715R). A dual-cone speaker (Tannoy System 600, frequency response 50 Hz to 15 kHz \pm 3 dB), suspended in air, was mounted 1 m above the test subject. A hydrophone (Brüel & Kjaer 8101, frequency range: 1 Hz to 80 kHz \pm 2 dB; voltage sensitivity: -184 dB *re* 1 V/ μ Pa) was placed close to the right side of the animals (2 cm apart) in order to determine absolute stimulus SPLs underwater in close vicinity of the subjects.

For each test condition, 1000 stimuli at opposing polarities (90° and 270°) were presented and averaged by the Bio-Sig software. The ABR traces of signals presented at different polarities do not cancel out each other when averaged. This is contrary to sound pressure waveforms when averaged (see Fig. 3 in Kenyon *et al.*, 1998) and efficiently eliminates artifacts. SPL was reduced in 4-dB steps until the ABR waveform was no longer apparent. The lowest SPL, for which a repeatable ABR trace could be obtained, as determined by overlaying replicate traces, was considered the threshold. This method of visual inspection correlation is the traditional means of determining thresholds in ABR audiometry (Kileny and Shea, 1986; Gorga *et al.*, 1988; Hall, 1992) and proved also to agree well with the correlation coefficient method developed by Yan (1998).

Animals were tested at frequencies of 0.2, 0.3, 0.5, 0.8, 1.0, 2.0, and 4.0 kHz presented in random order.

D. ABR waveform analysis and statistics

ABR waveform characteristics (peak-to-peak amplitude and latency of first negative peak) were analyzed in three to six individuals of each species before and immediately after exposure to white noise at 20 dB above hearing level (20 dB HL, baseline) and tested with an unpaired *T*-test.

Hearing thresholds of six individuals of each species were determined at both exposure durations (12 and 24 h). Contrary to goldfish, catfish had to be measured five times due to longer recovery periods. In order to avoid mortality seven catfish were only measured twice, four thrice, five four times and three five times. Threshold values from all individuals of one group as measured at seven different frequencies were compared between pre- and all postexposure measurements by a two-way analysis of variance (ANOVA) using a general linear model where one factor was exposure and the other was frequency. The exposure factor should indicate overall differences between the pre- and postexposure measurements; in combination with the frequency factor, it should indicate if different tendencies exist at different frequencies of the audiograms. This was followed by Scheffe's multiple comparison procedure. Significant differences between measurements indicate that noise- and recovery effects outweigh interindividual differences.

One-way ANOVAs, followed by Scheffe's multiple comparison procedure, were applied to evaluate the different

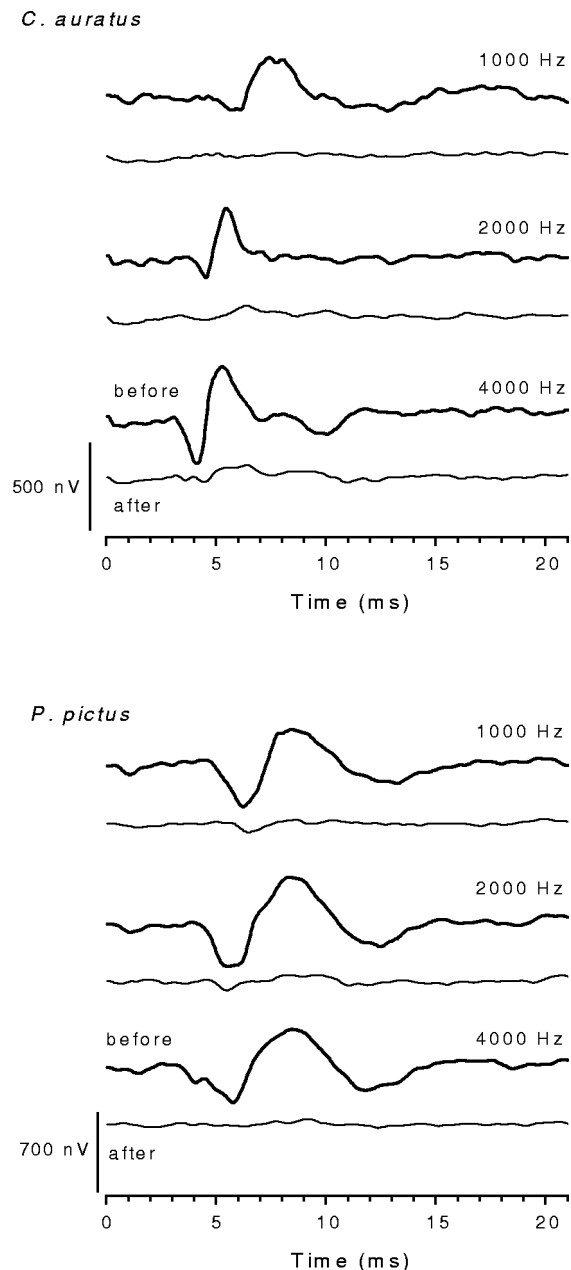


FIG. 1. Averaged ABR waveforms of *C. auratus* ($n=5-6$) and *P. pictus* ($n=3-6$) before (bold traces) and immediately after (thin traces) 12-h noise exposure, obtained in response to tone bursts of three representative frequencies 20 dB above baseline hearing levels (20 dB HL). Averaged traces of two different polarities are shown.

effects of noise exposure at different frequencies (*P*-value adjusted to 0.007, due to the seven frequencies tested) and to analyze possible exposure duration effects. Baseline auditory thresholds measured prior to noise exposure were pooled because a one-way ANOVA did not reveal any significant differences between the 12 and the 24 h group, either in *C. auratus* ($F_{1,90}=0.55$, n.s.) or in *P. pictus* ($F_{1,111}=3.01$, n.s.). All statistical tests were run using SPSS version 10.0.

III. RESULTS

A. ABR waveforms

ABR waveforms were obtained from all individuals of the two species, both before and after noise exposure. As an

TABLE I. Mean latencies, change of latencies ($\Delta \text{Lat} = \text{Lat 2} - \text{Lat 1}$), mean amplitudes and decrease of amplitudes ($\Delta \text{Ampl} = \text{Ampl 2} - \text{Ampl 1}$) of auditory brainstem responses (ABRs) measured before (Lat 1, Ampl 1) and immediately after exposure (Lat 2, Ampl 2). Latencies were defined as the distance between the beginning of the stimulus and the first negative peak of the ABR response. Amplitudes were measured from the lowest negative to the highest positive peak of the waveform. All measurements were performed at 20 dB HL.

		Lat 1 (ms)		Lat 2 (ms)		Δ Lat (ms)		Ampl 1 (nV)		Ampl 2 (nV)		Δ Ampl (nV)	
kHz	<i>n</i>	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
<i>C. auratus</i>													
12 h													
1.0	6	1.0	0.12	1.6	0.17	0.6	0.20	305	18.7	62	07.1	−243	24.5
2.0	5	1.8	0.16	2.4	0.91	0.7	0.11	555	130.4	90	24.4	−464	125
4.0	5	1.4	0.10	1.7	0.08	0.4	0.11	469	124.6	227	104.9	−242	75.7
24 h													
1.0	6	1.7	0.11	2.9	0.23	1.2	0.26	362	45.2	46	12.9	−316	38.6
2.0	6	1.6	0.04	2.3	0.21	0.7	0.23	354	28.6	79	17.6	−275	20.3
4.0	6	1.4	0.09	2.2	0.20	0.8	0.19	303	40.0	119	16.0	−184	50.6
<i>P. pictus</i>													
12 h													
1.0	6	2.4	0.13	2.68	0.17	0.29	0.11	716.8	104.2	162.1	41.4	−554.7	112.8
2.0	6	2.5	0.31	3.30	0.32	0.85	0.27	741.0	151.5	167.0	68.3	−574.0	118.6
4.0	6	1.1	0.14	1.70	0.15	0.59	0.16	657.2	127.3	103.3	08.4	−553.9	128.4
24 h													
1.0	3	2.2	0.15	2.50	0.09	0.27	0.15	1026.7	141.2	283.3	161.1	−743.6	299.4
2.0	3	0.9	0.18	3.35	...	2.55	...	591.3	277.5	179.7	...	−172.9	...
4.0	3	1.0	0.04	2.42	0.44	1.40	0.41	438.0	117.0	89.2	40.2	−348.8	156.5

example, Fig. 1 shows the averaged ABR traces of *C. auratus* and *P. pictus* following 12-h noise exposure. After exposure, the ABRs clearly diminished at all frequencies.

Noise exposure lowered the maximum ABR amplitudes by 240 to 320 nV in *C. auratus* and by 170 to 740 nV in *P. pictus* (Table I). In contrast, latencies increased after exposure by 0.6 ms to 1.2 ms in *C. auratus* and by 0.3 ms to 2.6 ms in *P. pictus* (Table I). The increase in latencies and the decrease of peak-to-peak amplitude did not differ between both exposure durations in either of the two species.

B. Auditory thresholds

1. *Carassius auratus*

The auditory thresholds of *C. auratus* shifted upwards after 12 h and after 24 h exposure (Fig. 2, Table II). Comparison between audiograms revealed significant overall differences between thresholds before and after both exposure durations (two-way ANOVA: 12 h: $F_{3,180}=202.28$, $P<0.001$; 24-h: $F_{3,168}=240.52$, $P<0.001$). A significant interaction between noise exposure and frequencies could only be found after 24-h exposure (two-way ANOVA: $F_{18,168}=3.43$, $P<0.001$), but not after 12-h exposure (two-way ANOVA: $F_{18,180}=1.43$, n.s.), indicating that only for the 24-h exposure were the changes of the auditory sensitivity different at different frequencies. There were no significant differences of overall threshold shift between the two exposure durations (one-way ANOVA: $F_{1,88}=0.001$, n.s.).

Thresholds at all seven frequencies tested were significantly elevated after noise exposure compared to baseline values, tested by a one-way ANOVA for each frequency separately (Fig. 2, Table III). After three days of recovery,

auditory thresholds of *C. auratus* were not significantly elevated compared to baseline values (Fig. 3, Table III), suggesting that hearing abilities recovered fully. Similar to day 0, no differences between the two exposure durations could be found after 3 days of recovery (one-way ANOVA: $F_{1,90}=0.27$, n.s.).

2. *Pimelodus pictus*

In *P. pictus* a significant elevation of the audiogram was observed (Fig. 4, Table IV), tested by a two-way ANOVA, after both 12 h ($F_{4,252}=127.01$, $P<0.001$) and 24-h exposure ($F_{4,245}=113.44$, $P<0.001$). In contrast to *C. auratus* a significant interaction between noise exposure and frequencies was found after both exposure durations (two-way ANOVA: 12 h: $F_{24,252}=10.01$, $P<0.001$; 24 h: $F_{24,245}=7.81$, $P<0.001$). Again, there were no significant differences of overall threshold shift between the two exposure durations (one-way ANOVA: $F_{1,90}=0.83$, n.s.).

Contrary to the goldfish, where all seven frequencies tested were significantly elevated after both exposure durations, this was not the case for *P. pictus*. After 12-h exposure, six out of seven frequencies (except 0.3 kHz) were significantly elevated compared to preexposure values, which was tested with one-way ANOVA for each frequency. After 24-h exposure, only four frequencies (0.8–4.0 kHz) were significantly different from the baseline thresholds (Fig. 4, Table V).

In contrast to *C. auratus* the audiogram of *P. pictus* had not fully recovered at all seven frequencies within three days following exposure (Fig. 5, Table V). On day three, two out of seven frequencies had not returned to the baseline level following 12- and 24-h exposure. Even on day 7, one fre-

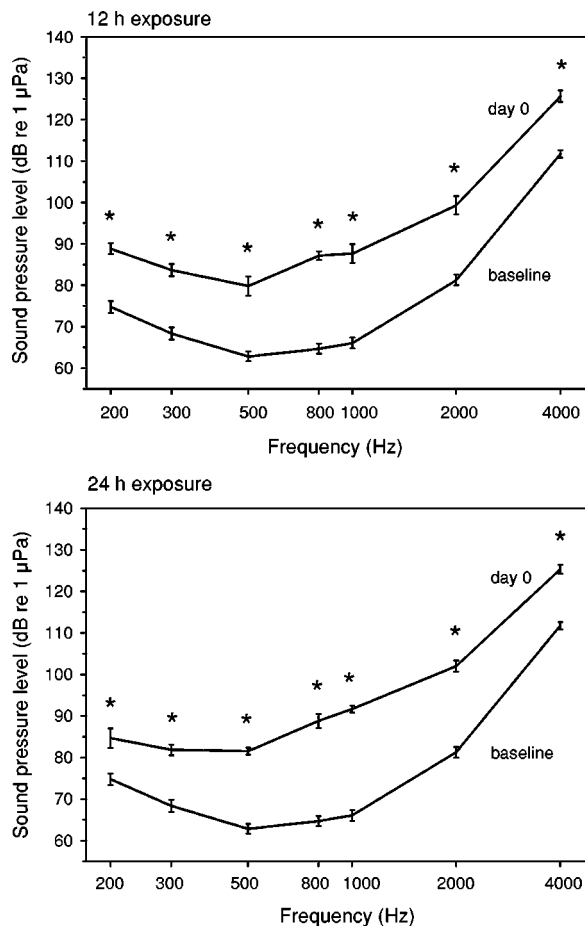


FIG. 2. Audiograms of *C. auratus* before (baseline) and immediately after (day 0) white noise exposure (upper figure: 12-h noise exposure; lower figure: 24-h exposure). Asterisks indicate statistically significant differences following adjustment to the number of frequencies tested ($P \leq 0.007$).

quency (4.0 kHz) was still elevated after both exposure durations. Another seven days later the same frequency was still significantly elevated following the 24-h exposure duration (Table V), but there were no significant overall differences on that day when comparing the whole audiogram. Comparison of recovery audiograms between the two exposure durations did not reveal any significant differences (one-way ANOVA: day 3: $F_{1,83}=3.58$, n.s.; day 7: $F_{1,83}=1.99$, n.s.; day 14: $F_{1,83}=0.00$, n.s.).

IV. DISCUSSION

A. ABR waveforms

The decrease in the maximum ABR amplitude as well as the increase in latency of the first negative peak after noise exposure clearly indicates that the auditory sensitivity of both species diminished. This resulted in higher auditory thresholds.

Scholik and Yan (2001, 2002a) exposed fathead minnows *Pimephales promelas* to white noise (0.3–4.0 kHz) and to boat engine noise (0.3–6.0 kHz with a major peak at 1.3 kHz) and revealed a decrease of ABR amplitudes after exposure. Unfortunately, they did not quantify the amount of amplitude shift or report whether the latencies shifted, so a further comparison with our data was not possible.

TABLE II. Hearing threshold values (dB re 1 μ Pa) of *C. auratus* before (baseline), immediately after (day 0) noise exposure and after 3 and 7 days of recovery (day 3, day 7).

kHz	Baseline		Day 0		Day 3		Day 7	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
12-h exposure								
0.2	74.75	1.39	88.83	1.35	72.17	2.50	68.67	2.60
0.3	68.33	1.46	83.67	1.50	70.33	1.89	68.33	4.48
0.5	62.83	1.17	79.83	2.30	64.33	1.05	65.33	1.20
0.8	64.67	1.23	87.17	1.01	66.50	1.77	64.00	1.55
1.0	66.08	1.28	87.67	2.26	68.33	1.31	63.00	1.73
2.0	81.25	1.27	99.33	2.20	81.67	2.16	77.33	2.40
4.0	111.75	0.90	125.67	1.36	112.17	1.62	109.67	0.33
24-h exposure								
0.2	74.75	1.39	84.67	2.39	71.67	1.56	69.33	2.60
0.3	68.33	1.46	81.83	1.30	68.33	1.56	63.33	0.33
0.5	62.83	1.17	81.50	0.85	61.83	1.05	62.00	1.00
0.8	64.67	1.23	88.83	1.70	65.17	0.17	59.33	1.20
1.0	66.08	1.28	91.67	0.80	68.33	1.89	63.00	3.51
2.0	81.25	1.27	102.00	1.37	80.17	1.60	81.00	1.15
4.0	111.75	0.90	125.33	1.09	109.50	1.88	109.67	3.33

In a mammalian study on the effects of intense sounds on hearing, Griffiths *et al.* (1994) found small latency increases in sheep fetuses exposed to 120 dB (re 20 μ Pa) broadband noise for 16 h in utero. They also observed recovery of the latencies to preexposure values.

B. Noise effects on auditory thresholds

Exposing the two otophysan species *C. auratus* and *P. pictus* to white noise at 158 dB SPL (re 1 μ Pa) for 12 and 24 h significantly elevated the auditory thresholds in both species. The amount of threshold shift clearly differed between both species, being more pronounced in the catfish (5–32 dB versus 10–24 dB in the goldfish). In *C. auratus* the greatest threshold shift was observed at 0.8 and 1.0 kHz, while in *P. pictus* the higher frequencies (2.0 and 4.0 kHz) were most affected by noise exposure. For both species the greatest effects of noise exposure were seen in their most sensitive hearing range. The larger effect in the catfish was most likely due to the larger sensitivity at higher frequencies. As reviewed by Ladich and Bass (2003) catfishes in general possess very flat hearing curves and a high frequency hearing sensitivity, which is even unusual among hearing specialists such as cyprinids.

Similar to our data Scholik and Yan (2001) observed in the cyprinid *P. promelas* after exposure to white noise for 1 to 24 h at 142 dB SPL (re 1 μ Pa), the greatest threshold shift in the fish's most sensitive frequency range (0.8–2.0 kHz). The amount of threshold shift in *P. promelas* ranged from 11 to 20 dB, which is comparable to the shift observed in our study in *C. auratus*. Contrary to our results, where all frequencies of *C. auratus* were significantly elevated immediately after noise exposure, in *P. promelas* 0.5, 2.5, and 4.0 kHz were not significantly affected. This difference is most likely due to different noise bandwidths applied. We exposed the fish to unfiltered white noise, while Scholik and Yan (2001) used white noise with a bandwidth ranging from 0.3

TABLE III. Differences of mean hearing thresholds (dB) of *C. auratus* at the seven frequencies tested calculated by one-way ANOVA for each frequency followed by pair wise multiple comparison procedure (Scheffé).

kHz	Differences						$F_{3,23}$	P
	Baseline—day 0		Baseline—day 3		Baseline—day 7			
	Mean	SE	Mean	SE	Mean	SE		
12-h exposure								
0.2	14.1 ^a	2.20	−2.6	2.63	−6.1	3.08	17.73	<0.001
0.3	15.3 ^a	2.34	2.0	2.47	0.0	3.60	13.72	<0.001
0.5	17.0 ^a	2.31	1.5	1.83	2.5	2.47	24.98	<0.001
0.8	22.5 ^a	1.89	1.8	2.14	−0.67	2.57	53.06	<0.001
1.0	21.6 ^a	2.40	2.3	2.04	−3.1	2.73	38.39	<0.001
2.0	18.1 ^a	2.37	0.4	2.35	−3.9	2.82	23.63	<0.001
4.0	13.9 ^a	1.59	0.4	1.70	−2.1	1.85	29.91	<0.001
24-h exposure								
0.2	9.9 ^a	2.58	−3.1	2.27	5.4	3.08	10.09	<0.001
0.3	13.5 ^a	2.28	0.0	2.36	5.0	3.01	18.95	<0.001
0.5	18.7 ^a	1.78	−1.0	1.83	−0.8	2.45	53.92	<0.001
0.8	24.2 ^a	2.11	0.5	1.77	−5.3	2.58	76.55	<0.001
1.0	25.6 ^a	1.91	2.3	2.24	−3.1	3.04	56.53	<0.001
2.0	20.8 ^a	2.05	−1.1	2.13	−0.3	2.67	44.94	<0.001
4.0	13.6 ^a	1.49	−2.3	1.82	−2.1	2.35	24.66	<0.001

^aStatistically significant differences following adjustment to the number of frequencies tested ($P < 0.007$). Positive threshold differences mean a higher auditory threshold of the second group of a pairing at this particular frequency.

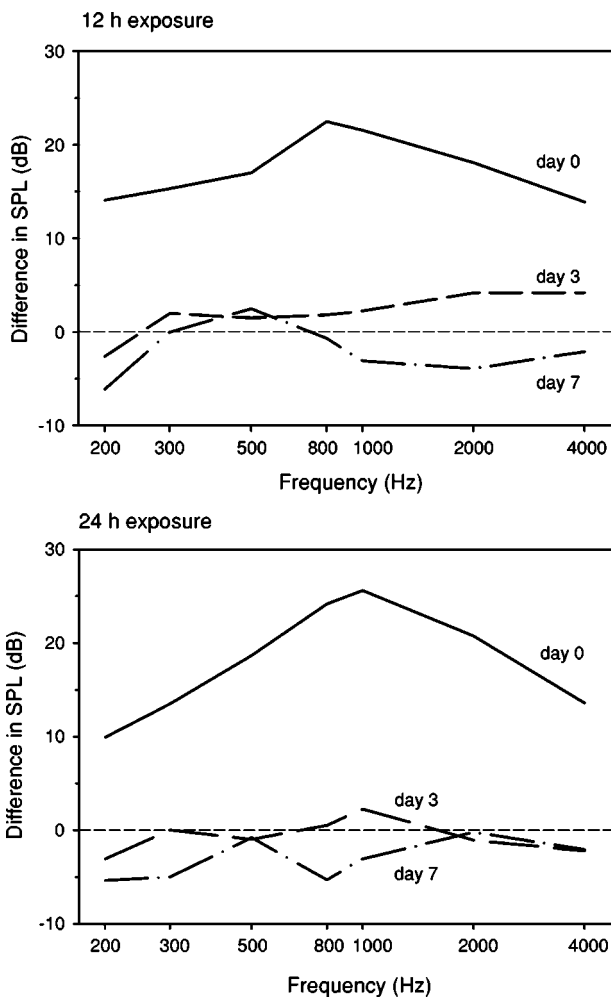


FIG. 3. Amount of threshold shift in *C. auratus* immediately after exposure (day 0) and after three (day 3) and seven (day 7) days of recovery. The dashed line at zero indicates baseline values.

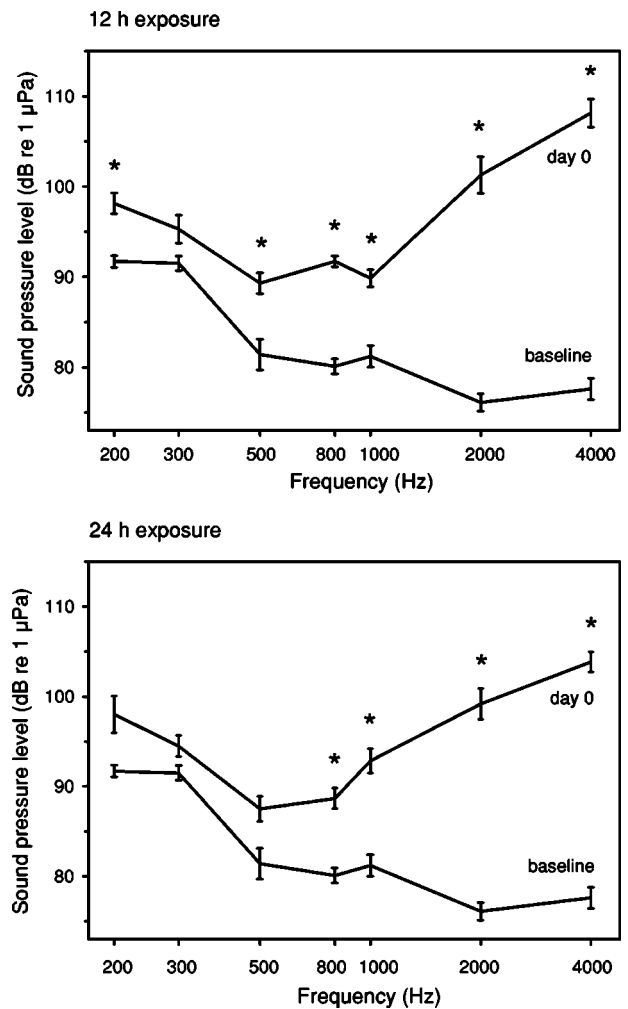


FIG. 4. Audiograms of *P. pictus* before (baseline) and after (day 0) white noise exposure. Asterisks indicate statistically significant differences following adjustment to the number of frequencies tested ($P \leq 0.007$).

TABLE IV. Hearing threshold values (dB *re* 1 μ Pa) of *P. pictus* before (baseline), immediately after (day 0) noise exposure and after 3, 7, and 14 days of recovery (day 3, day 7, day 14).

Frequency (kHz)	Baseline		Day 0		Day 3		Day 7		Day 14	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
12-h exposure										
0.2	91.13	0.55	98.14	1.16	92.00	0.97	91.33	1.54	89.67	1.15
0.3	89.88	0.80	95.29	1.55	89.17	1.05	93.17	1.05	89.00	1.65
0.5	78.19	1.67	89.29	1.15	79.00	2.19	80.67	0.99	79.33	0.84
0.8	79.13	0.87	91.71	0.61	81.50	1.61	82.33	1.61	81.00	1.73
1.0	78.25	1.40	89.86	0.94	81.00	1.41	82.50	0.99	79.33	0.95
2.0	73.19	1.47	101.29	2.03	87.33	2.35	80.50	1.78	75.83	1.33
4.0	76.44	0.94	108.14	1.55	95.33	2.54	87.33	1.89	77.83	0.79
24-h exposure										
0.2	91.13	0.55	98.00	2.05	93.17	1.40	90.33	2.20	87.83	1.17
0.3	89.88	0.80	84.50	1.18	93.83	1.14	88.67	0.92	86.33	1.23
0.5	78.19	1.67	87.50	1.38	80.50	1.06	78.67	1.99	78.00	0.97
0.8	79.13	0.87	88.67	1.15	84.17	0.70	82.00	1.90	80.00	0.68
1.0	78.25	1.40	92.83	1.35	85.83	1.56	79.50	0.99	80.67	2.32
2.0	73.19	1.47	99.17	1.72	92.17	0.75	80.17	1.38	76.00	1.29
4.0	76.44	0.94	103.83	1.11	94.50	1.48	86.00	1.37	83.17	1.05

to 4.0 kHz. This yields less noise energy at the lower and higher frequencies of the hearing range and thus results in smaller threshold shifts.

In another approach, Popper and Clarke (1976) exposed goldfish to pure tones (0.3, 0.5, 0.8, and 1.0 kHz; SPL of 149 dB *re* 1 μ Pa) for 4 h and tested the auditory thresholds at 0.5 and 0.8 kHz behaviorally. They observed shifts ranging from 6 to 26 dB at 0.5 kHz and from 4 to 29 dB at 0.8 kHz. Threshold shifts were similar for 0.3 and 0.5 kHz and for 0.8 and 1.0 kHz, respectively.

A subsequent study using the bluegill sunfish *Lepomis macrochirus*, a hearing generalist of the family Centrarchidae, revealed no significant changes in auditory sensi-

tivity (Scholik and Yan, 2002b). This could be due to the low noise level in relation to the hearing threshold of the generalist, which was approximately 20 dB above HL. To assess the effects on hearing, noise levels of at least 60 dB above HL need to be applied.

Utilizing a different method, some authors observed elevation of auditory thresholds in the presence of another signal. Buerkle (1968) reported that varying background noise levels resulted in thresholds shifts up to 25 dB in the cod *Gadus morhua*. Hawkins and Chapman (1975) revealed that thresholds were elevated up to 14 dB in the presence of a broad-band noise (0.03–1.0 kHz; 20–30 dB above ambient sea noise level). The signal and second tone need not be

TABLE V. Differences of mean hearing thresholds (dB) of *P. pictus* at the seven frequencies tested calculated by one-way ANOVA for each frequency followed by pair wise multiple comparison procedure (Scheffé).

kHz	Differences								$F_{4,40}$	P
	Baseline—day 0		Baseline—day 3		Baseline—day 7		Baseline—day 14			
	Mean	SE	Mean	SE	Mean	SE	Mean	SE		
12-h exposure										
0.2	7.0 ^a	1.21	0.9	1.07	0.2	1.28	−1.5	1.13	10.40	<0.001
0.3	5.4	1.57	−0.7	1.45	3.3	1.45	−0.9	1.64	4.96	<0.005
0.5	11.1 ^a	2.66	0.8	3.05	2.5	2.83	1.1	2.82	6.31	<0.005
0.8	12.6 ^a	1.39	2.4	1.72	3.2	1.72	1.9	1.76	16.20	<0.001
1.0	11.6 ^a	2.22	2.8	2.46	4.3	2.39	1.1	2.38	10.02	<0.001
2.0	28.1 ^a	2.60	14.1 ^a	2.80	7.3	2.66	2.6	2.56	37.95	<0.001
4.0	31.7 ^a	1.74	18.9 ^a	2.15	10.9 ^a	1.91	1.4	1.62	83.13	<0.001
24-h exposure										
0.2	6.9	1.51	2.0	1.22	−0.8	1.58	−3.3	1.14	7.04	<0.001
0.3	4.6	1.49	4.0	1.48	−1.2	1.42	−3.5	1.50	8.47	<0.001
0.5	9.3	2.89	2.3	2.84	0.5	3.01	−0.2	2.83	4.10	<0.005
0.8	9.5 ^a	1.59	5.0	1.50	2.9	1.82	0.9	1.50	11.01	<0.001
1.0	14.6 ^a	2.45	7.6	2.49	1.3	2.39	2.4	2.69	11.83	<0.001
2.0	26.0 ^a	2.64	19.0 ^a	2.48	7.0	2.57	2.8	2.56	45.99	<0.001
4.0	27.4 ^a	1.68	18.1 ^a	1.78	9.6 ^a	1.75	6.7 ^a	1.67	83.97	<0.001

^aStatistically significant differences following adjustment to the number of frequencies tested ($P < 0.007$). Positive threshold differences mean a higher auditory threshold of the second group of a pairing at this particular frequency.

presented simultaneously to change hearing thresholds. Popper and Clarke (1979) investigated nonsimultaneous masking in the goldfish and revealed threshold shifts of up to 35 dB for pure tone pulses of different durations (15–50 ms), regardless of whether these were presented before or after a 250 ms white noise pulse. Fay (1974) noted that a 10-dB noise increment produced an approximately 10-dB masking increment, independent of frequency. In natural environments thresholds may often be elevated due to the presence of background noise.

A comparison of the effects of the two noise exposure durations applied in our study did not reveal any significant differences in either species. In contrast, Scholik and Yan (2001) did observe duration-dependent effects on threshold shifts in *P. promelas*. Threshold elevation was smaller after 1 h compared to 2–24 h of duration. They termed this asymptotic threshold shift (ATS), which is a well-known phenomenon in mammalian auditory studies. These results indicate that auditory sensitivity diminishes quickly (1–2 h) following noise exposure.

The course of recovery varied between species and seems to depend on exposure duration and amount of threshold shift. In *C. auratus* all frequencies returned to baseline values within 3 days and no effect of exposure duration on recovery was observed. In *P. pictus*, on the other hand, it took 14 days until thresholds returned to baseline values following 12 h noise exposure. After 24 h exposure, one (4.0 kHz) out of seven frequencies tested had not returned to baseline values even after 14 days. Scholik and Yan (2001) also revealed that longer exposure durations seem to influence recovery: at the highest frequencies the thresholds had not fully recovered to baseline values even 14 days after 24 h exposure, while at the same frequencies following 2-h exposure, recovery was observed within 6 days. This could be because the responses to higher frequencies tend to recover more slowly than those to lower frequencies or because at frequencies where large threshold changes occurred, such as in *P. pictus*, recovery of the responses takes more than 2 weeks.

The effects of noise exposure on the microanatomy and physiology of fish hair cells are less clear. Enger (1981) exposed the cod *G. morhua* to high-level sounds (50–400 Hz, 180 dB *re* 1 μ Pa) for 1–5 h and observed destroyed regions with missing sensory cilia on the saccular macula. Hastings *et al.* (1996) found only damage to the sensory hair cells of the oscar *Astronotus ocellatus* after stimulation with 300 Hz continuous tones at 180 dB (*re* 1 μ Pa) and allowed to survive for 4 days. Damage was limited to small regions of the utricle and lagena, but not the saccule. Fish are able to replace or regenerate hair cells in the ear after treatment with ototoxic drugs after several days (Lombarte *et al.*, 1993). The recovery of auditory sensitivity in our goldfish within three days indicates that threshold changes were most likely due to hair cell fatigue. The much longer recovery time in our catfish (14 days) might additionally be explained by microanatomical injuries and hair cell replacement. In all cases loss of sensitivity in our animals was temporary and parallels temporary thresholds shifts (TTS) in other animal taxa (Clark, 1991).

C. Behavior and acoustic communication

The observed threshold elevations following noise exposure may be crucial for the survival of the examined species because fish obtain information about predators and prey, competitors and potential mates by listening to their acoustic environment. If the hearing abilities of fishes were impaired by exposure to noise, the distances over which such information could be obtained would decrease, and the quality of the obtained information would deteriorate.

One of the target species, the pimelodid catfish *P. pictus*, is known to produce two types of sound: low-frequency drumming sounds and high-pitched stridulatory sounds, which are used during agonistic encounters. Comparison between the audiograms and sound spectra in *P. pictus* revealed that, before the exposure to intense noise, drumming sound energies were up to 20 dB above hearing thresholds and stridulatory sound energies were up to 31 dB above hearing thresholds [sound spectra after Ladich (1999)]. Immediately after noise exposure the differences between drumming sound energies and the audiogram were maximally 9 dB (12 h) and 11 dB (24 h), and between stridulatory sound energies and the audiogram maximally 11 dB (12 h) and 6 dB (24 h) (Fig. 6). Hearing impairment could alter a fish's possibility to

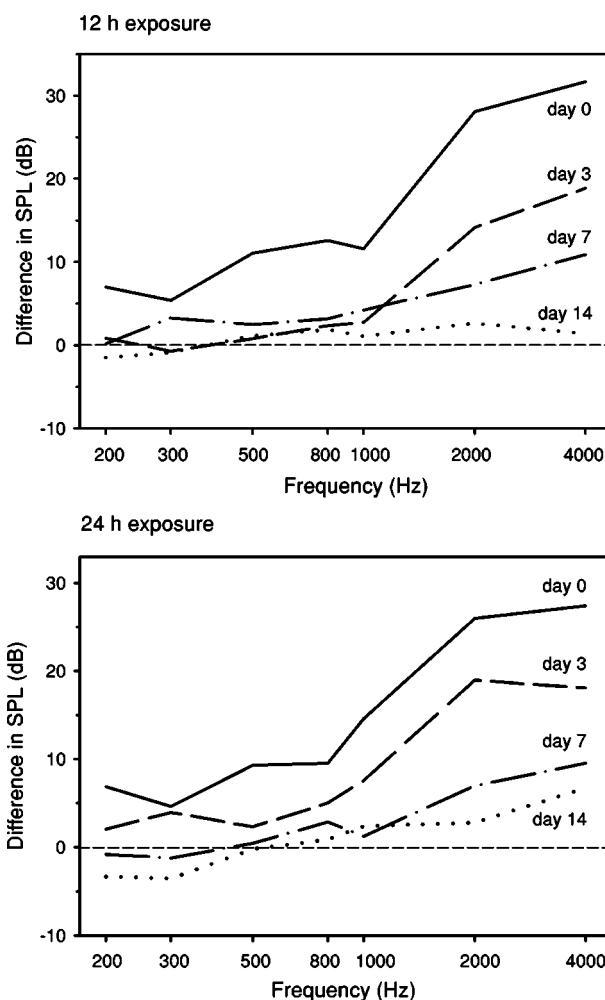


FIG. 5. Threshold shift in *P. pictus* immediately after noise exposure (day 0) and after 3, 7, and 14 days of recovery (day 3, day 7, day 14). The dashed line at zero indicates baseline values.

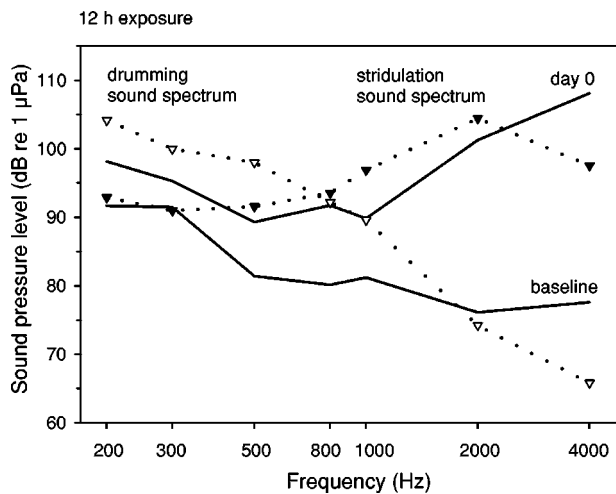


FIG. 6. Audiogram of *P. pictus* before and directly after 12-h noise exposure compared with spectra of conspecific drumming (open triangle) and stridulation (closed triangle) sounds. Audiogram: solid lines. Sound spectra: dotted lines. Sound spectra data after Ladich (1999).

assess the opponent's quality acoustically and therefore influence the outcome of a fight (Ladich *et al.*, 1992).

Despite the lack of data about effects on acoustic communication, existing data show behavioral changes of fishes exposed to intense sounds. Pearson *et al.* (1992) observed startle responses of rockfish *Sebastes* spp. during playback of air-gun-like sounds at 200–205 dB SPL (*re* 1 μ Pa), indicating that alarm as well as startle responses could be elicited by sound from actual survey operations. In contrast, Klimely and Beavers (1998) were unable to observe significant behavioral changes of rockfish exposed to ATOC-like signals, but this may be due to the relatively low SPL of 145 dB (*re* 1 μ Pa at 1 m). Schwarz and Greer (1984) observed behavioral responses of Pacific Herring *C. harengus pallasi* during playback of vessel noise and certain electronically synthesized sounds. Blaxter *et al.* (1981) observed startle responses (sharp bends to one side and subsequent escape) in shoaling herring *C. harengus* when the stimulus was above threshold.

More data exist for behavioral changes in several whale species due to anthropogenic noise. Lesage *et al.* (1999) monitored the vocal behavior of belugas *Delphinapterus leucas* before, during, and after exposure to two different vessel types and revealed reduced calling rates, brief increases in the emission of certain call types and a shift in frequency bands from a mean of 3.6 kHz to 5.2–8.8 kHz when vessels were approaching. Frankel and Clark (2000) observed subtle short-term behavioral changes in humpback whales *Megaptera novaeangliae* in the presence of ATOC transmission (received SPL 98–109 dB *re* 1 μ Pa).

V. CONCLUSIONS

Our experiments show that exposure to intense white noise affects the hearing abilities of two otophysan fishes. This is expressed by a decrease in amplitudes and an increase in latency of the auditory evoked responses and subsequently by an upward shift of hearing thresholds in both species.

Differences in hearing loss and recovery between both species, the goldfish *C. auratus* and the catfish *P. pictus*, were remarkable: the threshold shift was more pronounced in *P. pictus*, particularly at higher frequencies (32 vs 14 dB), and hearing thresholds recovered more slowly than in *C. auratus*. We suggest two possible explanations for this observation: (1) the auditory system of the catfish could be more susceptible to auditory fatigue and therefore may need more time to fully recover or (2) the higher auditory sensitivity in the catfish resulted in a greater threshold shift at higher frequencies which simply needed more time to return to baseline values in comparison to the goldfish.

The shift in auditory thresholds in the vocalizing catfish and a comparison with the spectra of conspecific sounds revealed that acoustic communication is severely impaired when fish live in noisy environments. Although it is assumed that hearing thresholds are always masked under natural conditions, high noise levels such as those applied here would reduce the effective communication distance to a few centimeters at most. Besides impacting intraspecific communication, elevated hearing thresholds would also limit the detection of predators or prey, thus reducing fitness.

ACKNOWLEDGMENTS

We would like to thank Lidia Eva Wysocki and Werner Timischl for statistical advice and Hannes Eder for providing a random numbers generating program. This research was supported by the Austrian Science Fund (FWF No. 12411 to FL).

- Andrew, R. K., Howe, B. M., Mercer, J. A., and Dzieciuch, M. A. (2002). "Ocean ambient sound: Comparing the 1960s with the 1990s for a receiver off the California coast," *ARLO* **3**, 65–70.
- Bauer, P., Körpert, P., Neuberger, M., Raber, A., and Schwetz, F. (1991). "Risk factors for hearing loss at different frequencies in a population of 47 388 noise-exposed workers," *J. Acoust. Soc. Am.* **90**, 3086–3098.
- Blaxter, J. H. S., and Hoss, D. E. (1981). "Startle response in herring: the effect of sound stimulus frequency, size of fish and selective interference with the acoustico-lateralis system," *J. Mar. Biol. Assoc. U.K.* **61**, 871–879.
- Blaxter, J. H. S., Gray, J. A. B., and Denton, E. J. (1981). "Sound and startle responses in herring shoals," *J. Mar. Biol. Assoc. U.K.* **61**, 851–869.
- Buerkle, U. (1968). "Relation of pure tone thresholds to background noise level in the Atlantic cod (*Gadus morhua*)," *J. Fish. Res. Board Can.* **25**, 1155–1160.
- Buerkle, U. (1969). "Auditory masking and the critical band in the Atlantic cod, *Gadus morhua*," *J. Fish. Res. Board Can.* **26**, 1113–1119.
- Clark, W. W. (1991). "Recent studies of temporary threshold shifts (TTS) and permanent threshold shift (PTS) in animals," *J. Acoust. Soc. Am.* **90**, 155–163.
- Dooling, R. J., Ryals, B. M., and Manabe, K. (1997). "Recovery of hearing and vocal behavior after hair-cell regeneration," *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14206–14210.
- Enger, P. S. (1981). "Frequency discrimination in teleosts—central or peripheral," in *Hearing and Sound Communication in Fishes*, edited by W. N. Tavolga, A. N. Popper, and R. R. Fay (Springer, New York), pp. 269–318.
- Fay, R. R. (1974). "Masking of tones by noise for the goldfish (*Carassius auratus*)," *J. Comp. Physiol. Psychol.* **87**, 708–716.
- Fay, R. R., Ahroon, W. A., and Orawski, A. A. (1978). "Auditory masking patterns in the goldfish (*Carassius auratus*): Psychophysical tuning curves," *J. Exp. Biol.* **74**, 83–100.
- Fay, R. R., and Popper, A. N. (2000). "Evolution of hearing in vertebrates: the inner ears and processing," *Hear. Res.* **149**, 1–10.

- Frankel, A. S., and Clark, C. W. (2000). "Behavioral responses of humpback whales (*Megaptera novaeangliae*) to full-scale ATOC signals," J. Acoust. Soc. Am. **108**, 1930–1937.
- Gorga, M. P., Kaminski, J. R., Beauchaine, J. R., and Jesteadt, W. (1988). "Auditory brainstem response to tone bursts in normally hearing subjects," J. Speech Hear. Res. **31**, 87–97.
- Griffiths, S. K., Pierson, L. L., Gerhardt, K. J., Abrams, K. J., and Peters, A. J. M. (1994). "Noise induced hearing loss in fetal sheep," Hear. Res. **74**, 221–230.
- Hall, J. W. (1992). *Handbook of Auditory Evoked Responses* (Allyn and Bacon, Boston).
- Hamernik, R. P., Henderson, D., and Salvi, R. (1982). *New Perspectives on Noise-induced Hearing Loss* (Raven, New York).
- Hastings, M. C., Popper, A. N., Finneran, J. J., and Lanford, P. J. (1996). "Effects of low-frequency underwater sound on hair cells of the inner ear and lateral line of the teleost fish *Astronotus ocellatus*," J. Acoust. Soc. Am. **99**, 1759–1766.
- Hawkins, A. D., and Chapman, C. J. (1975). "Masked auditory thresholds in the cod, *Gadus morhua* L.," J. Comp. Physiol. **103**, 209–226.
- Hawkins, A. D., and Myrberg, A. A. (1983). "Hearing and sound communication underwater," in *Bioacoustics, a Comparative Approach*, edited by B. Lewis (Academic, London), pp. 347–405.
- Kenyon, T. N., Ladich, F., and Yan, H. Y. (1998). "A comparative study of hearing ability in fishes: the auditory brainstem response approach," J. Comp. Physiol., A **182**, 307–318.
- Kileny, P., and Shea, S. L. (1986). "Middle-latency and 40-Hz auditory evoked responses in normal hearing subjects: clicks and 500-Hz thresholds," J. Speech Hear. Res. **29**, 20–28.
- Klimely, A. P., and Beavers, S. C. (1998). "Playback of acoustic thermometry of ocean climate (ATOC)-like signal to bony fishes to evaluate phonotaxis," J. Acoust. Soc. Am. **104**, 2506–2510.
- Ladich, F. (1999). "Did auditory sensitivity and vocalization evolve independently in otophysan fishes?" Brain Behav. Evol. **53**, 288–304.
- Ladich, F. (2001). "Acoustic communication and the evolution of hearing in fishes," Philos. Trans. R. Soc. Lond. B Biol. Sci. **355**, 1285–1288.
- Ladich, F., and Bass, A. H. (2003). "Audition," in *Catfishes*, edited by B. G. Kapoor, G. Arratia, M. Chardon, and R. Diogo (Oxford & IBH, New Delhi), pp. 701–730.
- Ladich, F., Brittinger, W., and Kratochvil, H. (1992). "Significance of agonistic vocalization in the croaking gourami (*Trichopsis vittatus*, Teleostei)," Ethology **90**, 307–314.
- Lesage, C., Barette, C., Kingsley, M. C. S., and Sjare, B. (1999). "The effect of vessel noise on the vocal behavior of belugas in the St. Lawrence River estuary, Canada," Mar. Mammal Sci. **15**, 65–84.
- Lombarte, A., Yan, H. Y., Popper, A. N., Chang, J. S., and Platt, C. (1993). "Damage and regeneration of hair cell ciliary bundles in a fish ear following treatment with gentamicin," Hear. Res. **64**, 166–174.
- Miller, P. J. O., Biassoni, N., Samuels, A., and Tyack, P. L. (2000). "Whale songs lengthen in response to sonar," Nature (London) **405**, 903.
- Pearson, W. H., Skalski, J. R., and Malme, C. I. (1992). "Effects of sounds from a geophysical survey device on behavior of captive rockfish (*Sebastes* spp.)," Can. J. Fish. Aquat. Sci. **49**, 1343–1356.
- Popper, A. N., and Clarke, N. L. (1976). "The auditory system of the goldfish (*Carassius auratus*): effects of intense acoustic stimulation," Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol. **53**, 11–18.
- Popper, A. N., and Clarke, N. L. (1979). "Non-simultaneous auditory masking in the goldfish, *Carassius auratus*," J. Exp. Biol. **83**, 145–158.
- Popper, A. N., and Fay, R. R. (1993). "Sound detection and processing by fish: critical review and major research questions," Brain Behav. Evol. **41**, 14–38.
- Richardson, W. J., Greene, C. R., Malme, C. J., and Thomson, D. H. (1995). *Marine Mammals and Noise* (Academic, San Diego).
- Richardson, W. J., Würsig, B., and Greene, C. R. (1986). "Reactions of bowhead whales, *Balaena mysticetus*, to seismic exploration in the Canadian Beaufort Sea," J. Acoust. Soc. Am. **79**, 1117–1128.
- Saunders, J., and Dooling, R. J. (1975). "Noise-induced threshold shift in the parakeet (*Melopsittacus undulatus*)," Proc. Natl. Acad. Sci. U.S.A. **71**, 1961–1965.
- Scholik, A. R., and Yan, H. Y. (2001). "Effects of underwater noise on auditory sensitivity of a cyprinid fish," Hear. Res. **152**, 17–24.
- Scholik, A. R., and Yan, H. Y. (2002a). "Effects of boat engine noise on the auditory sensitivity of the fathead minnow, *Pimephales promelas*," Environ. Biol. Fish. **63**, 203–209.
- Scholik, A. R., and Yan, H. Y. (2002b). "The effects of noise on the auditory sensitivity of the bluegill sunfish, *Lepomis macrochirus*," Comp. Biochem. Physiol. A **133**, 43–52.
- Schwarz, A. L., and Greer, G. L. (1984). "Responses of Pacific herring, *Clupea harengus pallasi*, to some underwater sounds," Can. J. Fish. Aquat. Sci. **41**, 1183–1192.
- Silman, S., and Silverman, C. A. (1991). *Auditory Diagnosis: Principles and Applications* (Academic, San Diego).
- Wysocki, L. E., and Ladich, F. (2001). "The ontogenetic development of auditory sensitivity, vocalization and acoustic communication in the labyrinth fish *Trichopsis vittata*," J. Comp. Physiol., A **187**, 177–187.
- Yan, H. Y. (1998). "Auditory role of the suprabranchial chamber in gourami fish," J. Comp. Physiol., A **183**, 325–333.

Erratum: “Benchmark solutions of plane wave bottom reflection loss” [J. Acoust. Soc. Am. 104, 3305–3312 (1998)]

Michael A. Ainslie^{a)} and Alvin J. Robins^{b)}

BAeSEMA Ltd., Apex Tower, 7 High Street, New Malden, Surrey KT3 4LH, England

(Received 14 November 2002; revised 10 January 2003; accepted 14 January 2003)

[DOI: 10.1121/1.1558355]

PACS numbers: 43.30.Ma, 43.30.Dr, 43.10.Vx

Equation (A12) is missing a factor involving the square root of the ratio of the densities at the top and bottom of the sediment layer. The equation should read:

$$p_2^\pm(z) = \left[\frac{\rho_2(z)}{\rho_2(0)} \right]^{1/2} \frac{\text{Ai}[-\chi(z)] \pm i \text{Bi}[-\chi(z)]}{\text{Ai}[-\chi(0)] \pm i \text{Bi}[-\chi(0)]}. \quad (\text{A12})$$

^{a)}Now at: TNO Physics and Electronics Laboratory, Underwater Acoustics Group, Oude Waalsdorperweg 63, 2509 JG, The Hague, The Netherlands.

^{b)}Now at: CORDA Limited, Chester House, Farnborough Aerospace Centre, Farnborough, Hampshire GU14 6YU, England.

In Sec. IA the sub-heading “1. Tabulated results” is incorrect and should be deleted.

The title and scope of Ref. 10 have changed. As a result the sentence beginning “The acoustic properties of...” on line 2 of p. 3306 should now read “The acoustic properties of the BLUG profile are described in Refs. 1, 2, and 10.” An updated reference for Ref. 10 is:

¹⁰A. J. Robins and M. A. Ainslie, “Sediment wave function calculations for two generic sound speed profiles using the Langer approximation” (in preparation).

PROGRAM OF

The 145th Meeting of the Acoustical Society of America

Nashville, Tennessee • Nashville Convention Center • 28 April–2 May 2003

1a MON. AM

NOTE: All Journal articles and Letters to the Editor are peer reviewed before publication. Program abstracts, however, are not reviewed before publication, since we are prohibited by time and schedule.

MONDAY MORNING, 28 APRIL 2003

ROOMS 103/104, 8:00 A.M. TO 12:00 NOON

Session 1aUW

Underwater Acoustics, Acoustical Oceanography and Physical Acoustics: Frederick Tappert Memorial Session on Propagation Phenomena and the Parabolic Equation

David R. Palmer, Chair

*Atlantic Oceanographic and Meteorological Laboratory, 4301 Rickenbacker Causeway,
Miami, Florida 33149*

Chair's Introduction—8:00

Invited Papers

8:05

1aUW1. The summer of 1974. Stanley M. Flatté (Phys. Dept., Univ. of California at Santa Cruz, Santa Cruz, CA 95064, flatté@physics.ucsc.edu)

In the early 1970s, Garrett and Munk summarized oceanographers' knowledge of ocean internal waves in a reference statistical spectrum. Almost at the same time, Hardin and Tappert introduced the parabolic equation into ocean acoustics. In the summer of 1974, Flatté and Tappert [*J. Acoust. Soc. Am.* **58**, 1151–1159 (1975)] helped bring these two ideas together by simulating 100-Hz acoustic propagation through an approximate version of the Garrett–Munk spectrum to a range of 100 km, in order to help understand the effects caused by internal waves on acoustic fluctuations. The history of that summer will be discussed, along with a movie made at that time. The influence of that work on subsequent developments will be discussed.

8:30

1aUW2. Waves, rays, and the predictability of underwater sound fields. Michael G. Brown (RSMAS-AMP, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

Recent results relating to the application of ray-based methods to long-range underwater sound propagation are discussed, including Fred Tappert's influence on the development of some of the ideas. Understanding limitations on predictability linked to chaotic motion of ray trajectories is a central issue. Related issues that are discussed include semiclassical breakdown and the important role played by the background sound-speed structure in controlling ray and wavefield stability. [Work supported by ONR.]

8:55

1aUW3. Computing rough surface Doppler effects on broadband pulse propagation using a split-step Fourier parabolic equation model. Kevin B. Smith (Code PH/Sk, Dept. of Phys., Naval Postgraduate School, Monterey, CA 93943, kbsmith@nps.navy.mil) and Richard Ead (Autonomous Undersea Vehicles Div. Naval Undersea Warfare Ctr., Newport, RI 02841)

The current interest in broadband pulse propagation in shallow water is increasing with the need for improved active sonar systems and the growth of applications utilizing underwater acoustic communications. Such shallow-water propagation is dominated by boundary interactions. If the ocean surface is rough on the scale of the acoustic wavelength, considerable scattering can occur that can significantly influence the coherent propagation. Because the rough ocean surface is also evolving dynamically, such scattering can introduce Doppler shifting and spreading of the acoustic pulse spectrum. Following the method of F. D. Tappert and L. Nghiem-Phu [*J. Acoust. Soc. Am. Suppl.* **1** **77**, S101 (1985)], an exact formulation for scattering from a rough surface is introduced into the

Monterey-Miami Parabolic Equation Model [JCA 9, 243–285 (2001)]. An algorithm is then developed for computing the solution of pulse propagation as the rough surface evolves dynamically. The result can be shown to produce the proper spatial scattering as well as introduce the correct Doppler shifts from the dynamic surface. A simple sinusoidal rough surface interface will be employed to test the algorithm followed by a more complex, realistic rough ocean interface. Spatial and temporal properties of the scattered field will be examined and discussed.

9:20

1aUW4. Can chaotic rays be used for long-range propagation? Michael A. Wolfson and Frank S. Henyey (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105-6698)

Fred Tappert was skeptical of the applicability of chaotic rays to long-range propagation experiments. In this talk, the best ray predictions we can make are compared to PE predictions. In particular, one expects the predictions of “deep arrivals” to be the most robust to uncertainties in how to use ray information, and the ray predictions fail to agree with the PE distribution of the intensity of the deep arrivals. The discrepancy can be seen in the propagation over the first 50 km from an assumed single mode source array. A comparison with a toy model shows that the difficulty with the rays is that the scattering is too weak, in the sense that almost all of the phase fluctuations in a correlation length are smaller than a cycle.

Contributed Papers

9:45

1aUW5. Chaos and wave propagation regimes. John Colosi (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Ray chaos theory and parabolic equation numerical modeling were two thrusts of Fred Tappert’s research that were perpetually in tension. Fred was interested in the problem of identifying wave propagation regimes, most notably the strong focusing caustic regime and its evolution into the saturation regime. On the one hand, chaos theory held the seed of the complexity Fred believed existed in ocean acoustic wavefields; on the other hand ocean acoustic ray chaos theory (which Fred helped to pioneer) was a disdainful approximation to the full wave treatments offered by parabolic equation calculations. Fred was convinced that the saturation limit could not be obtained using ray theory and therefore he examined a new field of inquiry: a blend of chaotic ray insight and full wave dynamics called wave chaos. This talk will discuss some of Fred’s insights on this topic and how they relate to observations from basin scale acoustic transmissions.

10:00–10:15 Break

10:15

1aUW6. Dozier–Tappert theory for long-range propagation. Frank S. Henyey and Michael A. Wolfson (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105-6698)

Among Fred Tappert’s contributions to ocean acoustics is the mode transport equation in the paper he wrote with Dozier. In light of the results reported by Wolfson (this session), we need to have a theoretical approach other than chaotic rays for long-range propagation experiments; mode transport is a promising candidate. The principal assumptions made by Dozier and Tappert are the Markov approximation and the assumption that cross-mode coherence is small. Although eminently reasonable at the time they wrote their paper, the latter approximation must be questioned due to newer data and simulations. Deep arrivals constitute a phenomenon directly predictable from mode transport equations that extend that of Dozier and Tappert by including travel time information.

10:30

1aUW7. Low-frequency phase rate, source–receiver motion, and the parabolic approximation. George V. Frisk (Appl. Ocean Phys. & Eng. Dept., M.S. #11, Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, gfrisk@whoi.edu) and Travis L. Poole (MIT/WHOI Joint Prog. in Oceanogr. and Oceanogr. Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

A theory relating the phase rate of low-frequency cw signals to the range rate between the source and receiver is presented. A key component of this theory is the parabolic approximation, which was introduced into underwater acoustics by Tappert. It is shown that, after shifting the re-

ceived signal to the base band, the leading-order behavior of the residual phase rate is simply equal to the product of a typical wavenumber in the water column and the source–receiver range rate. This result holds true even for situations where the acoustic field magnitude clearly displays a complicated multimodal (or multipath) interference pattern. The theory is supported by a variety of experimental measurements obtained in a broad range of ocean environments. These include deep- and shallow-water situations, range-independent and range-dependent cases, and short- and long-range scenarios. Implications of the theory for phase tracking and localization of quasi-cw sources are also discussed. [Work supported by ONR.]

10:45

1aUW8. Statistics of normal-mode amplitudes in shallow water. Daniel Rouseff (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Dozier and Tappert published a statistical theory of acoustic propagation in a random ocean [J. Acoust. Soc. Am. **63**, 353–365 (1978)]. Their theory predicts the statistical moments of the acoustic mode amplitudes. The acoustic modes are coupled by random variability in the water column. Creamer extended the theory to include bottom loss, an important factor in shallow-water applications [J. Acoust. Soc. Am. **99**, 2825–2838 (1996)]. In the present paper, the extended theory is applied to propagation through a realistic shallow-water internal wave model. An efficient computational algorithm is developed that makes practical the theory’s application to observed sound speed and buoyancy profiles. Model parameters such as bottom loss and internal-wave strength can be varied with minimal computational cost. Numerical results are presented that demonstrate the competing effects of modal attenuation and mode coupling. [Work supported by ONR.]

11:00

1aUW9. Coupled-mode solutions in the complex plane and benchmarking with the parabolic equation method. Steven A. Stotts, David P. Knobles, and Robert A. Koch (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, stotts@arlut.utexas.edu)

A complex plane extension of a previously developed two-way coupled mode model is presented. Coupling coefficients based on horizontal layer propagators using complex Airy solutions are evaluated analytically [Stotts, J. Acoust. Soc. Am. **111**, 1623–1643 (2002)]. This approach significantly reduces memory storage requirements and improves the computational efficiency over numerical depth integration for the coefficients. The integral equation method introduced by Knobles [J. Acoust. Soc. Am. **96**, 1741–1747 (1994)] is used to solve the coupled equations. Benchmarking with the parabolic equation (PE) approach will be presented for several examples including the ASA continental shelf benchmark and up-slope/down-slope wedge propagation. Since Tappert’s introduction, PE has been accepted as the de-facto model for range variable calculations. Nevertheless, having an independent verification by a normal mode model for

general environments is advantageous. Differences between real and complex k -plane solutions are emphasized. Comparisons to numerically calculated key quantities are presented. Application to a range-dependent data set recorded recently off the coast of Florida is discussed.

11:15

1aUW10. Vertical modes with sediment absorption in generalized Pekeris waveguides. Frank S. Henyey, Dajun Tang, and Stephen A. Reynolds (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

The eigenvalue spectrum for vertical modes in a system consisting of a water layer overlying sediments modeled as an infinitely deep fluid region with absorption is investigated. These eigenvalues are the square of the horizontal (complex) wavenumbers. The sound speed in the water is assumed smaller than that in the sediments. The dependence of modes on the absorption coefficient as a parameter is of primary concern. The nature of the eigenvalues is related to the amount of absorption in sediments. The main result is that as the amount of absorption is raised, leaky modes (which are not eigenvalues in the absence of absorption) become true discrete modes. This behavior, not present in perturbation theory, is physically explained. [Work supported by ONR.]

11:30

1aUW11. Coupled hydrodynamic ship wake and PE-based acoustic propagation model. Aditya Balani, R. Lee Culver, David L. Bradley, Eric G. Paterson (Appl. Res. Lab., Penn State Univ., P.O. Box 30, N. Atherton St., State College, PA 16804-0030, agb127@psu.edu), Xiao Di, and Robert F. Kunz (Penn State Univ., State College, PA 16804-0030)

A full-wave, one-way, 2D parabolic equation (PE) code is used to predict sound propagation through a complex hydrodynamic bubble field. The normalized velocity component obtained from a hydrodynamic ship wake model is the basis for constructing the bubble distribution. A three-dimensional field is used to represent the complex environment: down wake (toward or away from the ship), cross wake (perpendicular to the wake axis), and with depth. Although the hydrodynamic wake model gen-

erates three-dimensional wake-bubble fields, the PE program is a 2D model. The problem is assumed to be two-dimensional by taking various two-dimensional cuts of the wake field, which are then used in the PE simulation. The wake of a ship is a complicated and challenging environment in which to model acoustic propagation. Predictions are made for the representation of the wake at various distances from the source ship. From this study a few results are seen: (1) The sound level corresponding to the strength of a spherically spreading sound field can be obtained. (2) Sound propagation through the wake bubble field is a three-dimensional problem. (3) Small changes in void fraction can produce noticeable differences in the propagation. [Work supported by ONR under Award No. N00014-02-1-0156.]

11:45

1aUW12. Propagation modeling issues for macro/micro scale environments. John McCoy (The Catholic Univ. of America, Washington, DC)

The propagations of acoustic signals through a porous solid and through a bubble cloud are both representative of a class of experiments for environments that are observable on macro and micro length scales. Intuitive models for these propagations are described by theories for poroelastic solids and for frothy fluids, respectively. These theories reflect a different intuitive understanding of the underlying physics that applies to the micro scale, and accommodate different phenomenology that can be observed on the macro scale. A more precise understanding of the linkage between micro and macro scale observations is a fundamental problem that has attracted a number of different approaches that are usually left unrelated. Moreover, there are other phenomenology that, while observable on a macro scale, are not modeled by either theory, and other models that are developed for these other phenomenology. An example of another phenomenology is a continuous backscattering occasioned by micro scale heterogeneity, which results in a kind of localization in one dimension. An example of another model is a radiative transport model. These issues are discussed in the light of still another approach to accommodating macro/micro scale variation, this one based on a phase space, spatial filtering. [Work supported by ONR.]

MONDAY AFTERNOON, 28 APRIL 2003

ROOMS 108/109, 2:00 TO 4:45 P.M.

Session 1pAA

Architectural Acoustics: Architectural Acoustics—Research and Application

Martha M. Larson, Cochair

Kirkegaard Associates, 801 West Adams Street, 8th Floor, Chicago, Illinois 60607-3034

Lily M. Wang, Cochair

Architectural Engineering, University of Nebraska—Lincoln, 200B Peter Kiewit Institute, 1110 South 67th Street, Omaha, Nebraska 68182-0681

Contributed Papers

2:00

1pAA1. A modal approach to the prediction of the sound reduction index. Alain Tisseyre, Cécile Courné, Thomas Buzzy, and André Moulinier (Tisseyre & Associés, 16, chemin de Manel, F 31400 Toulouse, France)

The calculation of the sound reduction index in modal analysis is presented in a general way; different possible approaches are described. These calculations are done in two steps: a vibratory study to determine the transverse displacement of the plate and a study of radiation. The specificity of orthotropic plates is presented. This study led to programming a calculation algorithm. Initial hypotheses are indicated, as well as

results obtained for various plates or partitions. Modal analysis calculation results are then compared to the Cremer–Sewell approach results.

2:15

1pAA2. Experimental determination of porous media properties for predicting absorption of bulk-reacting surfaces. Jerry W. Rouse, Linda P. Franzoni, and Justin W. Jaworski (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708, jwrouse@me1.egr.duke.edu)

Porous materials are commonly used and located along enclosure boundaries for the purpose of sound attenuation. An accurate representation of the boundary condition is needed for the analytical modeling of

these enclosures. The proper characterization of porous materials requires the determination of the effective complex phase speed and effective complex density of the fluid during propagation through the porous material. This work presents a two-step process in which these characteristic properties are determined using a standing wave tube device. The normal incidence impedance of a porous material is obtained for two cases: a zero velocity condition, and a zero pressure condition, behind the material sample. The effective phase speed and density are then obtained from formulas containing the normal incidence impedances. For verification, standing waves are predicted theoretically using the previously obtained parameters as inputs for the case of a two-dimensional slender duct with the porous material lining a side wall. These predictions are compared to experimental measurements of sound pressure in the two-dimensional duct and good agreement is observed for the open-cell foams considered in this work.

2:30

1pAA3. Dependence of reverberant energy on source distance in enclosures. Erik Larsen and Albert S. Feng (Beckman Inst. for Adv. Sci. and Technol., Univ. of Illinois at Urbana-Champaign, 405 N. Mathews Ave., Urbana, IL 61801, elarsen@uiuc.edu)

It is known that the direct-to-reverberant energy ratio has a strong influence on speech intelligibility, and also on the performance of adaptive beamforming algorithms. It is also a cue for source distance perception. According to traditional acoustic assumptions, reverberation is homogeneous and isotropic, yielding a decay of -6 of direct-to-reverberant energy ratio per doubling of source distance. However, the extent to which real-world data deviates from this ideal prediction and the underlying physical parameters that affect the outcome are unknown. A series of experiments and simulations were conducted in which the reverberant energy was determined at various distances from the sound source, in various listening rooms. It was found that reverberant energy decreased systematically with an increase in source distance, while reverberation time increased orderly. Also, source distance at which direct-to-reverberant energy ratio was zero corresponded well with the critical distance as given by standard acoustic theory. The decay of reverberant energy with distance varied for the different rooms tested, roughly between -1 and -3 dB per doubling of source distance. An evaluation of these results in terms of the room's acoustic parameters will be given.

2:45

1pAA4. 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 Dept. of Architecture, 201 Cowgill Hall 0205, Blacksburg, VA 24061, mermann@vt.edu), and Byron W. Harrison (The Talaske Group, Oak Park, IL)

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.

3:00

1pAA5. Relating double slope decay in coupled volumes with volume ratio, absorption ratio, and aperture size. David T. Bradley and Lily M. Wang (Architectural Eng., Univ. of Nebraska-Lincoln, 1110 S. 67th St., Omaha, NE 68182-0681, dbradley@mail.unomaha.edu)

Coupled volumes typically consist of two spaces connected through an opening known as a coupling aperture. If the absorption characteristics in the auxiliary space are such that its decay time is longer than that of the main space, late-arriving energy will be fed back into the main space by the secondary volume. This energy coupling leads to the phenomenon known as the double slope decay. This project uses computer modeling to study the effect of three architectural variables on the double slope profile from a simple coupled volume system. The three variables studied are the volume ratio between the main and secondary space, the absorption ratio between the two spaces, and the aperture size. The resulting energy decay profiles are categorized according to their degree of double slope behavior and cross-referenced with the independent variables. Quantitative results from this analysis are presented.

3:15

1pAA6. A comparison of background noise levels and reverberation times measured in unoccupied elementary classrooms. Richard D. Godfrey (Owens Corning, Sci. and Technol., 2790 Granville Rd., Granville, OH 43023)

The key performance criteria listed in ANSI S12.60-2002, Acoustical Performance Criteria, Design Requirements and Guidelines for Schools, are that the maximum background noise is limited 35 dBA, and that the maximum reverberation time is limited to 0.6 seconds in the most common classroom size. Limits on sound transmission properties of the room envelope are also made. If these performance criteria are met, each student, no matter where he or she sits in the classroom, will be in an environment that affords adequate speech intelligibility, i.e., an adequate opportunity to understand the teacher's words correctly. During the standard development process, the author had the opportunity to work with his colleges at The Ohio State University. The OSU group made measures in 34 classrooms in the Columbus area (inter-city, suburbs, and rural; public and private). The author analyzed these data, and compared the results with ANSI S12.60-2002 performance criteria. It was found that the background noise levels were around 50 dBA. This is 15 dBA above the standard requirement. A majority of the reverberation times were in line with the 0.6-second criteria, but a significant number had reverberation times around 1 second.

3:30-3:45 Break

3:45

1pAA7. The effects of stage configurations in a recital hall. Wonyoung Yang (School of Occupational and Environ. Hygiene, Univ. of British Columbia, 2206 East Mall, Vancouver, BC V6T 1Z3, Canada), Kathleen K. Hodgdon, and Jiry Tichy (Penn State Univ., State College, PA 16804)

The room acoustical parameters, the reverberation time (RT), the early decay time (EDT), the bass ratio (BR), the clarity factor (C80), the initial time delay gap (ITDG), and the interaural cross-correlation coefficient (IACC) were measured for acoustical evaluation both in the audience seats and on the stage, using six different stage configurations in a 450-seat unoccupied hall. The stage configurations consisted of combinations of varying the position of the drapes with and without the reflector at the back of the stage. The listening test was conducted in a laboratory using high fidelity headphones to verify the subjective preference for the seats and the presence of the reflector. Analysis showed that the draped area of the stage was related to the RT, EDT, C80, BR, while the ITDG was strongly dependent on the presence of the reflector. There was not a strong impact observed in the IACC for the varying configurations. In general,

the reflectors improved most of the acoustical parameters. The loudness of the musical sound was the primary parameter used to decide the subjective preference, which correlated with the C80 and the ITDG with the reflector.

4:00

1pAA8. Comparison of echo criteria for a large fan-shaped auditorium. Lance L. Locey, Timothy W. Leishman, Scott D. Sommerfeldt, and Brian E. Anderson (Dept. of Phys. and Astron., Brigham Young Univ., N281 Eyring Sci. Ctr., Provo, UT 84602, LLL6@email.byu.edu)

Complaints of perceived echoes at specific locations in a 21,000 seat fan-shaped auditorium have prompted the measurement and analysis of numerous impulse responses. The responses were first processed using the echo criterion of Niese, then using the criterion of Dietsch and Kraak. This paper compares the two criteria and explores their abilities to assess whether peaks and anomalies of the measured responses were likely to produce audible echoes in the hall. The Niese criterion was found to better predict the perception of echoes produced by broad irregular decay trends. The Dietsch and Kraak criterion was able to better predict echoes produced by sufficiently strong specular reflections. Neither criterion alone was able to fully characterize these perceptions. Results obtained for both criteria at various seat locations will be presented and compared with subjective findings.

4:15

1pAA9. Measurement of the acoustic characteristics of the concert hall at the Sydney Opera House. John Bassett and Densil Cabrera (School of Architecture, Design Sci. and Planning, The Univ. of Sydney, Sydney NSW 2006, Australia)

The Sydney Opera House Trust is considering making changes to its Concert Hall. Prior to any alteration, the Acoustics group at the University of Sydney has sought to document the hall. The main measurements were

made for 48 receiver locations and 6 source locations, using omnidirectional measurement microphones, B-format (Soundfield) microphones and dummy head microphones in every receiver position. The measurements included impulse responses, anechoic music recordings, and recordings of a calibrated sound power source. Results documented by previous practitioners and researchers are described and comparison is made with the recent results.

4:30

1pAA10. A full simulation of the Quetzal echo at the Mayan pyramid of Kukulkan at Chichen Itza in Mexico. Nico F. Declercq, Joris Degrieck (Soete Lab., Dept. of Mech. Constr. & Prod., Ghent Univ., St Pietersnieuwstr 41, 9000 Gent, Belgium), Rudy Briers (KATHO, Sint Jozefstraat 1, 8820 Torhout, Belgium), and Oswald Leroy (IRC-KULAK Univ., E. Sabbelaan 53, B-8500 Kortrijk, Belgium)

It is well known that a handclap in front of the staircase of the pyramid produces an echo that sounds similar to the chirp of the Quetzal bird. This phenomenon occurs due to diffraction. There exist some publications concerning this phenomenon and even some first attempts are reported to simulate it. However, no full simulation (amplitude, frequency, time) has ever been reported before. The present work presents a simulation which is based on the theory of the diffraction of plane waves and which takes into account continuity conditions. The latter theory is the building block for an extended theory that tackles the diffraction of a spherical sound pulse. By means of these principles it is possible to entirely simulate the echo following a handclap in front of the staircase. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.)]

MONDAY AFTERNOON, 28 APRIL 2003

ROOMS 103/104, 1:00 TO 5:20 P.M.

Session 1pAO

Acoustical Oceanography, Underwater Acoustics and Signal Processing in Acoustics: Geoacoustic Inversion I

Peter Gerstoft, Chair

Scripps Institution of Oceanography, Marine Physical Laboratory, University of California, San Diego, 9500 Gillman Drive, La Jolla, California 92093-0238

Chair's Introduction—1:00

Invited Papers

1:05

1pAO1. Probabilistic geoacoustic inversion. Stan Dosso (SACLANT Undersea Res. Ctr., Viale S. Bartolomeo, 400 19138 La Spezia, Italy and Univ. of Victoria, Victoria, BC, Canada, sdosso@uvic.ca)

The problem of estimating seabed geoacoustic parameters from ocean acoustic measurements has received considerable attention in recent years. Geoacoustic inversion represents a convenient alternative to direct measurements (e.g., coring) and provides sensitivity relevant to acoustic source localization applications; however, it requires solving a strongly nonlinear inverse problem. A variety of approaches have been developed (by a number of researchers) based on seeking geoacoustic parameters that provide the optimal match to measured acoustic fields using global search techniques. Other approaches include inversion of bottom-loss or seabed-reflectivity data and ambient noise. Topics of current interest include range-dependent inversion, coherent spatial/temporal processing, and uncertainty estimation. This paper reviews the above approaches in terms of a general probabilistic formulation for geoacoustic inversion. The goals of the probabilistic approach are to fit the acoustic data and available prior information to within their uncer-

tainties, and to estimate geoacoustic parameters, their uncertainties, and inter-relationships. This is accomplished using a Bayesian formulation and Markov chain Monte Carlo approach (Gibbs sampling) to extract features of the posterior probability density such as the maximum *a posteriori* estimate, marginal probability distributions, and correlations. The approach is illustrated for matched-field inversion, inversion of seabed reflectivity, and source localization with environmental uncertainty.

2:05

1pAO2. Using measured acoustic data to obtain estimates of seabed physical parameters. D. P. Knobles (Appl. Res. Labs., Univ. of Texas, P.O. Box 80, Austin, TX 78713-8029, knobles@arlut.utexas.edu)

Inversion for seabed parameters is part of a broader analysis that seeks to characterize the nature of and to quantify the predictability of sound propagation in complex shallow seas. This talk examines results obtained from numerous inversions using real data and attempts to place inversion techniques in perspective as a tool for understanding the physics of the interaction of sound with various types of seabeds. Data sets include those obtained in the Gulf of Mexico, off the eastern coast of the United States, and off the Korean Peninsula. Acoustic data include multifrequency measurements of transmission loss versus range and time series generated by both impulsive sources and surface ships received on both HLA and VLA. The issue of uniqueness is addressed by a combination of methods that include the comparison of inverted results to geophysical measurements, cost function sensitivity, and a comparison of modeled results to independent data not used in the inversion and localization. Of particular interest is the attenuation structure in both hard and soft sediments and the acoustic data required to extract this information. [Work supported by ONR.]

2:25

1pAO3. Measures of uncertainty in geoacoustic inversion. N. Ross Chapman (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada, chapman@uvic.ca)

Inversion methods for estimating geoacoustic model parameters from acoustic field data have been applied with considerable success over the past decade. The most effective methods that have been developed generally fall into two categories, nonlinear methods that are posed as optimization problems, and linearized methods that invert quantities such as horizontal wave numbers that are derived from the acoustic field data. For either case, the complete solution of the geoacoustic inverse problem requires a measure of the error for the estimated parameter, as well as the estimate itself. This paper describes an approach for specifying an error measure in nonlinear inversion processes based on matched field processing. The inversion uses an optimization algorithm that combines global and local search processes to sample the model parameter space. An effective error measure is obtained from information in scatter plots of the cost function versus parameter values for models that were tested in the search process. Examples are presented to demonstrate the application of the method to test cases from the recent Geoacoustic Inversion Benchmark Workshop, and from experimental data from vertical line arrays and seafloor horizontal arrays. [Work supported by ONR.]

2:45

1pAO4. Geoacoustic inversion of data collected in two “typical” shallow-water environments. Peter Nielsen (SACLANT Undersea Res. Ctr., Viale San Bartolomeo 400, 19138 La Spezia, Italy) and Martin Siderius (SAIC, San Diego, CA 92121)

Two distinct types of shallow-water experiments were recently conducted by SACLANTCEN in different shallow-water areas (<150 m) to perform geoacoustic inversion of received acoustic signals. A towed sound source–horizontal line array was used in one of the experiments, and in the other the source and vertical array were moored on the bottom. The two experimental configurations were deployed in two areas with soft and hard bottom properties in order to investigate (1) the performance of geoacoustic inversion using horizontal and vertical line arrays, and (2) the robustness of the inversion method to different bottom characteristics. The transmitted signals were 1-s linear frequency modulated sweeps in the band from 200–800 Hz. The source–horizontal array separation was 300 m and the fixed path between the source and the vertical array was 2 km. This short source–receiver separation minimizes the impact of the range-dependent water-column and seabed properties on the acoustic propagation. The applied geoacoustic inversion technique of the data is the traditional model-based matched-field processing utilizing a global search algorithm. Results from successful geoacoustic inversion in the regions with soft-bottom properties are presented for the two experimental configurations. Severe difficulties appeared when inverting data from the hard-bottom regions and possible causes are given.

3:05–3:20 Break

Contributed Papers

3:20

1pAO5. Optimal search strategies for nonlinear, multi-parameter geoacoustic inverse problems. Gopu R. Potty, James H. Miller, Colin J. Lazauski, and Chuen-Song Chen (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882)

A mode travel time based technique for geoacoustic inversions has been developed using broadband explosive sources [Potty *et al.*, J. Acoust. Soc. Am. **108**, 973–986 (2000)]. In this inversion scheme we search for various parameters including water depth, sound speed in the water column represented by Empirical Orthogonal Function (EOFs), compressional wave speed in the sediment and sub-base, thickness of the sediment and range. The sensitivity of the objective function to perturbations in

these parameters is different, i.e., different parameters have different sensitivities. The most sensitive parameters can be estimated with higher accuracy compared to parameters with low sensitivities. In this study we aim to link the search process to parameter sensitivities by matching the fineness of search to sensitivities. The highly sensitive parameters will be searched with a fine scale sampling of the model space with a coarser sampling adopted for less sensitive model parameters. This condition will be incorporated into the search tool (Genetic Algorithm) used for the inversion scheme. The objective of this approach is to utilize the computational resources in an optimum manner by giving more emphasis to parameters which are more sensitive as opposed to a uniform weighting for all parameters. This approach will be applied to data from the ASIAEX-2001 and Shelf Break Primer data sets. [Work supported by ONR.]

1pAO6. Geoacoustic inversion techniques (GAIT) Version 1.0 global search (GS). Peter Neumann (Planning Systems, Inc., 408 Alleghany St., Blacksburg, VA 24060-5007) and Gregory Muncill (Planning Systems, Inc., Reston, VA 20191-3453)

Geoacoustic Inversion Techniques (GAIT) Version 1.0 is a PEO (C4I and Space) PMW 155 funded product that accepts measured acoustic data and produces an optimized estimate of the bottom environment that produced the observed acoustic data. The Global Search (GS) segment of GAIT pairs the Adaptive Simulated Annealing (ASA) algorithm with a variety of Navy standard propagation loss models (PE, ASTRAL and Nautilus) and an active sonar performance prediction model (ASPM). The goal of the GS segment of GAIT is to provide a best estimate of the geoacoustic properties of the ocean bottom that, when paired with a selected model, result in the observed acoustic data. An overview of the GS segment of GAIT 1.0 will be presented with details on the ASA algorithm, component models, cost functions and geoacoustic parametrizations. Inversion results will be shown for synthetic test cases from the Inversion Technique Workshop (ITW) held in May 2001 and from both narrowband and broadband measured data test cases. [Work supported by PEO (C4I and Space) PMW 155 and uses the products of a Phase I and II SBIR from the ONR (Code 321US).]

3:50

1pAO7. Active rapid geo-acoustic characterization. Kevin D. Heaney (Orincon Corp., 4350 N. Fairfax Dr., Ste. 470, Arlington, VA 22203)

There is a growing body of evidence that the use of archival database values to estimate acoustic propagation and reverberation leads to substantial errors. To address this issue, ideas have been proposed to use mono-static reverberation to perform simultaneous geo-acoustic and scattering strength inversions. The difficulty with this technique has to do with the underlying physics of propagation and reverberation. The current techniques of using mono-static reverberation to invert for scattering strength and the geo-acoustic parameters are fundamentally limited by the fact that all reverberant energy must propagate from the source to the bottom and then back to the receiver. Thus all reverberant energy contains twice the transmission loss (TL). This ambiguity is un-resolvable from only a direct mono-static reverberation measurement. This ambiguity represents itself as a 1 dB error in estimation of the TL yields a 2 dB error in the scattering strength (SS). Further, an accurate sediment characterization and scattering strength determination, handling the range-dependence of these variables, is required. To address this problem, and to break the inherent degeneracy, a combined reverberant direct blast inversion is proposed. This approach is applied to several sets of data from the Harsh Environmental Program (HEP).

4:05

1pAO8. Frechet derivatives for shallow water ocean acoustic inverse problems. Robert I. Odom (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

For any inverse problem, finding a model fitting the data is only half the problem. Most inverse problems of interest in ocean acoustics yield nonunique model solutions, and involve inevitable trade-offs between model and data resolution and variance. Problems of uniqueness and resolution and variance trade-offs can be addressed by examining the Frechet derivatives of the model-data functional with respect to the model variables. Tarantola [*Inverse Problem Theory* (Elsevier, Amsterdam, 1987), p. 613] published analytical formulas for the basic derivatives, e.g., derivatives of pressure with respect to elastic moduli and density. Other derivatives of interest, such as the derivative of transmission loss with respect to attenuation, can be easily constructed using the chain rule. For a range independent medium the analytical formulas involve only the Green's function and the vertical derivative of the Green's function for the medium. A crucial advantage of the analytical formulas for the Frechet derivatives over numerical differencing is that they can be computed with a single pass of any program which supplies the Green's function. Various

derivatives of interest in shallow water ocean acoustics are presented and illustrated by an application to the sensitivity of measured pressure to shallow water sediment properties. [Work supported by ONR.]

4:20

1pAO9. Estimation of error bounds in geoacoustic inversions. William Sanders (Naval Res. Lab., Code 7185, Stennis Space Center, MS 39529)

Geoacoustic inversion has been shown to yield accurate estimates of effective ocean bottom parameters for simple environments when high fidelity models are used for both environmental deconstruction and signal propagation. But as the environment becomes more complex, some parameters are estimated with less certainty than others. Uncertainty, as represented by errors in inverse problems stem from measurement inaccuracies, model imperfections and prior assumptions. Whereas the propagation of errors in inverse problems is not generally possible, the problem becomes tractable if the forward equation can be linearized about some (preferably the true) set of values and, if further, all errors are assumed Gaussian. Under these assumptions, the covariances of the *a posteriori* errors can be formulated, thus providing bounds on the uncertainty resulting from the inverse process. These are ultimately expressed in terms of bounds on measurement errors, modeling errors and the linearization of the forward model. This effort analyzes the errors involved in some well understood and benchmarked cases and compares results to other published analyses. This is done by utilizing a parabolic equation (PE) propagation model. Derivatives of the field with respect to the environmental variables are derived in order to calculate the error bounds.

4:35

1pAO10. Two methods for solving a 3D acoustic inverse scattering problem. Fadoulourahmane Seydou, Nail A. Gumerov, and Ramani Duraiswami (Perceptual Interfaces and Reality Lab., Inst. for Adv. Computer Studies, Univ. of Maryland, College Park, MD 20742)

We consider the problem of finding the refractive index of a buried object by using far-field measurements in an inhomogeneous medium. We describe two methods for solving the inverse problem. Both methods are implemented in two steps in order to better deal with the ill-posedness of the problem. In the first method an integral equation of the first kind is derived for the far-field operator which is solved via least-squares and Tikhonov regularization. We then use the solution of the integral equation to derive an over-posed boundary value problem, i.e., the Helmholtz equation in a bounded domain with Cauchy data on the boundary. The index that satisfies this over-posed problem most closely is obtained via the Levenberg-Marquardt algorithm. The second method is an iterative method and is based on the Lippmann-Schwinger equation. It is implemented via the Newton method. The first step consists, as for the other method. Here we use a Fourier integral approach and regularization via discretization. The second step is to obtain the index by iterating the Lippmann-Schwinger equation starting with the Born approximation.

4:50

1pAO11. Ray-based geoacoustic inversion for high frequency broadband data. Martin Siderius, Paul Hursky, and Michael Porter (SAIC, 10260 Campus Point Dr., San Diego, CA 92121)

One of the difficulties in making reliable acoustic propagation predictions in shallow water is the lack of good information about the seabed type. In recent years, matched field processing- (MFP-) based geoacoustic inversion has been shown as a practical technique for estimating properties of the seabed. The MFP inversion method compares measured acoustic fields to those generated using an acoustic propagation model. Often, thousands of forward model calculations are required to find a set of seabed parameters that correlate well with the measured data. The large number of forward model calculations is computationally demanding and this is made worse when matching at higher frequencies or over broad band data. Ray-based propagation modeling relieves some of the computational burden since the calculation time is fairly insensitive to frequency and is inherently broad band. Further, the ray arrival amplitudes and delays are

well suited for interpolation and this allows the seabed parameter search space to be explored using just a few ray trace calculations. The broad band nature of the modeled data provides flexibility in choosing correlation functions and this allows for more robust inversions. In this presentation, techniques using ray-based propagation modeling will be applied to the geoaoustic inversion problem.

5:05

1pAO12. Geoaoustic inversion of range-dependent shallow-water data using a complex plane-wave reflection coefficient approach. Steven A. Stotts, David P. Knobles, and Robert A. Koch (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, stotts@arlut.utexas.edu)

The inversion of range-dependent broadband data in shallow water is presented using a previously developed approach with rays and complex bottom plane-wave reflection coefficients [Stotts, J. Acoust. Soc. Am. **109**,

2334 (2001)], allowing separation of seabed physics from water column contributions. Using geometrical optics, waterborne eigenray characteristics are calculated once and stored. Model fields are calculated by including reflection coefficient contributions from each bottom interaction and are evaluated for each perturbation of the seabed properties using simulated annealing to minimize a Bartlett-type cost function. Inversion results from range-independent and -dependent data taken during recent experiments off the coast of Florida are presented using a two-layer model. Implosive source broadband data and XBTs were recorded out to 10 km from a bottom-mounted, 52-element array 229-m deep. Grab samples obtained during the experiment show a thin, hard, crusted surface sediment. Inversions reveal an underlying softer sediment. Model time series using one inverted parameter set are compared to the data obtained from sources at different ranges. Data arrivals are compared to eigenray arrivals providing further propagation insight. Consistent inversion results from other forward models are discussed.

MONDAY AFTERNOON, 28 APRIL 2003

ROOM 208, 1:15 TO 3:50 P.M.

Session 1pBB

Biomedical Ultrasound/Bioresponse to Vibration, Physical Acoustics and ASA Committee on Standards: Acoustic Cavitation

E. Carr Everbach, Chair

Department of Engineering, Swarthmore College, 500 College Avenue, Swarthmore, Pennsylvania 19081

Chair's Introduction—1:15

Invited Papers

1:25

1pBB1. Acoustic cavitation movies. Lawrence A. Crum (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

Acoustic cavitation is a phenomenon that occurs on microsecond time scales and micron length scales, yet, it has many macroscopic manifestations. Accordingly, it is often difficult, at least for the author, to form realistic physical descriptions of the specific mechanisms through which it expresses itself in our macroscopic world. For example, there are still many who believe that cavitation erosion is due to the shock wave that is emitted by bubble implosion, rather than the liquid jet created on asymmetric collapse... and they may be right. Over the years, the author has accumulated a number of movies and high-speed photographs of cavitation activity, which he uses to form his own visual references. In the time allotted, he will show a number of these movies and photographs and discuss their relevance to existing technological problems. A limited number of CDs containing the presented materials will be available to interested individuals. [Work supported in part by the NIH, USAMRMC, and the ONR.]

1:45

1pBB2. Microbubble dynamics in the ultrasound field. Yoichiro Matsumoto (Dept. of Mech. Eng., The Univ. of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan, ymats@tkd.att.ne.jp)

Ultrasonic biomedical applications such as ultrasound imaging with microbubble contrast agents, sonodynamic therapy, and so on, have attracted much attention in recent years. These phenomena have a close relationship to the behavior of microbubbles, so that it is essential to understanding the microbubble dynamics. A microbubble's motion in ultrasound field is simulated in detail by the full equations for mass, momentum, and energy in both phases. It is revealed that the bubble motion is much influenced by internal phenomena such as thermal diffusion, mist formation, mass diffusion, heat and mass transfer through the bubble wall. As the driving acoustic amplitude is increased, steep increases of the oscillation amplitude occur with the largest radius observed around the resonant bubble radius. Subharmonic and higher harmonic oscillations become obvious. The acoustic and thermal radiation from the microbubbles are analyzed and it is found that considerably stronger higher harmonic emissions are obtained from smaller bubbles than for the resonant bubbles.

1pBB3. Quantitative theoretical explanation of Apfel's experimental phase diagrams for sonoluminescing bubbles. Detlef Lohse (Dept. of Phys., Univ. of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands)

Robert Apfel had an enormous impact on the research of single bubble sonoluminescence [M. P. Brenner, S. Hilgenfeldt, and D. Lohse, *Rev. Mod. Phys.* **74**, 425 (2002)], the light emission of a single sound-driven bubble. In 1996, at the ASA Meeting in Hawaii, he posed a challenge to the theoreticians in the field: make experimentally testable predictions on single bubble sonoluminescence. Apfel collected the predictions and gave a wonderful review talk on the state of the field. Later, he came back several times to that list, comparing the predictions with latest experimental results. Our own predictions in those days referred to the phase diagrams of single bubble sonoluminescence. Later Apfel himself, together with Ketterling, measured those phase diagrams experimentally [J. A. Ketterling and R. E. Apfel, *Phys. Rev. Lett.* **81**, 4991 (1998); *Phys. Rev. E* **61**, 3832 (2000)]. Though qualitatively our 1996 predictions turned out to be correct, a full quantitative model could only be developed recently [R. Toegel and D. Lohse, *J. Chem. Phys.* (in press); B. D. Storey and A. J. Szeri, *Phys. Rev. Lett.* **88**, 074301 (2002)]. In the presentation we will compare the model predictions with Apfel's data.

1pBB4. The oscillations of vapor bubbles. Andrea Prosperetti, Z. Yin, and B. Yang (Dept. of Mech. Eng., Johns Hopkins Univ., 119 Latrobe Hall, 3400 N. Charles St., Baltimore, MD 21218, prosperetti@jhu.edu)

Bob Apfel had so many interests that it is impossible—however fitting and desirable—to pay homage to his work as a whole. Some of his early studies were devoted to bubble nucleation at high superheats. In the first part of this paper a recent application of this phenomenon is described. Once a vapor bubble is generated, its subsequent oscillations (free and forced) present analogies and differences with those of a gas bubble: the second part of the paper focuses on this topic. [Work supported by NSF and NASA.]

1pBB5. What are the limits of energy focusing in sonoluminescence? Seth Putterman, C. Camara, B. Kappus, C. K. Su (Phys. Dept., UCLA, Los Angeles CA 90095, putherma@ritva.physics.ucla.edu), and E. Kirilov

Sonoluminescence [SL] is amazing for the extraordinary degree by which ultrasonic energy can be focused by a cavitating bubble. Local energy dissipation exceeds Kirchhoff's law by $1E15$ and the acoustic energy density concentrates by 12 orders of magnitude to create picosecond flashes of broadband ultraviolet light. At the minimum bubble radius, the acceleration exceeds $1E11$ g and a megabar level shock wave is emitted into the surrounding fluid. For single bubbles driven at 30 KHz, SL is nature's smallest blackbody. This implies that the bubble's interior is such a dense plasma that the photon-matter mean free path is shorter than the wavelength of light, and suggests that SL originates in an unusual state of matter. Excitation of a vertical column of fluid [~ 10 Hz] so as to create a water hammer leads to the upscaling of SL and generation of flashes of light with $3E8$ photons and peak powers approaching 1 W. At 1 MHz, the spectrum resembles bremsstrahlung from a transparent plasma with a temperature ~ 1 MK. At 10 MHz the collapsed size of the SL bubble approaches 10 nm, which raises the possibility that the SL parameter space may extend to the domain of quantum mechanics. [Research supported by DARPA and DOE.]

Contributed Papers

1pBB6. Light-scattering measurements of cavitation from ultrasound contrast agent bubbles. Jingfeng Guan, Wen-Shiang Chen, and Thomas Matula (Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, matula@apl.washington.edu)

Ultrasound contrast agents are typically micron-sized gas bodies with stabilizing shell coatings. The shell prevents the gas bubble from dissolving. The shell also changes a bubble's scattering properties and affects a bubble's destruction. Optimization of the shell can be beneficial for diagnostic imaging and therapeutic ultrasound applications. Thus, understanding agent response to pulsed ultrasound is important. However, because the agents are so small, their response to pulsed ultrasound is difficult to measure. Until recently, the only tool available for measuring contrast agent dynamics was a very expensive high-speed camera. We have been investigating the use of light-scattering to measure aspects of cavitation from contrast agents in which a small HeNe laser beam scatters off a bubble and is focused onto a photomultiplier tube detector. Previous experiments have involved clusters of agents, where collective bubble oscillations and destruction were observed. Our current apparatus is designed to observe single-bubble oscillations from a diagnostic ultrasound system. The advantages for using light-scattering include small cost and the ability to monitor bubble oscillations for extended times. The technique will be described, and representative data samples will be shown.

1pBB7. Use of intensity interferometry to determine the size of a cavitation hot-spot. Carlos Camara, Seth Putterman (Phys. Dept., Univ. of California, Los Angeles, CA 90095-1547, camara@physics.ucla.edu), Keith Weninger (Stanford Univ., Stanford, CA 94305-4060), and Paul Evans (Dept. of Mater. Sci. Eng., Wisconsin Univ., Madison, WI 53706)

Although it is generally accepted that a sonoluminescing bubble collapses to a volume determined by the van Der Waals hard core of its contents, the size of the light-emitting region remains undetermined. Photos of the hot-spot indicate that the resolution of its size is diffraction limited. Intensity interferometry can resolve smaller sizes as well as the flash width if the light is incoherent. An experiment aimed at exploiting this is underway.

1pBB8. Nonlinear bubble dynamics in viscoelastic media. Stanislav Y. Emelianov (Dept. of Biomed. Eng., Univ. of Texas, Austin, TX 78712-1084), Mark F. Hamilton, Yuri A. Ilinskii, and Evgenia A. Zabolotskaya (Univ. of Texas, Austin, TX 78712-1063)

Microbubbles are already widely used for biomedical and clinical purposes, and are further being investigated for innovative medical and industrial applications. In most cases, the Rayleigh-Plesset equation for a gas bubble in liquid is used. In this paper we present a more general equation for gas bubble dynamics in viscoelastic media such as soft tissue.

The equation, describing oscillations of a spherical bubble in an isotropic, incompressible medium, contains an integral over an elastic energy density function that depends on two invariants of the strain tensor. For moderate bubble oscillations one may use Landau's cubic expansion of energy density, which contains five elastic moduli—shear and bulk moduli corresponding to quadratic terms, and three moduli (A , B , and C) corresponding to cubic terms. With energy density expressed through invariants of

the strain tensor in spherical coordinates, the total elastic energy for an incompressible medium becomes an explicit function of bubble radius and contains only two of the five moduli, the shear modulus and modulus A . Compared to the Rayleigh–Plesset equation, the resulting equation has additional terms associated with shear force and elastic nonlinearity. Further modifications of this equation account for weak compressibility, viscosity, surface tension, radiation loss, and elastic shells.

MONDAY AFTERNOON, 28 APRIL 2003

ROOMS 209/210, 1:00 TO 4:35 P.M.

Session 1pNS

Noise, ASA Committee on Standards, Signal Processing in Acoustics and Architectural Acoustics: Dan Johnson Retrospective

Mary M. Prince, Cochair

National Institute for Occupational Safety and Health, 4676 Columbia Parkway, M.S. R-16, Cincinnati, Ohio 45226

John P. Seiler, Cochair

Mine Safety and Health Administration, Cochran Mill Road, Building 038, P.O. Box 1823, Pittsburgh, Pennsylvania 15236

Chair's Introduction—1:00

Invited Papers

1:05

1pNS1. Dan Johnson—Mr. Standards. Paul Schomer (Schomer and Assoc., Inc., 2117 Robert Dr., Champaign, IL 61821, schomer@SchomerAndAssociates.com)

Dan Johnson had a long history of involvement in the standards program of ASA. He has chaired and/or participated in many working groups, primarily in the areas of hearing conservation and instrumentation. He was one of a rare set that was an individual expert to all four S-committees, being very familiar with instrument design from his days at Larson-Davis, and being familiar with physiological and psychological effects, noise, and vibratory effects from his work with the Air Force at Wright–Patterson AFB. Dan was the chair of S-12, the vice-chair of ASACOS, and, most recently, the chair of ASACOS and Standards Director. As Standards Director, Dan guided us through many turbulent times including establishment of the Standards Store, the change of the standards manager, the budgetary trials of 2000 and 2001, and the appeal of ANSI S12.60-2002 (Classroom Acoustics) by ARI before Accredited Standards Committee S12. He is very missed.

1:25

1pNS2. Monitoring audiometry for occupational hearing loss: A case for eliminating 500 Hz. Mark Stephenson (Natl. Inst. for Occupational Safety and Health, 4676 Columbia Pkwy., Cincinnati, OH 45226)

Audiometric monitoring is an important element in hearing conservation programs. Nearly every existing hearing conservation standard dictate that hearing thresholds should be measured at specific frequencies, and that 500 Hz be among those frequencies tested. Actual and estimated noise-induced permanent threshold shifts were evaluated as a function of exposure duration and exposure level. The results demonstrate 500 Hz to be of little value in assessing noise-induced hearing loss for typical industrial noise exposures of up to 40 years, at least for time-weighted average exposures of up to 100 dB A. Furthermore, few hearing conservation programs currently require audiometric monitoring to be performed in an environment that meets ANSI standards for maximum permissible background noise levels. This is particularly likely to compromise hearing testing at 500 Hz. As a result, this paper argues against the need for testing at 500 Hz, and recommends it be eliminated as a required test frequency in audiometric monitoring for noise-induced occupational hearing loss.

1:45

1pNS3. Predicting noise-induced hearing loss in human populations. The contribution of the ISO-1999 and ANSI S3.44 standards. Mary M. Prince (Centers for Disease Control, Natl. Inst. for Occupational Safety and Health, 4676 Columbia Pkwy., R-16, Cincinnati, OH 45226)

This paper reviews the ISO-1999 and ANSI S3.44 standards and the literature related to prediction of noise-induced hearing loss (NIHL) in humans. The dearth of contemporary population-based studies of hearing loss among industrial, low noise-exposed individuals have made the ISO-1999 and ANSI S3.44 standards a key source of comparative data for assessing the risk of NIHL. This paper will highlight how these standards have been used in several published papers to examine the magnitude of NIHL risk, predictions of noise-induced thresholds in hearing conservation data, and to generate hypotheses regarding the biologic plausibility of hearing loss due to chemicals or other ototoxic agents.

1pNS4. Dan Johnson's impact on hearing research. Lawrence I. Shotland (James H. Quillen VA Medical Ctr., Mountain Home, TN and Dept. of Communicative Disord., East Tennessee State Univ., Johnson City, TN)

Daniel L. Johnson is well known for his many technical contributions to noise research. Throughout a long and distinguished career at Wright–Patterson Air Force Base, Dan published the results of several significant experiments, including his landmark experiments on asymptotic threshold shift and exposure to impulse noise. His work in the area of noise exposure laid much of the groundwork for a greater understanding of the physiologic response to hazardous noise, much of which has since been incorporated in national and international standards. Dan is highly regarded for his tireless work on technical and advisory committees in noise, and most recently, ototoxicity. Throughout his career, Dan has adhered to a self-imposed standard of intellectual honesty and discovery. Dan's most recent endeavor, the development of a personal noise dosimeter designed for self-monitoring by the employee, is characteristic of his creativity and energy. Perhaps less well known are his contributions over the years to the success of his younger colleagues. He has accomplished this in an unselfish and egalitarian manner, oftentimes challenging and even contradicting his own research. The focus of this talk will elaborate on these facets of Dan's professional contributions.

2:25

1pNS5. A CHAT with Dan Johnson. Robert A. Bertrand (Bertrand Johnson Acoust., Inc., 5995 Gouin Blvd., Montreal, QC H4J 2P8, Canada)

Dan's preoccupation with preventing NIHL is well known. For those of us who had the occasion to CHAT with him, we often heard him say "if ears would bleed when exposed to noise, people would pay more attention to the harmful effect of noise upon hearing." His objective, as he often stated, was to eradicate NIHL so that in a few decades, it would become a historical footnote. Among his many preoccupations in his illustrious career, approaches and techniques to prevent NIHL were of primordial importance. One approach he advocated is the use of TTS instead of the STS in HCP's. His intention was to use the identification of TTS as an easy approach to introduce appropriate measures to prevent NIHL at a stage of reversibility rather than waiting for a confirmed permanent hearing loss, as noted with a confirmed STS. One of his last projects was developing the CHAT (Change of Hearing Audio Test) for easy use both in industrial and environmental settings to identify subjects with a TTS. Several groups are interested in pursuing his aim of using the TTS in HCP's, hoping to fulfill his comment that in a few decades, NIHL will become a historical footnote.

2:45

1pNS6. Dan Johnson the mentor. Richard McKinley (Air Force Res. Lab., 2610 Seventh St., Wright–Patterson Air Force Base, OH 45433)

I first met Dan Johnson in early 1975 as I was interviewing for an engineering job with Henning von Gierke's bioengineering and bionics laboratory at Wright–Patterson Air Force Base. From the very beginning Dan was always direct and forthright. Over the ensuing next 27 years my knowledge and respect of Dan constantly grew. This presentation will review Dan's technical and personal contributions while at the laboratory at Wright–Patterson Air Force Base. He was instrumental in the development of a national noise exposure criteria with the equal-energy-rule, an accurate single number hearing protector attenuation measure based on "C-A," an impulse noise exposure criteria, a longitudinal study of hearing loss in children, development of noise dosimeters, and description of hearing damage risk from nonoccupational noise exposures such as disco's, bowling alleys, lawn mowers, and school buses. Dan has had a significant effect on my career. I and the many people who knew him at the laboratory miss him greatly.

3:05–3:20 Break

Contributed Papers

3:20

1pNS7. Personal noise exposure assessment from small firearms. Chucui A. Kardous, William J. Murphy, and Robert D. Willson (NIOSH, Hearing Loss Prevention Section, 4676 Columbia Pkwy., M.S. C-27, Cincinnati, OH 45226-1998)

The National Institute for Occupational Safety and Health conducted noise exposure evaluations of law-enforcement personnel during firearms training at indoor and outdoor firing ranges. A representative cross section of weapons used by officers was measured. Shooters participated in live-fire exercise at an indoor firing range using three different weapons: a Beretta .400 caliber pistol, a Remington 12-gauge shotgun, and an M4 .223-caliber assault rifle. Indoor and outdoor measurements were obtained for the Smith and Wesson .357 pistol and Colt .450 and 9-mm pistols, the Glock .400 pistol, and the Heckler and Koch and Colt AR15 .223 rifles. Impulses were measured using a Bruel and Kjaer 4136 1/4-in. microphone and TASCAM digital audio tape recorder. Relevant impulse noise metrics were calculated. Peak levels ranged from 155 to 168 dB SPL. A-weighted equivalent levels ranged from 124 to 128 dBA. The contributions of the secondary weapon firings were approximately 1 to 9 dBA. Other parameters such as A/B durations, number and mixture of impulses, spectral content, energy, kurtosis, temporal spacing, and hearing protectors' effectiveness were examined. Comparisons of applicable damage risk criteria are presented. Further studies are needed to establish an occupational impulse noise damage risk criterion.

3:35

1pNS8. Attenuation measurements of passive linear and nonlinear hearing protectors for impulse noise. William J. Murphy and Chucui A. Kardous (NIOSH, Hearing Loss Prevention Section, 4676 Columbia Pkwy., M.S. C-27, Cincinnati, OH 45226-1998)

As a part of a NIOSH Health Hazard Evaluation of law-enforcement personnel, the attenuation of several types of earplugs were measured in response to impulse noise produced by small-arms gunfire. The earplugs were primarily flanged premolded plugs produced by EAR/Aearo and Bilsom/Baccou-Dalloz. Measurements for the North Sonic Ear Valve, EAR Classic earplugs, and EAR Ultra 9000 passive nonlinear ear muff were conducted. The EAR premolded earplugs were the Combat Arms passive linear and nonlinear, HiFi and Ultratech earplugs. The Bilsom devices were the 555, 655 NST, and 655 ISL earplugs. The Combat Arms and 655 ISL earplugs both utilize a cartridge developed by the French German Research Institute de Saint Louis that provides nonlinear attenuation. The peak reduction of these devices ranged between 10 and 28 dB.

The slope of peak reduction with peak level for the Ultra9000 device was about 0.5 dB/dB, while the slopes for most earplugs were about 0.1 to 0.3 dB/dB for weapons impulses between 159- and 170-dB peak level. The peak reductions ranged from 6 dB for the North Ear valve to 30 dB for the EAR Classic foam earplug.

3:50

1pNS9. Relation between peak pressure and spectrum for small caliber muzzle report and ballistic shock. Michael J. White and Larry L. Pater (U.S. Army ERDC/CERL, P.O. Box 9005, Champaign, IL 61826)

It might seem impossible to reconstruct the peak signal level of an impulsive sound from its spectrum if the phase information is not kept. Without such an estimate, sound propagation assessments made in the frequency domain are at a disadvantage when it is the peak level that is desired. From a set of experimental measurements of small caliber firearm noise at Camp Guernsey, muzzle report and ballistic shock signals were time-gated to obtain 1/3-octave sound exposure spectra and peak levels from each signal. The rms bandwidth of each signal was determined by summing the squared frequency, weighted by the relative 1/3-octave band exposures. Finally, a relation was presumed only such that the sound exposure was equal to a product of squared peak pressure, the reciprocal of the rms bandwidth, and an undetermined constant. Constants equal to 3 dB for muzzle blast and 3.5 dB for ballistic shock fit the data sets, with approximately 3 dB standard error. Greater error tended to coincide with smaller measured bandwidths, by overestimating the peak level.

4:05

1pNS10. Noise sampling issues for impact/impulse noise surveys. Mary M. Prince (NIOSH, IWSB/DSHEFS, 4676 Columbia Pkwy., R-16, Cincinnati, OH 45226) and Jeffrey S. Viperman (Dept. of Mech. Eng., Univ. of Pittsburgh, 531 Benedum Hall, Pittsburgh, PA 15261)

Noise-induced hearing loss (NIHL) has been recognized as a serious health concern for decades. ISO Standard 1999:1990 provides a means to predict noise-induced hearing loss (NIHL) based on L_{Aeq} measurements in the working environments of workers. This standard seems to work well for predicting hearing loss in continuous noise fields. However, it is pos-

sible that ISO 1999 does not apply well to impact, impulsive, or other transient noise fields. NIOSH and University of Pittsburgh are currently developing noise-sampling strategies to measure impact and impulse noise in a manufacturing environment with the aim of developing new impulsive noise metrics. As part of the study, broadband impact/impulse pressure measurements will be made. Issues such as instrumentation, data quality, repeatability, spatial sampling, equipment portability, and calibration are addressed. Also, the annotation, digitization, and editing of the waveforms will be discussed. As part of the project, an archival database of manufacturing impulse/impact will be created to support the future algorithmic development. The ultimate goal of the project is to develop new metrics to characterize the hazards of impact/impulse noise that will complement ISO 1999 for predicting NIHL.

4:20

1pNS11. Analysis of impact/impulse noise for predicting noise induced hearing loss. Jeffrey S. Viperman (Dept. of Mech. Eng., Univ. of Pittsburgh, 531 Benedum Hall, Pittsburgh, PA 15261), Mary M. Prince (NIOSH, IWSB/DSHEFS, Cincinnati, OH 45226), and Angela M. Flamm (Univ. of Pittsburgh, Pittsburgh, PA 15261)

Studies indicate that the statistical properties and temporal structure of the sound signal are important in determining the extent of hearing hazard. As part of a pilot study to examine hearing conservation program effectiveness, NIOSH collected noise samples of impact noise sources in an automobile stamping plant, focusing on jobs with peak sound levels (L_{pk}) of greater than 120 dB. Digital tape recordings of sounds were collected using a Type I Precision Sound Level Meter and microphone connected to a DAT tape recorder. The events were archived and processed as .wav files to extract single events of interest on CD-R media and CD audio media. A preliminary analysis of sample wavelet files was conducted to characterize each event using metrics such as the number of impulses per unit time, the repetition rate or temporal pattern of these impulses, index of peakedness, crest factor, kurtosis, coefficient of kurtosis, rise time, fall time, and peak time. The spectrum, duration, and inverse of duration for each waveform were also computed. Finally, the data were evaluated with the Auditory Hazard Assessment Algorithm (AHAAH). Improvements to data collection for a future study examining different strategies for evaluating industrial noise exposure will be discussed.

MONDAY AFTERNOON, 28 APRIL 2003

ROOM 204, 2:00 TO 4:40 P.M.

Session 1pPP

Psychological and Physiological Acoustics: In Memory of Evan Relkin

Robert L. Smith, Chair

Institute for Sensory Research, Syracuse University, 621 Skytop Road, Syracuse, New York 13244-5290

Chair's Introduction—2:00

Invited Papers

2:10

1pPP1. Estimating the similarity of temporal discharge patterns of auditory neurons. William Shofner (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, wshofne@luc.edu)

Behavioral data obtained from stimulus generalization experiments are often interpreted to imply how similar animal perceptions of test stimuli are to a training stimulus. An appropriate decision variable would be useful for evaluating physiological responses in the context of stimulus generalization tasks. The decision variable $P(A)$ obtained from ROC analysis of spike count distributions is useful for evaluating rate coding schemes [e.g., Relkin and Pelli, J. Acoust. Soc. Am. **82**, 1679–1691 (1987)], but it ignores temporal information which may exist in responses to periodic sounds. Neural autocorrelograms show temporal firing patterns following a spike

and were generated for cochlear nucleus unit responses to aperiodic, quasiperiodic, and periodic stimuli. Variance from average rate was computed as a function of time for each autocorrelogram; similarity between two autocorrelograms was estimated as the ratio of these variances. A value of 1 indicates that the autocorrelograms are identical, whereas larger values indicate less similarity. Ratios were around 1 if both autocorrelograms were from responses to wideband noise and were greater than 1 if one autocorrelogram was from responses to iterated rippled noises or harmonic complex tones. This ratio may be a useful decision variable for evaluating temporal coding schemes. [Work supported by NIDCD P01 DC00293.]

2:35

1pPP2. Can we learn anything new from the animal's perspective?: Bayesian stimulus estimation based on neural spike trains. B. Scott Jackson (Inst. for Sensory Res., Syracuse Univ., 621 Skytop Rd., Syracuse, NY 13244-5290, Scott_Jackson@isr.syr.edu)

The classical approach to understanding a neural coding scheme is to characterize spike trains elicited by multiple presentations of the same stimulus, an example being the rate-intensity function. However, the task for the organism is deciphering the stimulus content from spike trains elicited by single stimulus presentations, a process that is more analogous to Bayesian stimulus estimation. The usefulness of Bayesian stimulus estimation in the neural context, using the auditory periphery as a model system, was investigated. More specifically, the encoding of stimulus intensity in the spike rate of single primary auditory neurons was examined. It was found that the results of this method are heavily influenced by the *a priori* stimulus distribution and that apparent benefits of this approach, such as linearization and, in some instances, increased stimulus dynamic range, are offset by concomitant disadvantages, such as increased estimation error and decreased stimulus dynamic range in other instances. Hence, in this context, Bayesian stimulus estimation does not contribute meaningful additional knowledge to our understanding of neural coding, and these results suggest that related methodologies, such as stimulus reconstruction and information theoretic methods, be used and interpreted with caution. [Work supported by Syracuse University and NIH Grant 5-P01-DC000380.]

3:00

1pPP3. Monaural and diotic detection of tones in wideband and narrow-band reproducible noise maskers. Sean A. Davidson and Laurel H. Carney (Dept. of Bioengineering and Neurosci. and Inst. for Sensory Res., Syracuse Univ., 621 Skytop Rd., Syracuse, NY 13244, sadavids@syr.edu)

In order to test monaural and binaural processing components of binaural detection, diotic and monaural detection were studied using reproducible noises. Correlations across masker bandwidth and interaural configuration were computed. Preliminary results reveal high correlations for probability of detection across N_0S_0 and N_mS_m configurations for both wideband and narrow-band maskers, suggesting that the binaural system does not influence results across reproducible noises for diotic detection. As reported previously [Evilsizer *et al.*, J. Acoust. Soc. Am. **111**, 336–345 (2001)], significant differences in correlations between wideband and narrow-band results were observed for the diotic case, suggesting cross-frequency interactions. Preliminary findings indicate that the same result holds for the monaural case. These data will be used to test models that incorporate monaural processing based on the phase-opponency model [Carney *et al.*, Acta Acustica United with Acustica **88**, 334–347 (2002)], and binaural processing based on the Jeffress model and on the two-channel model [McAlpine *et al.*, Nature Neurosci. **4**, 396–401 (2001)]. Predictions will be made for N_mS_m , N_0S_0 , and N_0S_π interaural configurations, and for wideband and narrow-band maskers. [Work supported by NIDCD.]

3:25–3:40 Break

Contributed Papers

3:40

1pPP4. The mid-duration hump in the intensity-difference limen as a function of frequency: Further evidence for the frequency–time listening window of the human ear. Lance Nizami and Walt Jesteadt (Ctr. for Hearing Res., Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

It has been suggested that the ear has a frequency–time listening window and that the stimulus that occupies the fewest frequency–time windows, i.e., that which has the smallest product of effective bandwidth and effective duration, should result in the largest intensity-difference limen (DL) [Van Schijndel *et al.*, J. Acoust. Soc. Am. **105** (1999)]. Van Schijndel *et al.* identified the Gaussian-shaped tone-pip as the optimal stimulus, obtaining DLs at 1 kHz and 4 kHz. The DLs were greatest for the tone-pips whose frequency–time window was one critical band wide and 4 cycles long in equivalent square duration D (envelope duration/ 3.19). Here, DLs for Gaussian-shaped 250-Hz tone-pips were obtained from 5 subjects for Ds of 1.25–22.57 ms at levels of 40, 50, 60, 70, 80, and 90 dB SPL. These data were combined with earlier results showing a mid-duration rise at 40–60 dB SPL for Gaussian-shaped tone-pips of 6.5 kHz, 2 kHz, and 500 Hz. The peak in the DL is sharp for 6.5 kHz but flattens as frequency

drops, being replaced by a broad rise for 250 Hz. The highest DLs correspond to approximately 4 cycles in D, not far from the ringing duration calculated as the inverse of the critical bandwidth.

3:55

1pPP5. Growth of interleaved masking patterns for cochlear implant listeners at different stimulation rates. Bom Jun Kwon, Chris van den Honert, and Wendy Parkinson (Cochlear Americas, 400 Inverness Pkwy., Ste. 400, Englewood, CO 80112)

This study investigates the pattern of growth of masking (GOM) for interleaved masking with Nucleus cochlear implant users. For an interleaved masking paradigm, where the masker and probe overlap in a same time window, the masker may have contrasting effects: it may increase the threshold (as a masker normally does) or decrease it due to a neural summation effect, facilitating detection of the probe. Several stimulation rates and masker levels were tested to examine under what conditions what phenomenon would occur. The results indicated that, in most of the conditions, the amount of masking was positive, i.e., the facilitating effect was not consistently observed. However, the slope of the GOM appears to be

dependent upon the stimulation rate: the higher the stimulation rate, the lower the slope, implying that the facilitating effect might be always present and make a bigger impact on overall masking as the stimulation rate becomes high. The amount of masking was also often nonzero (positive) even when the masker was below the threshold level. Overall, the present findings indicate that interleaved masking should be handled with care to understand cochlear implant users speech perception and improve speech coding, as it contains some nontraditional aspects of masking.

4:10

1pPP6. Biologically inspired robust onset detection. Leslie S. Smith (Dept. of Computing Sci. and Mathematics, Univ. of Stirling, Stirling FK9 4LA, Scotland, UK)

Onsets are rapid increases in signal strength. The common onset time in different frequency bands provides an important cue for dynamically grouping sound energy, and hence for sound streaming. Onsets are important for segmenting sounds [Smith, J. New Music Res. **23**, 11–23 (1994)] and for determining where to measure IIDs and ITDs for the sound direction finding [Smith, J. Acoust. Soc. Am. **111**, 2467 (2002)]. Effective onset detection requires low latency and the capacity to cope with wide variation in the dynamic range. Many neurons in the auditory brainstem are sensitive to onsets. The system filters sound into cochlea-like bands (using a Gammatone filterbank), then spike codes positive-going zero-crossings. A wide dynamic range is achieved by using multiple spike trains per filter band, each with different sensitivity. The spike trains from each band innervate a leaky integrate-and-fire neuron. The excitatory synapses from the spike trains are fast and depressing: the shunting inhibitory synapses are facilitating and slower. The combined effect is that the neuron pro-

duces a single spike for each onset over a wide dynamic range with very low latency. The use of both inhibitory and excitatory synapses improves onset detection over purely excitatory synapses, leading to a better sound direction finding than previously reported.

4:25

1pPP7. Otoacoustic emissions and the “set of the center:” Covariances linking ear, brainstem, and cortex, documented with the AXS Test Battery. Judith L. Lauter (Stephen F. Austin State Univ., Nacogdoches, TX 75962, jlauter@sfasu.edu)

Although otoacoustic emissions (OAEs) are believed to reflect afferent processes connecting the central auditory system with the ear, such relations have not been demonstrated in humans, for central components rostral to the lower pons. The Auditory Cross-Section (AXS) Test Battery (Lauter, 2000, 2002) provides a straightforward way to do this, with the result that one can observe covariances between any of a number of OAE variables (TEOAE amplitude, DPOAE stability, etc.), and measures of the physiological status of the auditory nerve, brainstem centers from caudal pons to midbrain, and auditory cortex. The testing methodologies employed are relatively inexpensive and should be within the means of many more clinics and laboratories than can afford access to brain-imaging technologies such as fMRI and PET. In addition, a new statistical means of testing covariation in multivariate, within-subject data sets (Ninness and Lauter, 2003) will be described that offers a means of mathematically testing the results. Data illustrating this approach will be presented, documenting covariances observed during spontaneous day-to-day systemic fluctuations, as well as responses to medications.

MONDAY AFTERNOON, 28 APRIL 2003

ROOM 203, 1:15 TO 2:45 P.M.

Session 1pSP

Signal Processing in Acoustics and Speech Communication: Speech Processing

Douglas O’Shaughnessy, Cochair

INRS-Telecommunications, Place Bonaventure, Suite 6900, 900 de la Gauchetiere West, Montreal, Quebec H5A 1C6, Canada

Deborah M. Grove, Cochair

Zargis Medical Corporation, 755 College Road East, Princeton, New Jersey 08540

Contributed Papers

1:15

1pSP1. Robust automatic recognition of telephone speech. Douglas O’Shaughnessy and Selouani Sid-Ahmed (INRS-Telecommunications, 800 de la Gauchetiere West, Ste. 6900, Montreal, QC H5A 1K6, Canada)

A method is presented to compensate cepstral coefficients (MFCCs) in a speech recognition system for degraded telephone channel conditions. The technique proposed is based on a combination of the Karhonen–Loeve Transform (KLT) and Genetic Algorithms (GA). The idea consists of projecting the band-limited MFCCs onto a subspace generated by the genetically optimized KLT principal axes. Experiments show a clear improvement when the method was applied to the NTIMIT speech database. Word recognition results obtained on the HTK toolkit platform using *N*-mixture tri-phone models and a bigram language model are presented and discussed.

1:30

1pSP2. Speech enhancement using dynamic synapse neural networks. Hassan H. Namarvar and Theodore W. Berger (Dept. of BME, Univ. of Southern California, OHE-500, Los Angeles, CA 90089-1451, heidarin@usc.edu)

An idea of speech enhancement using a Dynamic Synapse Neural Network (DSNN) with an extended Kalman filtering (EKF) training method is described. The goal of this study is to introduce a new methodology in better speech enhancement in the presence of continuous environment background noise, such as fans and air-conditioning units. The efficiency of this method is shown by applying it to noisy speech signals to remove recorded laboratory noise from signals at different signal-to-noise ratio levels. The preliminary results have been encouraging enough to justify our idea. To provide more noise robustness, this could be used as

a pre-processing level in automatic speech recognition (ASR) systems. The proposed method would have a profound impact on the performance of ASR systems. [Work supported by DARPA CBS, NASA, and ONR.]

1:45

1pSP3. Text-to-speech from concatenation of articulatory units derived from natural speech. Daniel J. Sinder and M. Mohan Sondhi (Avaya Labs, 233 Mt. Airy Rd., Basking Ridge, NJ 07920, sinder@avaya.com)

It has been conjectured that articulatory synthesis possesses the great potential for generating high quality synthetic speech. However, for text-to-speech (TTS), waveform concatenation techniques have proven more practical due in part to the challenge of generating appropriate trajectories of articulatory parameters. A waveform generation method for TTS that combines the practical success of concatenative methods with the quality potential of articulatory synthesis is under development. The system concatenates articulatory units derived from natural speech using an articulatory voice mimic. The mimic estimates articulatory parameters by minimizing a cost function that includes a spectral distance between natural and synthetic speech and a geometric distance that penalizes rapid or discontinuous changes in articulator positions. A database of articulatory trajectories representing phonetic units is constructed from the estimated parameters. For TTS, phonetic units generated by text analysis are used to select the corresponding articulatory units from the database. Duration modification, concatenation, and smoothing across units are performed in the articulatory domain resulting in a single articulatory trajectory for the complete utterance. Speech is synthesized from the trajectory using a two mass model for voicing, achieving a high degree of acoustic continuity across unit boundaries while also allowing for source-tract interaction.

2:00

1pSP4. A measure of aperiodicity content in a speech signal. Om D. Deshmukh and Carol Y. Espy-Wilson (ECE Dept., Univ. of Maryland, College Park, MD 20782)

Most of the current aperiodicity detectors measure aperiodicity indirectly, in which the absence of periodic component in a nonsilent region is termed as aperiodicity. Such indirect measurements are inadequate and can be misleading especially in cases like voiced fricatives or breathy vowels. This motivated us to develop a direct measure of aperiodicity which is independent of the periodicity measure. The speech signal is passed through a 60-channel gamma tone auditory filterbank. Average Magnitude Difference Function (AMDF) is computed on the envelope of each channel. The randomness in the distribution of the AMDF dips is the basis for the measure of the aperiodicity whereas the measure of the periodicity is based on the occurrence of the dips at multiple locations. The system was evaluated on the MOCHA database, which has simultaneous recording of the EGG data, and on the TIMIT database. Preliminary analysis shows that the aperiodicity and voicing accuracy on a per frame basis are 95% and 90.3%, respectively. In voiced fricatives and voiced stops, high aperiodicity and high periodicity were detected in 21.6% of the frames. Note that not all these sounds necessarily had both the sources prominent.

2:15

1pSP5. Spectral moments and alternative methods of characterizing fricatives. Oliver S. Blacklock and Christine H. Shadle (Dept. of Elec. and Computer Sci., Univ. of Southampton, Southampton SO17 1BJ, UK)

Parameters that both distinguish fricatives in a given language, and characterize normal and disordered productions would be useful. Various parametrizations have been tried since the 1950s with only partial success. Forrest *et al.* [J. Acoust. Soc. Am. **84**, 115–123 (1988)] defined spectral moments of speech power spectra, but found that these failed to distinguish the voiceless English fricatives. They have however proven useful for distinguishing sibilants, and for characterizing long-term place changes of sibilants in disordered speech. As discussed in this paper, other changes, and nonsibilants, are not so easily characterized. Relative amplitude differences have to date not been incorporated in spectral moment methodology. The quality of the spectral estimate (whether consistent, unbiased) affects the variability of the moments, and means of obtaining a good estimate usually involve assuming stationarity or ergodicity. Finally, moments capture little that is known about fricative production. Modern techniques for consistent PSD estimation (e.g., multitaper and wavelet analysis) have been examined; their suitability in analyzing fricatives is discussed. While these modern PSD estimation methods may improve the performance of spectral moments, they also lend themselves to new parametric methods which do not rely on assumptions of ergodicity or stationarity. [Research supported by EPSRC.]

2:30

1pSP6. Speaker identification by difference sum and correlation coefficients of narrow-band spectrum. Byunggon Yang (English Dept., Dongeui Univ., 24 Kayadong Pusanjingu, Pusan 614-714, Korea) and SunMee Kang (Seokyeong Univ., Jungneung-Dong Sungbuk-Ku, Seoul 136-704, Korea)

Some problems in speaker identification procedures were examined: transformation of acoustic parameters into auditory scales, invalid measurement values, and comparability of spectral energy values across the frequency range. To resolve those problems, the acoustic spectral energy of three Korean numbers produced by ten female students from narrow-band spectrograms at 19 proportional time points of each voiced segment were analyzed. Then, cells of the first five spectral matrices were averaged to form a matrix model for each speaker. The correlation coefficients and sum of the absolute amplitude difference in each pair of the spectral models of the ten subjects were obtained. Also, some individual matrix models were compared to those of the same subject or the other subject with a similar spectral model. Results showed that in numbers “2” and “9” subjects could not be clearly distinguished from the others but in number “4” it shed some possibility of setting threshold values for speaker identification if the coefficients and the sum of absolute difference were employed. Further studies would be desirable on various combinations of the range of long-term average spectra and the degree of signal pre-emphasis. [Work supported by grant No. R01-1999-000-00229-0 from the Korea Science & Engineering Foundation.]

1p MON. PM

NOTE: Separate registration required. See page A24.

MONDAY EVENING, 28 APRIL 2003

RENAISSANCE, TENNESSEE ROOM, 7:00 TO 9:00 P.M.

Session 1eID

Interdisciplinary: Tutorial Lecture on Hearing in Three Dimensions

Beverly A. Wright, Chair

*Department of Communication Sciences and Disorders, Northwestern University, 2299 North Campus Drive,
Evanston, Illinois 60208-3550*

Chair's Introduction—7:00

7:05

1eID1. Hearing in three dimensions. Barbara Shinn-Cunningham (Boston Univ., Boston, MA 02215, shinn@bu.edu)

One of the key functions of hearing is to help us monitor and orient to events in our environment (including those outside the line of sight). The ability to compute the spatial location of a sound source is also important for detecting, identifying, and understanding the content of a sound source, especially in the presence of competing sources from other positions. Determining the spatial location of a sound source poses difficult computational challenges; however, we perform this complex task with proficiency, even in the presence of noise and reverberation. This tutorial will review the acoustic, psychoacoustic, and physiological processes underlying spatial auditory perception. First, the tutorial will examine how the many different features of the acoustic signals reaching a listener's ears provide cues for source direction and distance, both in anechoic and reverberant space. Then we will discuss psychophysical studies of three-dimensional sound localization in different environments and the basic neural mechanisms by which spatial auditory cues are extracted. Finally, "virtual reality" approaches for simulating sounds at different directions and distances under headphones will be reviewed. The tutorial will be structured to appeal to a diverse audience with interests in all fields of acoustics and will incorporate concepts from many areas, such as psychological and physiological acoustics, architectural acoustics, and signal processing.

Session 2aAA

Architectural Acoustics and Engineering Acoustics: Multi-Channel Sound Reinforcement Systems

Robert C. Coffeen, Chair

School of Architecture and Urban Design, University of Kansas, Marvin Hall, Lawrence, Kansas 66045

Chair's Introduction—8:00

Invited Papers

8:05

2aAA1. Systems for stereophonic sound reinforcement. Jim Brown (Audio Systems Group, Inc., 4875 N. Ravenswood, Chicago, IL 60640)

Although stereophonic loudspeaker systems for large spaces were developed in well-documented work at Bell Labs nearly 70 years ago, most contemporary designers appear to be ignorant of the critically important basic principles as applied to modern reinforcement systems. In fact, many designers don't believe stereo can even work in a room larger than a living room. This paper draws on both the literature and the author's experience over nearly 20 years with both permanent and portable systems using two and three front-referenced channels. Design criteria and examples of both good and problematic design practices are presented.

8:25

2aAA2. Providing audible information relating to the differences in monophonic and multichannel sound reinforcement systems using computer auralization. Robert C. Coffeen (School of Architecture & Urban Design, Marvin Hall, The Univ. of Kansas, Lawrence, KS 66045)

When selecting the sound distribution system for a particular venue, there is often a decision to be made as to the audible differences between a monophonic and potentially less expensive loudspeaker system and a multi-channel and potentially more expensive system. Computer auralization is a tool that can be used for auditioning the approximate differences between single channel and multi-channel loudspeaker systems.

8:45

2aAA3. Using 3-D modeling to design multichannel audio systems. Bruce C. Olson (Olson Sound Design, 8717 Humboldt Ave. N., Brooklyn Park, MN 55444, BCO@OlsonSound.com)

Implementation of multichannel sound systems requires a balancing of a number of different parameters and criteria to produce a cohesive sound field that is appropriate to the venue. Use of a 3-D modeling application allows for an iterative and interactive approach to optimizing the system. This design process will be explored and explained using some recently completed projects. Coordination with specific acoustical features of the room is inherently part of this process.

9:05

2aAA4. Microphone and production techniques for stereophonic sound reinforcement. Jim Brown (Audio Systems Group, Inc., 4875 N. Ravenswood, Chicago, IL 60640)

Producing a stereo image for many listeners in a large room is a far more demanding task than doing so for a few listeners in a living room, and very different acoustic systems of both microphones and loudspeakers are required for large room listening as compared to small room listening. As long ago as the 1930s, William B. Snow and his team at Bell Labs clearly understood these differences. Their thoughts on the subject are very different from those of Blumlein, and are well documented. Nearly all thinking and writing about microphone technique since Snow has focused on recording and playback for a single listener centered between two or three loudspeakers in a small room. Stereophonic reinforcement imposes a very different set of requirements, in that both listeners and performers are distributed over a wide area, and the system is a closed loop (that is, the potential for acoustic feedback must be considered).

Contributed Papers

9:25

2aAA5. Derivation of moving-coil loudspeaker parameters using acoustical testing techniques: Theoretical developments. Timothy W. Leishman and Brian E. Anderson (Dept. of Phys. and Astron., Brigham Young Univ., N335 Eyring Sci. Ctr., Provo, UT 84602, tim_leishman@byu.edu)

Moving-coil loudspeaker driver parameters are generally derived through the measurement of electrical impedances. Nevertheless, because these drivers are electro-mechano-acoustical transducers, their parameters may also be determined from measurements taken in other physical do-

main. This paper presents theoretical concepts that show how they may be determined acoustically using plane wave tube techniques. A driver is mounted in a baffle to form a composite partition between a source tube and a receiving tube. The frequency-dependent transmission loss of the partition is determined using upstream and downstream sound field decompositions that compensate for possible nonanechoic receiving tube conditions. A transmission loss curve based on an equivalent circuit model of the system is then fit to the measured curve to extract specific driver parameters. Different electrical conditions are imposed at the driver terminals to modify the transmission loss in ways that allow the determination of additional parameters.

2aAA6. Derivation of moving-coil loudspeaker parameters using acoustical testing techniques: Experiment results. Brian E. Anderson and Timothy W. Leishman (Dept. of Phys. & Astron., Brigham Young Univ., N281 ESC Brigham Young Univ., Provo, UT 84602, LoudspeakerDesign@hotmail.com)

A unique 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. Electrical conditions are easily controlled and automated to allow for the derivation of additional parameters. Current parameter measurement techniques require measurement of the electrical impedance of the DUT. This paper will discuss the acoustical measurement apparatus, experimental measurement techniques, and compare its measured parameters to those derived using electrical techniques.

9:55

2aAA7. Analysis of nonlinear behavior of loudspeakers using the instantaneous frequency. Hai Huang (College of Biomed. Eng. and Instrumentation Sci., Zhejiang Univ., Hangzhou 310027, PROC, hhai@zjuem.zju.edu.cn) and Finn Jacobsen (Tech. Univ. of Denmark, DK-2800 Lyngby, Denmark)

It is well known that the weakest link in a sound reproduction chain is the loudspeaker. The most significant effect on the sound quality is nonlinear distortion of loudspeakers. Many methods are applied to analyze the nonlinear distortion of loudspeakers. Almost all of the methods are based on the Fourier transform. In this work, a new method using the instantaneous frequency is introduced for describing and characterizing loudspeaker nonlinearities. First, numerical integration is applied to simulate the nonlinearities of loudspeakers caused by two nonlinear parameters, force factor and stiffness, separately. Then, a practical loudspeaker is used in an experiment and its nonlinear characteristics are analyzed with the instantaneous frequency. The results provide a clear physical interpretation of the nonlinearities of loudspeakers and will be useful for understanding the nonlinear behavior of loudspeakers. It will also be helpful for compensating for the nonlinearities and for improving the quality of sound reproduction. [Work supported by Sino-Danish International Co-operative Project, No. AM13: 66 and DANIDA (Denmark).]

10:10–10:25 Break

10:25

2aAA8. Real-time control of sound diffusion parameters. Colby Leider (School of Music, Univ. of Miami, 1314 Miller Dr., Coral Gables, FL 33124)

Much electronic and computer music relies extensively on real-time diffusion of electronic sound in surround loudspeaker configurations. The most common method of sound diffusion in practice today is to use standard sound-reinforcement mixers whereby each fader controls the playback volume of a corresponding loudspeaker in a concert hall. Using the mixer this way, however, presents problems when one attempts to create sound trajectories, because complex, precise, and repeatable control of the individual faders is required. To address these interface issues, two different handheld controllers were created for real-time sound diffusion. The first controller is equipped with accelerometers, force-sensing resistors, and joysticks and maps using input into diffusion parameters. The second controller mimics the operation of a standard mixer but allows more rapid movement of fader positions by replacing each fader with a bend sensor. To test the validity of each controller, twenty undergraduate music engineering students were asked to repeatedly perform various diffusion tasks (such as moving a monaural sound in a circle around the audience) using a standard mixer and each of the controllers. The accuracy and speed of their performances were tracked. The study concludes with lessons learned from the statistics gathered. [Research supported by the University of Miami.]

2aAA9. The virtual microphone technique in active sound field control systems. Iraklis E. Lampropoulos and Yasushi Shimizu (Prog. in Architectural Acoust., School of Architecture, Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180-3590, lampri@rpi.edu)

Active Sound Field Control (AFC) has been proven very useful in reverberation enhancement applications in large rooms. However, feedback control is required in order to eliminate peaks in the frequency response of the system. The present research closely follows the studies of Shimizu in AFC, in which smoothing of the rooms transfer function is achieved by averaging the impulse responses of multiple microphones. "The virtual or rotating microphone technique" reduces the number of microphones in the aforementioned AFC technology, while still achieving the same acoustical effects in the room. After the impulse responses at previously specified pairs of microphone positions are measured, the ratio of transfer functions for every pair is calculated, thus yielding a constant K . Next, microphones are removed and their impulse responses are reproduced by processing the incoming signal of each pair through a convolver, where the computed K constants have been previously stored. Band limiting, windowing and time variance effects are critical factors, in order to reduce incoherence effects and yield reliable approximations of inverse filters and consequently calculations of K . The project is implemented in a church lacking low frequency reverberation for music and makes use of 2 physical and 2 virtual microphones.

10:55

2aAA10. Analysis of multiple listener equalization performance due to displacement effects. Sunil Bharitkar, Philip Hilmes, and Chris Kyriakakis (Immersive Audio Lab., EEB 428, Integrated Media Systems Ctr., Univ. of Southern California, Los Angeles, CA 90089, bharitka@usc.edu)

Traditionally, room response equalization is performed to enhance sound quality, by reducing the effects of reverberation, at a given listener. However, room responses vary with source and listener positions. Hence, in a multiple listener environment, equalization may be performed through averaging the room responses measured at multiple listener locations. However, the performance of averaging based equalization, at the listeners, may be affected when listener positions change or due to microphone-listener position mismatch (i.e., displacement effects). In this paper, we present a statistical approach to map variations in listener positions to equalization performance of spatial average methods. The analysis is done at frequencies above the Schroeder frequency where the direct and the reverberant sound fields are uncorrelated, and the results are presented in a 3-D plot to clearly show the changes in equalization performance (in dB's) versus mismatch and frequencies for various listener configurations relative to a fixed loudspeaker source. The results indicate that, for the analyzed listener configurations, the zone of equalization depends on distance of microphones/listeners from the source and the frequencies in the sound. [Work supported by the U.S. Dept. of Army.]

11:10

2aAA11. Open-loop dereverberation of multichannel room impulse responses. Bowon Lee, Mark A. Hasegawa-Johnson (Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana-Champaign, 1406 W. Green St., Urbana, IL 61801, bowonlee@uiuc.edu), and Camille Goudeseune (Integrated Systems Lab., Beckman Inst., Urbana, IL 61801)

We are developing the audio display for a CAVE-type virtual reality theater, a 3-m cube with displays covering all six rigid faces. The user's headgear continuously reports ear positions so headphones would be possible, but we nevertheless prefer loudspeakers because this enhances the sense of total immersion. Because sounds produced at the loudspeakers are distorted by the room impulse responses, we therefore face the problem of controlling the sound at the listener's two ears. Our proposed solution consists of open-loop acoustic point control, i.e., dereverberation. The room impulse responses from each loudspeaker to each ear of the listener are inverted using multichannel inversion methods, to create exactly the desired sound field at the listener's ears. Because the actual room impulse responses cannot be measured in real time (as the listener walks

around), instead the impulse responses simulated by the image-source method is used. A new evaluation criterion is proposed to quantitatively evaluate both the simulation and the open-loop dereverberation. The actual impulse responses used for this evaluation are measured with a starter pistol, since this best approximates the point source assumed by the image-source method.

11:25

2aAA12. Multichannel sound reinforcement systems at work in a learning environment. John Malek and Colin Campbell (10489 E. Grand River Ste. I, Brighton, MI 48116, jmalek@annarbaud.com)

Many people have experienced the entertaining benefits of a surround sound system, either in their own home or in a movie theater, but another application exists for multichannel sound that has for the most part gone unused. This is the application of multichannel sound systems to the learning environment. By incorporating a 7.1 surround processor and a touch panel interface programmable control system, the main lecture hall at the University of Michigan Taubman College of Architecture and Urban Planning has been converted from an ordinary lecture hall to a working audio-visual laboratory. The multichannel sound system is used in a wide variety of experiments, including exposure to sounds to test listeners' aural perception of the tonal characteristics of varying pitch, reverberation, speech transmission index, and sound-pressure level. The touch panel's custom interface allows a variety of user groups to control different parts of the AV system and provides preset capability that allows for numerous system configurations.

11:40

2aAA13. Virtual AM stereo and surround sound to setup AM/FM radio theatre. Selvakumaran Vadivelmurugan, K. Veeraraghavan (Dept. of I.T., Sri Venkateswara College of Eng., Anna Univ., Pennalur, Sriperumbudur, Tamil Nadu, India 602105, vselvakumaran@vselvakumaran.com), and Sharavan V. Govindan (Univ. of Connecticut, Storrs, CT 06269)

Introduction of virtual surround sound and stereo to AM radio has been proposed in this study. This technology can be further applied to aid the construction of an AM radio theatre. Adding to the advantages of AM, the lower bandwidth, higher range, and simpler circuitry, AM can now offer excellent sound effect with the post-transmission process. The motivation for the introduction of virtual surround sound is the poor quality of AM sound. In this study, the response by human ear has been thoroughly investigated and the methodology to create virtual surround sound has been developed. The elements essential to setup audio theatre such as the components of audio chain, multiple unit audio speaker, inner section of the ear, psychological effect of different ranges of frequencies, and radio theatre design have been extensively studied on the basis of Helmholtz audition theory. The vital changes include the different frequency division multiplexing of message at the transmitting end, three phases of the process, resulting in the vertical and horizontal digital connection, espresso program, and the 3×12 speaker design theatre.

2a TUE. AM

TUESDAY MORNING, 29 APRIL 2003

ROOMS 103/104, 8:00 TO 10:00 A.M.

Session 2aAO

Acoustical Oceanography, Underwater Acoustics and Signal Processing in Acoustics: Geoacoustic Inversion II

N. Ross Chapman, Chair

School of Earth and Ocean Sciences, University of Victoria, P.O. Box 3055, Victoria, British Columbia V8W 3P6, Canada

Invited Papers

8:00

2aAO1. Modal inverse methods: An overview. Kyle M. Becker (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804-0030) and George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Geoacoustic inversion using modal inverse methods is based on a very simple experimental geometry. Data are acquired from synthetic aperture measurements of a point source acoustic field using a single hydrophone translating through the water column at constant depth. The modal composition of the propagating field is extracted by exploiting the Hankel transform relationship between the acoustic field, measured as a function of range, and the horizontal wavenumber spectrum for horizontally stratified media. Modal content, estimated as discrete values of the horizontal wavenumber, are used as input data for a general linear inverse problem, where small perturbations to a background sound speed profile are related to changes in individual horizontal wavenumbers. An overview of the method is presented with an emphasis placed on the application to range-dependent shallow water environments. Inversion results are presented from a recent workshop on geoacoustic inversion for range-dependent environments, as well as from recent at-sea measurements. [Work supported by ONR.]

8:20

2aAO2. Estimating the low-frequency (0.1–1 kHz) sound speed in marine sediments using the harmonics from the propeller of a light aircraft. Michael J. Buckingham and Eric M. Giddens (Marine Physical Lab., Scripps Inst. of Oceanogr., 8820 Shellback Way, La Jolla, CA 92093-0238)

During ONR's Sediment Acoustics Experiment 1999 (SAX99) in the northeastern Gulf of Mexico, several research groups made high-precision, *in situ* measurements of dispersion in the medium-sand sediment at frequencies greater than 20 kHz. Comparable precision at lower frequencies is difficult to achieve with *in situ* time-of-flight techniques because of wavelength issues which, *inter alia*, dictate an inconveniently large and costly acoustic source. Yet low-frequency (1 kHz) sound speed measurements are sorely needed to distinguish between competing theoretical predictions. An alternative to the traditional travel-time approach employs a single hydrophone buried in the sediment and, instead of an *in situ* sound source, the low-frequency harmonics from the propeller of

a light aircraft. Essentially, the airborne-source technique relies on the difference between the Doppler-shifted frequencies on aircraft approach and departure, as detected on the buried hydrophone, to yield a direct measure of the local sound speed at the detector. Experiments recently conducted about 1.5 km north of Scripps pier, using a single-engine Tobago TB10 aircraft with a two-blade propeller, will be described and the resultant estimates of the low-frequency sound speed in the fine-sand sediment at the site will be presented. [Work supported by ONR.]

8:40

2aAO3. Inversion of sediment property using ambient noise. Dajun Tang (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Recently, there have been a couple of groups started using ambient noise and a vertical array to invert for bottom geo-acoustic properties. The basic idea is to select beam outputs from the array of the noise field which is generated near the ocean surface. The ratio of the two beams with angles symmetrical to the horizontal gives the reflection loss of the bottom. The reflection loss can be obtained in large and small grazing angles and over a frequency band limited only by the array length and spacing. In this paper, both modeling and data are presented to demonstrate the applicability and limitations of the approach. Emphasis will be on the resolution and uncertainty of the method when the array is of finite length and has a tilt angle relative to the vertical. The noise field will be generated using numerical methods based on Kraken. Wind as well as shipping noise is considered.

Contributed Papers

9:00

2aAO4. Seabottom geoacoustic inversion from reverberation vertical coherence in shallow water. Ji-Xun Zhou and Xue-Zhen Zhang (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405 and Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, jixun.zhou@me.gatech.edu)

Shallow-water reverberation from one shot offers a continuous spatial sampling of the surrounding sound field up to several tens of kilometers from the source. It involves a bottom-controlled two-way sound propagation and bottom scattering process, and brings back rich information on seabottom geoacoustic parameters. Thus, geoacoustic inversion from reverberation in shallow water is very attractive for saving time and cost compared with inversion from propagation measurements. In this paper, the reverberation vertical coherence in shallow water, derived by ray-mode analogies [Zhou, *Acta Oceanol. Sinica* **1**, 212–218 (1979)], is converted back to a more familiar summation of normal-modes. Measured reverberation cross-correlation coefficients as a function of time, frequency and hydrophone separation at different areas with a flat seabed are in good agreement with theoretical predictions. From the best match between the measurements and predictions of the reverberation vertical coherence, sound velocities and attenuations in sediments from China seas, including the ASIAEX01 site, have been inverted at low and mid-frequencies. [Work supported by ONR and CAS.]

9:15

2aAO5. Source depth and array tilt effects on seabed inversion of ambient noise. Francine Desharnais, David J. Thomson (DRDC Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada), and Chris A. Gillard (Defence Sci. and Technol. Organisation, Salisbury SA 5108, Australia)

Ambient noise coherence between two vertically separated sensors in shallow water relates to the directionality of the noise field and is sensitive to the reflective properties of the surficial sediments. In earlier work, an energy-flux model was developed to calculate the noise coherence over a multilayered seabed; in this model the surface sources were assigned a frequency-independent dipole directivity pattern. The model was subsequently combined with a hybrid nonlinear inversion procedure to effectively search the space of geoacoustic properties that parametrize a multilayered seabed. In order to benchmark the energy-flux model against a wave-theoretical formulation, the issue of the unknown but finite source depth of the noise sources must be addressed. These near-surface sources result in a frequency-dependent beam pattern that is not dipolar above a few hundred hertz [M. J. Buckingham and N. M. Carbone, *J. Acoust. Soc. Am.* **102**, 2637–2644 (1997)]. In this paper, we extend the energy-flux model to include the effect of finite source depth as well as sensor tilt on the observed directivity of the ambient noise field. These extra degrees of freedom are added to the search space of the inversion procedure and their influence on estimating seabed properties from noise coherence data is investigated.

9:30

2aAO6. Depth partitioning of modal energy of guided acoustic waves in shallow water as an additional input for geoacoustic inversion. Allan D. Pierce, George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Luiz L. Souza (MIT, Cambridge, MA 02139)

Modal mapping experiments (MOMAX) previously described by Souza, Frisk, and Becker [*J. Acoust. Soc. Am.* **112**, 2281 (2001)] employ a CW-source at a fixed water depth and a receiver also at a fixed depth, with the range between the two varying systematically. The received amplitude versus range over a limited range interval yields a k -space spectrum, where the individual peaks correspond to the eigen-wavenumbers of the individual guided modes. This extraction in conjunction with a knowledge of the sound speed profile within the water column yields discrete values of the impedance at the water-bottom interface for the corresponding frequency and horizontal wavenumber, and these values in turn can be used in an approximate geoacoustic inversion for the bottom properties. The areas under these peaks are shown to be especially robust and can be given a theoretical interpretation with reference to the modal sum solution for a point source in shallow water with weak range dependence. This yields the ratio of the modal energies in the water column to those in the bottom. Both types of information are combined into a theory that yields an improved geoacoustic inversion.

9:45

2aAO7. Autoregressive wave number estimation technique for range-dependent shallow-water waveguides with abrupt environmental variations. Luiz L. Souza (MIT/WHOI Joint Prog. in Oceanogr./Appl. Ocean Sci. and Eng., 77 Massachusetts Ave., Rm. 5-435, Cambridge, MA 02139) and George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

We extend the technique described by Becker and Frisk [*J. Acoust. Soc. Am.* **111**, 2388 (2002)], which uses an autoregressive spectral (AR) estimator to infer characteristic modal wave numbers in shallow-water, range-dependent waveguides. Their method offers high resolution over relatively small range apertures, an improvement over the short-range Fourier transform technique originally employed. When the environment varies continuously, e.g., due to a sloping bottom, the assumption of constant modal eigenvalues over a range aperture, intrinsic to the AR estimator, becomes invalid, resulting in increased bias. Abrupt changes in the modal content of the field can be detected, but not measured. The present work extends this method by allowing the modal structure to change with range on a sample by sample basis, and yields AR coefficients that are functions of range, analogous to the time-varying AR (TVAR) estimators used for nonstationary time series analysis. By processing the data twice, with increasing and decreasing ranges (forward and backwards), and combining the results, both slow and abrupt changes in the eigenvalues can be measured.

Session 2aBB**Biomedical Ultrasound/Bioresponse to Vibration and ASA Committee on Archives and History:
History of Bioresponse to Vibration in the Acoustical Society of America**

Stanley J. Bolanowski, Chair

*Institute for Sensory Research, Syracuse University, Syracuse, New York 13244***Chair's Introduction—11:00*****Invited Paper*****11:05****2aBB1. History of bioresponse to vibration in the Acoustical Society of America.** Janet M. Weisenberger (Speech & Hearing Sci., Ohio State Univ., Columbus, OH 43210, jan+@osu.edu)

Human response to vibratory stimulation of the skin surface has long been considered an aspect of the sense of touch; however, the debate over whether vibration was one aspect of pressure sensation, as espoused by von Frey in the late 1800s, or a separate sense, as argued by Katz (1925), focused attention on this mode of stimulation. Experimental investigations from the 1920s to the 1960s by Knudsen, Geldard, Sherrick, Verrillo, Mountcastle, and others provided basic data on vibrotactile perception and the neural transduction of vibratory stimulation. Within the Acoustical Society of America, work on bioresponse to vibration has included not only basic investigations of vibrotactile perception and physiology, but also studies of the loss of sensitivity resulting from intense hand-arm vibration induced by occupational use of chainsaws and jackhammers, studies of human response to whole-body vibration, and evaluations of the utility of vibrotactile devices for communication of speech to hearing-impaired persons. Contributions in each of these areas, as well as future research directions, are discussed.

Session 2aPA**Physical Acoustics: Sono(con)-fusion I: Evaluating the Chances and Claims of Bubble Fusion**

D. Felipe Gaitan, Cochair

Impulse Devices, Inc., 12731-A Loma Rica Drive, Grass Valley, California 95945

R. Glynn Holt, Cochair

*Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street,
Boston, Massachusetts 02215*

Thomas J. Matula, Cochair

*Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105***Chair's Introduction—7:25*****Invited Papers*****7:30****2aPA1. Update and clarifications on analytic studies for nuclear emissions during acoustic cavitation.** Robert I. Nigmatulin (Russian Acad. of Sci., Russia), R. T. Lahey (Rensselaer Polytechnic Inst., Troy, NY), R. P. Taleyarkhan, and C. D. West (Oak Ridge Natl. Lab., Oak Ridge, TN 37831)

A one-dimensional hydrodynamic shock (HYDRO) code was developed to numerically evaluate the conservation equations of each phase during bubble growth and collapse. This code includes the Mie-Grüniesen equations of state and Born-Mayer potential functions, which are known to be valid for highly compressed fluids. In particular, for acetone these equations of state are based on

the shock wave adiabat data, and they implicitly specify the effect of the induced radiation field and the dissociation and ionization processes that take place during plasma formation within imploding bubbles. Moreover, relevant energy losses and the effect of both molecular and electron conductivity were taken into account, and the resultant HYDRO code allowed for the evaluation of shock wave interaction using the well-established Godunov numerical technique. Bubble dynamics were studied in deuterated acetone for conditions typical of those in our experiments. It was found that highly compressed conditions suitable for thermonuclear fusion were predicted, and the results were sensitive to the values of the phase change (that is, accommodation) coefficient, a , and the liquid pool temperature T_o .

8:15

2aPA2. Why seek fusion from cavitation: Molecular dynamic simulations and a detector capable of time correlated single neutron counting. Carlos Camara, Robert Cousins, Brian Naranjo, Seth Putterman (Phys. Dept., UCLA, Los Angeles, CA 90095), Barry Merriman (UCLA, Los Angeles, CA 90095), and Steven Ruuth (Simon Fraser Univ., Burnaby, BC V5A 1S6, Canada)

The blackbody spectra, and similar sonoluminescence intensities of He and Xe bubbles suggest that the interior of a sonoluminescing bubble is highly stressed and dense. Molecular dynamic simulations indicate interior temperatures which are enhanced by thermal conduction and can approach 1 MK. Furthermore the gas passes through states where the mean free path is larger than the distance over which temperature varies and so calls into question the value of theories based on hydrodynamics. To search for rare fusion events a neutron detector with 25% total discriminated quantum efficiency has been built. It can time stamp neutron arrival and sonoluminescence to better than 1 ns and record tracks on the fly. [Work supported by DARPA.]

9:00–9:15 Break

9:15

2aPA3. Acoustically driven spherical implosions and the possibility of thermonuclear reactions. D. Felipe Gaitan and Ross Tessien (Impulse Devices, 12731-A Loma Rica Dr., Grass Valley, CA 95945)

Acoustically driven, gas-filled cavities in liquids have been known to collapse violently, generating short flashes of light of ~ 100 -ps duration. More recently, the possibility of generating fusion reactions using acoustics (acoustic inertial confinement fusion) has been considered. Results of computer simulations using the HYADES hydrocode (Cascade Applied Sciences, Inc) plus the SESAME equations of state for free collapsing and acoustically driven cavities in molten metals will be presented as well as experimental data at high ambient pressures in different liquids. 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 acoustic ICF as an energy source.

10:00

2aPA4. Chemical control of single bubble cavitation. Yuri T. Didenko and Kenneth S. Suslick (Univ. of Illinois at Urbana-Champaign, 600 S. Mathews Ave., Urbana, IL 61801, ksuslick@uiuc.edu)

Sonochemistry would be ideally studied with a single bubble with known size pulsating in known acoustic pressure field. Single bubble cavitation provides the means to make such studies. The promise that single bubble cavitation brought to the quantitative measurements of chemical activity of cavitation, however, has not been previously fulfilled due to the very small amount of reacting gas within a single bubble (typically <10 – 13 moles). We have now quantitated the chemical reactions inside a single cavitating bubble and established an inventory of energy dissipation during bubble collapse. The yields of nitrite ions, hydroxyl radicals, and photons from a single cavitation bubble have now been measured, and the first true measures of energy efficiencies during acoustic cavitation have been determined. The energy efficiency of the formation of hydroxyl radicals from single bubble is comparable to that in multibubble cavitation. The energy efficiency of light emission, however, is much higher for SBSL. The observed rate of nitrite formation is in good agreement with the calculated diffusion rate of nitrogen into the bubble. The temperatures reached during single bubble cavitation in liquids with significant vapor pressures will be substantially limited by the endothermic chemical reactions of the polyatomics inside the collapsing bubble.

10:45

2aPA5. Litho-Fusion and HIFusion: Alternative ways to generate hot bubbles? Thomas Matula, Paul Hilmo, Michael Bailey, and Lawrence Crum (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

The recent sonofusion experiments at Oak Ridge National Laboratory utilized standing acoustic wave fields that produced negative pressures of about 15 atm. The generated bubbles were thought to grow from around 100 nm to a maximum size of about 1 mm, before collapsing violently. There are other means for generating bubbles that grow to millimeter sizes. For example, in lithotripsy, a focused shock wave ($\text{Pa} \sim -100$ atm) creates bubbles that grow to millimeter sizes. High-intensity focused ultrasound (HIFU) pulses can also generate large bubbles. Can these systems be used to investigate extreme bubble collapse violence? We will describe our efforts in generating and observing bubbles from these systems. If time permits, we will also point out that bubbles in a standing wave undergo translational motion, and this motion may lead to instabilities.

11:30–12:30

Panel Discussion

Session 2aPP

Psychological and Physiological Acoustics: Developmental Psychoacoustics

Lynne A. Werner, Cochair

Department of Speech and Hearing, University of Washington, 1417 NE 42nd Street, Seattle, Washington 98105-6246

Prudence Allen, Cochair

*Department of Communication Disorders, University of Western Ontario, 1500 Elborn College,
London, Ontario N6G 1H1, Canada*

Chair's Introduction—8:30

Invited Papers

8:35

2aPP1. Development of infants' pitch perception. Marsha G. Clarkson (Dept. of Psych., Georgia State Univ., Atlanta, GA 30303)

Several studies demonstrate that by 7 months of age infants can hear the pitch of the missing fundamental for tonal complexes and that the same acoustic cues used by adults permit infants to perceive pitch. Infants and adults were tested in an operant conditioning procedure and learned to categorize tonal complexes with differing spectra according to the pitch of the missing fundamental. Responses to spectral manipulations that reduce the salience of pitch suggested that the strength of the pitch percept may be weaker and the use of temporal cues may be poorer for infants than for adults. Other work utilizing an iterated rippled noise stimulus (IRN) attempted to quantify these developmental differences. Applying progressively greater attenuation to the delayed noise in an IRN progressively decreases the strength of the pitch percept evoked by the stimulus. The maximum attenuation for which listeners detect changes in pitch estimates the strength of their pitch percept. Such thresholds revealed that the strength of the pitch percept evoked by IRN was significantly weaker for infants (6.7 dB) than for adults (19.1 dB). In combination, these findings suggest that temporal mechanisms for processing pitch may be particularly immature in infants.

9:00

2aPP2. Speech intelligibility in complex acoustic environments in young children. Ruth Litovsky (Waisman Ctr., Univ. of Wisconsin—Madison, Madison, WI 53705)

While the auditory system undergoes tremendous maturation during the first few years of life, it has become clear that in complex scenarios when multiple sounds occur and when echoes are present, children's performance is significantly worse than their adult counterparts. The ability of children (3–7 years of age) to understand speech in a simulated multi-talker environment and to benefit from spatial separation of the target and competing sounds was investigated. In these studies, competing sources vary in number, location, and content (speech, modulated or unmodulated speech-shaped noise and time-reversed speech). The acoustic spaces were also varied in size and amount of reverberation. Finally, children with chronic otitis media who received binaural training were tested pre- and post-training on a subset of conditions. Results indicated the following. (1) Children experienced significantly more masking than adults, even in the simplest conditions tested. (2) When the target and competing sounds were spatially separated speech intelligibility improved, but the amount varied with age, type of competing sound, and number of competitors. (3) In a large reverberant classroom there was no benefit of spatial separation. (4) Binaural training improved speech intelligibility performance in children with otitis media. Future work includes similar studies in children with unilateral and bilateral cochlear implants. [Work supported by NIDCD, DRF, and NOHR.]

9:25

2aPP3. Developmental MLD effects. Emily Buss, Joseph W. Hall III, and John H. Grose (Univ. of North Carolina at Chapel Hill, Chapel Hill, NC 27514)

Whereas the masking-level difference (MLD) in children is similar to that found in adults for wideband masking noise, children often have smaller MLDs for narrowband noise maskers. One possibility is that the small MLD of children in narrowband noise is related to the pronounced amplitude fluctuation of the masker. In the present study, adults and children (aged 5.1 to 10.7 years) were tested in an MLD paradigm in which the detection of brief signals was contrasted for signal placement in masker envelope maxima versus masker envelope minima. Maskers were 50-Hz wide noise bands centered on 500 Hz, and the signals were S_0 or $S_{\langle \pi \rangle}$ 30-ms, 500-Hz tones. In agreement with previous studies, it was found that MLDs were greater for masker envelope minima placement than for masker envelope maxima placement. This effect was due to variation across the $S_{\langle \pi \rangle}$ conditions. The results indicated that young children are relatively poor in taking advantage of the relatively good signal-to-noise ratios in masker minima when detecting $S_{\langle \pi \rangle}$ signals in No maskers. This finding is consistent with the possibility that binaural temporal resolution continues to develop through the first decade of life.

2aPP4. Learning problems, delayed perceptual development, and puberty. Beverly A. Wright, Steven G. Zecker, and Miriam D. Reid (Dept. of Commun. Sci. and Disord. and Inst. for Neurosci., 2240 Campus Dr., Northwestern Univ., Evanston, IL 60208-3550)

Language-based learning problems affect approximately one person in twelve with no other obvious signs of disorder. Many of these individuals have accompanying deficits in nonlinguistic perception. To determine whether age influences the magnitude of these deficits, thresholds on a set of auditory masking tasks were measured in individuals with learning problems and controls ranging in age from 6 years to adult. Performance improved with increasing age in both groups. However, the thresholds of the individuals with learning problems were most similar to those of controls approximately 2–4 years younger on every task, suggesting that the perceptual development of the affected individuals was delayed by a constant amount. Further, on the subset of conditions on which controls reached adult levels of performance after 10 years of age, the improvement of affected individuals halted at 10 years of age, suggesting that puberty may play a critical role in human perceptual development. Taken together, these data support the idea that some learning problems result from a neuromaturational delay, of unknown breadth, and indicate that neurological changes associated with puberty prevent the complete resolution of delayed perceptual development. [Work supported by NIH/NIDCD.]

10:15–10:30 Break

10:30

2aPP5. Variability and reduced performance of preschool- and early school-aged children on psychoacoustic tasks: What are the relevant factors? Prudence Allen (Nat'l. Ctr. for Audiol., Univ. of Western Ontario, 2262 Elborn College, London, ON N6G 1H1, Canada)

Young children typically perform more poorly on psychoacoustic tasks than do adults, with large individual differences. When performance is averaged across children within age groups, the data suggest a gradual change in performance with increasing age. However, an examination of individual data suggests that the performance matures more rapidly, although at different times for different children. The mechanisms of development responsible for these changes are likely very complex, involving both sensory and cognitive processes. This paper will discuss some previously suggested mechanisms including attention and cue weighting, as well as possibilities suggested from more recent studies in which learning effects were examined. In one task, a simple frequency discrimination was required, while in another the listener was required to extract regularities in complex sequences of sounds that varied from trial to trial. Results suggested that the ability to select and consistently employ an effective listening strategy was especially important in the performance of the more complex task, while simple stimulus exposure and motivation contributed to the simpler task. These factors are important for understanding the perceptual development and for the subsequent application of psychoacoustic findings to clinical populations. [Work supported by the NSERC and the Canadian Language and Literacy Research Network.]

10:55

2aPP6. Applications of psychophysical models to the study of auditory development. Lynne Werner (Dept. of Speech & Hearing Sci., Univ. of Washington, Seattle, WA 98105-6246)

Psychophysical models of listening, such as the energy detector model, have provided a framework from which to characterize the function of the mature auditory system and to explore how mature listeners make use of auditory information in sound identification. The application of such models to the study of auditory development has similarly provided insight into the characteristics of infant hearing and listening. Infants' intensity, frequency, temporal and spatial resolution have been described at least grossly and some contributions of immature listening strategies to infant hearing have been identified. Infants' psychoacoustic performance is typically poorer than adults under identical stimulus conditions. However, the infant's performance typically varies with stimulus condition in a way that is qualitatively similar to the adult's performance. In some cases, though, infants perform in a qualitatively different way from adults in psychoacoustic experiments. Further, recent psychoacoustic studies of children suggest that the classic models of listening may be inadequate to describe the children's performance. The characteristics of a model that might be appropriate for the immature listener will be outlined and the implications for models of mature listening will be discussed. [Work supported by NIH grants DC00396 and by DC04661.]

Contributed Papers

11:20

2aPP7. Infants' detection in the presence of masker uncertainty. Lori J. Leibold and Lynne A. Werner (Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98105-6246, ljl@u.washington.edu)

Most adults have difficulty detecting a fixed-frequency pure tone when the frequency content of the masker is varied on each presentation. This informational masking is thought to be the result of masker uncertainty. The purpose of the present study was to examine infant thresholds under conditions of masker uncertainty. The subjects were 7- to 9-month-old infants and 18- to 30-year-old adults with no risk factors for hearing loss. Detection thresholds were measured for a 300-ms, 1-kHz tone in the presence of a noise band (300–3000 Hz), a two-component constant-frequency masker, or a two-component random-frequency masker. Maskers repeated at 300-ms intervals throughout the testing session at an overall level of 60 dB SPL. The signal was played synchronously with one interval of the repeating masker. An observer-based testing method was used. Following

training to 80%-correct criterion, detection thresholds were determined adaptively. For adults, the random-frequency masker condition produced the greatest amount of masking, ranging from 30–60 dB. For infants, the random-frequency masker also produced the greatest amount of masking, ranging from 60–80 dB. In contrast to the adults, infant thresholds were also elevated in the constant-frequency masker condition, ranging from 35–55 dB across subjects. [Work supported by NIDCD RO1 DC000396 and NRSA DC006122.]

11:35

2aPP8. A cocktail-party listening experiment with children. Frederic Wightman, Michael Callahan, and Doris Kistler (Waisman Ctr., 1500 Highland Ave., Madison, WI 53705)

In an experiment modeled after one reported recently by Brungart and Simpson [J. Acoust. Soc. Am. **112**, 2985–2995 (2002)], 38 children (ages 4–16) and 10 adults responded to a monaural target speech signal in the

presence of one or two distracter speech signals. The target speaker was a male and the distracter speakers were females. When two distracters were present they were in different ears. Performance at several different target ear S/N was measured and psychometric functions were fitted to estimate threshold, or the 50% performance level. The youngest children required approximately 20 dB higher S/N than adults to achieve threshold with a single distracter. This difference disappeared by age 16. The impact of adding the contralateral distracter, which is thought to contribute only

informational masking, was roughly constant across age, however. Adult thresholds increased about 11 dB and the thresholds for the youngest children increased about 10 dB. This was surprising given previous experiments that showed much larger informational masking effects in young children. Also inconsistent with previous results is the lack of individual differences. Nearly all listeners showed almost the same contralateral distracter effect. [Work supported by NICHD.]

TUESDAY MORNING, 29 APRIL 2003

ROOM 205, 8:30 TO 11:45 A.M.

Session 2aSC

Speech Communication: Neural Models for Speech Perception

Robert F. Port, Cochair

Department of Linguistics, Indiana University, 330 Memorial Hall, Bloomington, Indiana 47405

Frank H. Guenther, Cochair

Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, Massachusetts 02215

Invited Papers

8:30

2aSC1. Introductory remarks on neural modeling in speech perception research. Frank H. Guenther (Dept. of Cognit. and Neural Systems, Boston Univ., Boston, MA 02215, guenther@bu.edu)

This talk will provide an overview of several neural models that have been used to address experimental data on speech perception, including psychophysical and neurophysiological data. Although these models differ greatly in their scope and mathematical detail, all are geared toward providing a better understanding of the neural mechanisms underlying the processing of speech sounds.

8:50

2aSC2. Resonant cortical dynamics of speech perception. Stephen Grossberg (CNS Dept., Boston Univ., 677 Beacon St., Boston, MA 02215)

What is the neural representation of a speech code as it evolves in time? How do listeners integrate temporally distributed phonemic information into coherent representations of syllables and words? How does the brain extract invariant properties of variable-rate speech? This talk describes a neural model that suggests answers to these questions, while quantitatively simulating speech and word recognition data. The conscious speech and word recognition code is suggested to be a resonant wave, and a percept of silence a temporal discontinuity in the rate that resonance evolves. A resonant wave emerges when sequential activation and storage of phonemic items in working memory provides bottom-up input to list chunks that group together sequences of items of variable length. The list chunks compete and winning chunks activate top-down expectations that amplify and focus attention on consistent working memory items, while suppressing inconsistent ones. The ensuing resonance boosts activation levels of selected items and chunks. Because resonance occurs after working memory activation, it can incorporate information presented after intervening silence intervals, so future sounds can influence how we hear past sounds. The model suggests that resonant dynamics enable the brain to learn quickly without suffering catastrophic forgetting, as described within Adaptive Resonance Theory.

9:20

2aSC3. Using neural modeling and functional neuroimaging to study the neural basis of auditory object processing. Barry Horwitz and Fatima T. Husain (Brain Imaging & Modeling Section, NIDCD, Natl. Institutes of Health, Bethesda, MD 20892, horwitz@helix.nih.gov)

The neural basis of auditory object processing in the human cerebral cortex was investigated by combining neural modeling and functional neuroimaging. We developed a large-scale, neurobiologically realistic network model of auditory pattern recognition that relates neuronal dynamics of cortical auditory processing of frequency-modulated (FM) sweeps to functional neuroimaging data obtained using functional magnetic resonance imaging (fMRI). FM sweeps are ubiquitous in animal communication. Areas included in the model extend from primary auditory to prefrontal cortex. The electrical activities of the model neuronal units were constrained to agree with data from the neurophysiological literature regarding FM sweep perception. A fMRI experiment using stimuli and tasks similar to those used in our simulations was performed. The regional integrated synaptic activities of the model were used to determine simulated regional fMRI activities, and generally agreed with the experimentally observed fMRI data. Our results demon-

2a TUE. AM

strate that the model is capable of exhibiting the salient features of both electrophysiological neuronal activities and fMRI values that are in agreement with empirically observed data. These findings provide support for our hypotheses concerning how auditory objects are processed by primate neocortex. This type of approach offers the potential for understanding the neural basis of human speech perception.

9:50–10:05 Break

10:05

2aSC4. Modeling the representation of speech sounds in auditory cortical areas. Frank H. Guenther, Jason A. Tourville, and Jason W. Bohland (Dept. of Cognit. and Neural Systems, Boston Univ., Boston, MA 02215, guenther@bu.edu)

Many studies have shown that sounds from near the center of a sound category (such as a phoneme from one's native language) are more difficult to discriminate from each other than sounds from near a category boundary. However, the neural processes underlying this phenomenon are not yet clearly understood. In this talk we describe neural models that have been developed to address experimental data from psychophysical and functional brain imaging experiments investigating sound representations in the cortex. Experiments investigating the effects of categorization and discrimination training with nonspeech sounds indicate that different training tasks have different effects on sound discriminability: discrimination training increases the discriminability of the training sounds, whereas learning a new sound category decreases the discriminability of the training sounds within the category. These results can be accounted for by a neural model in which categorization training causes a decrease in the size of the cortical representation of central sounds in the category, while discrimination training leads to an increase in the cortical representation of training sounds. This model is supported by brain imaging results for speech and nonspeech sounds. Experimental results further suggest preferential utilization of different auditory cortical regions when subjects perform identification versus discrimination tasks.

10:35

2aSC5. The functional anatomy of speech perception: Dorsal and ventral processing pathways. Gregory Hickok (Dept. of Cognit. Sci., Univ. of Calif., Irvine, CA 92697)

Drawing on recent developments in the cortical organization of vision, and on data from a variety of sources, Hickok and Poeppel (2000) have proposed a new model of the functional anatomy of speech perception. The model posits that early cortical stages of speech perception involve auditory fields in the superior temporal gyrus bilaterally (although asymmetrically). This cortical processing system then diverges into two broad processing streams, a ventral stream, involved in mapping sound onto meaning, and a dorsal stream, involved in mapping sound onto articulatory-based representations. The ventral stream projects ventrolaterally toward inferior posterior temporal cortex which serves as an interface between sound and meaning. The dorsal stream projects dorsoposteriorly toward the parietal lobe and ultimately to frontal regions. This network provides a mechanism for the development and maintenance of "parity" between auditory and motor representations of speech. Although the dorsal stream represents a tight connection between speech perception and speech production, it is not a critical component of the speech perception process under ecologically natural listening conditions. Some degree of bi-directionality in both the dorsal and ventral pathways is also proposed. A variety of recent empirical tests of this model have provided further support for the proposal.

11:05

2aSC6. Is there still a TRACE of trace? James McClelland, Daniel Mirman, and Lori Holt (Dept. of Psych. and Ctr. for Neural Basis of Cognition, Carnegie Mellon Univ., Pittsburgh, PA 15213)

According to the TRACE model [McClelland and Elman, *Cogn. Psychol.* **18**, 1–86 (1986)], speech recognition is an interactive activation process involving the integrated use of top-down (lexical) and bottom-up (acoustic) information. Although it is widely accepted that there are lexical influences on speech perception, there has been a disagreement over their exact nature. Two contested predictions of TRACE are that (a) lexical influences should delay or inhibit recognition of phonemes not consistent with lexical information and (b) a lexical influence on the identification of one phoneme can trigger compensation for co-articulation, affecting the identification of other phonemes. Others [Norris, McQueen, and Cutler, *BBS* **23**, 299–370 (2000)] have argued that the predicted effects do not occur, taking this to support an alternative to the TRACE model in which lexical influences do not affect perception, but only a post-perceptual identification process. We re-examine the evidence on these points along with the recent finding that lexical information may lead to a lasting adjustment of category boundaries [McQueen, Norris, and Cutler, *Psychonomics Abstract* **255** (2001)]. Our analysis indicates that the existing evidence is completely consistent with TRACE, and we suggest additional research that will be necessary to resolve unanswered questions.

11:35–11:45

Panel Discussion

Session 2aSP

Signal Processing in Acoustics and Physical Acoustics: Subspace Methods for Acoustical Imaging I

David H. Chambers, Cochair

Lawrence Livermore National Laboratory, L-154, P.O. Box 808, Livermore, California 94551-5508

Sean K. Lehman, Cochair

Lawrence Livermore National Laboratory, L-154, 7000 East Avenue, Livermore, California 94550

Chair's Introduction—9:55

Invited Papers

10:00

2aSP1. Imaging with eigenfunctions of a scattering operator. Robert C. Waag (Depts. of Elec. & Computer Eng. and Radiol., Univ. of Rochester, Rochester, NY 14642)

An inverse scattering method using eigenfunctions of a scattering operator is reviewed. The framework of the method encompasses the use of eigenfunctions for focusing and quantitative image reconstruction in arbitrary media. Scattered acoustic fields are described using an operator. The eigenfunctions correspond to far-field patterns of an effective source distribution. Incident-wave patterns specified by the eigenfunctions focus on the distribution. The eigenfunction focusing properties are employed to reconstruct an unknown scattering medium by using products of numerically calculated fields defined in terms of eigenfunctions. A linearized version of the method is equivalent to the filtered-backprojection method for Born inversion. The full range of frequencies present in an incident pulse waveform can be used in the method. The method is examined by using both calculated and measured data. In the calculations, an exact solution for scattering from nonconcentric cylinders was employed to obtain the scattered field. In the measurements, a novel ring-transducer system was used to obtain the incident and total fields. The results of simulations and experiments show that the method is accurate for objects large compared to the incident pulse center-frequency wavelength and that the point resolution of the method is about one-half the center-frequency wavelength.

10:30

2aSP2. Imaging with the DORT method. Claire Prada, Estelle Kerbrat, Jean-Louis Thomas, and Mathias Fink (Lab. Ondes et Acoustique, ESPCI, 10 rue Vauquelin, 75005 Paris, France)

The decomposition of the time reversal operator (DORT) method applies to active detection and focusing of acoustic waves using arrays of transmitters and receivers. This method allows detection and selective focusing on scatterers through complex media. It consists in the construction of the invariants of the time reversal process, that are also the singular vectors of the array response matrix. The DORT method is particularly interesting for detection in inhomogeneous media when the acoustic properties are poorly known. However, if an estimate of the medium's Green function is available, images of the medium can be formed by backpropagation of the dominant eigenvectors. For the detection of defect in scattering media like titanium, the method is useful to separate the echo of a defect from the microstructure contribution and thus to reduce speckle noise. We will also see that the time reversal operator can be interpreted as a pseudocovariance matrix like those encountered in classical array signal processing. And as shown by Anthony J. Devaney [J. Acoust. Soc. Am. **110**, 2617 (2001)], nonlinear estimators like MUSIC can be applied to achieve high resolution. These different points will be illustrated through several experimental results.

11:00

2aSP3. Vector subspace methods in inverse acoustic wave scattering. Anthony J. Devaney (ECE, Northeastern Univ., 360 Huntington Ave., Boston, MA 02115)

The problems of detecting, locating, and identifying acoustic scatterers embedded in a known inhomogeneous background from the multistatic data matrix collected from an arbitrary unstructured, mixed sensor array are addressed using the distorted wave Born approximation. Based on this formulation two classes of problems are considered: (i) detecting and locating a finite set of point scatterers, (ii) imaging an arbitrary distribution of point or extended scatterers. Problems of the first class are shown to admit solutions based on the singular value decomposition (SVD) of the multistatic data matrix considered as a linear mapping from the finite vector space of transmitter inputs to the finite vector space of receiver outputs. In this case the SVD of the data matrix is shown to lead

directly to generalized time-reversal algorithms that allow super-resolution location estimation. Problems of the second class are also addressed using the SVD but this time it is based on the multistatic data matrix considered as a linear mapping from the Hilbert space of scatterer distributions to the finite vector space of receiver outputs. In this case a generalized form of the filtered backpropagation algorithm is derived and shown to lead to a minimum norm image of the scatterer distribution.

11:30

2aSP4. Remote sensing and communications in random media. George Papanicolaou (Dept. of Math., Stanford Univ., Stanford, CA 94305)

Reliable, high-capacity communications in scattering media can be effectively established with some basic remote sensing techniques involving time reversal. I will formulate these problems and discuss the various mathematical approaches that can be used for analysis. It turns out that stochastic analysis plays an important role and, in some cases, gives very satisfactory results. One such result is the spectacular increase in communications capacity in a richly scattering environment. I will end with a discussion of applications and computational issues that arise in the realistic simulation of communication systems.

TUESDAY MORNING, 29 APRIL 2003

ROOMS 103/104, 10:15 A.M. TO 12:00 NOON

Session 2aUW

Underwater Acoustics, Signal Processing in Acoustics and Engineering Acoustics: Robust Passive Sonar I

Lisa M. Zurk, Cochair

Lincoln Laboratory, Massachusetts Institute of Technology, 244 Wood Street, Lexington, Massachusetts 02173-6426

Brian H. Tracey, Cochair

Lincoln Laboratory, Massachusetts Institute of Technology, 244 Wood Street, Lexington, Massachusetts 02173-6426

Contributed Papers

10:15

2aUW1. Detection of a weak target in the presence of loud moving noise sources. Edmund J. Sullivan (OASIS, Inc., 5 Militia Dr., Lexington, MA 02421), James V. Candy (Lawrence Livermore Natl. Lab., Livermore, CA 94550), and William M. Carey (Boston Univ., Boston, MA 02115)

A previous study [J. Acoust. Soc. Am. **112** (2002)] treated the detection of a weak target in an adverse shallow-water environment with ambient noise and "known" interfering ships by applying a model-based adaptive (recursive) technique. The shallow-water environment and noise sources were represented by a normal-mode model directly incorporated into the model-based processor, thereby allowing their effects to be removed from the decision function prior to target detection. This previous work assumed that the interfering ships were motionless. Here, a general state-space framework is developed, that includes a dynamic model of the moving interferers together with the adaptive signal processor in a self-consistent formalism. The decision function is the so-called "Weighted Sum Square Residual" (WSSR), which evolves directly from the innovation sequence of the model-based processor. An example is shown for a low level sound source in Gaussian ambient noise with a strongly interfering moving ship. It is shown further that, upon detection, the innovation sequence can be used as "prefiltered" data providing input that can be used to track the source after its detection.

10:30

2aUW2. Channel order selection in blind deconvolution based on eigenvector characteristics and using normal modes in conjunction with multipath compression for source identification. James P. LaRue, George E. Ioup, Juliette W. Ioup (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148, jlarue@uno.edu), and George B. Smith (Naval Res. Lab., Stennis Space Center, MS 39529)

Two related topics in blind deconvolution are presented. The first concerns channel order selection. Many blind deconvolution techniques assume the filter length to be known. One technique of filter length deter-

mination uses the dimension of a nullspace obtained by decomposing correlation matrices. Although the dimension is usually determined by calculations involving the eigenvalues, this alternative method relies on the known characteristics of eigenvectors among the subspaces. Multipath functions in underwater acoustics commonly have high kurtosis. The new method can give correct estimates on filter length for signal-to-noise ratios as low as 10 dB. The matrix methods described may not work as well when the sample length increases. However, the second topic shows how normal mode properties in conjunction with time compression in modeled multipath functions may be used to enhance blind deconvolution in a waveguide when the sample length is large. Given data from a vertical array of hydrophones, one may accurately obtain a source identification. The methods of the tracking and positioning of normal modes are incorporated with an input of a linear frequency-modulated pulse propagated through an ocean environment. [Research supported by ONR.]

10:45

2aUW3. Model-based source localization in the 8–16 kHz band using the channel impulse response function. Paul Hursky, Martin Siderius, Michael B. Porter (Sci. Applications Intl. Corp., 10260 Campus Point Dr., San Diego, CA 92121), and Vincent K. McDonald (Space and Naval Warfare Systems Ctr., San Diego, CA 92152)

SignalEx tests have been conducted in a variety of shallow water coastal environments to relate the performance of acoustic communications systems to the prevailing oceanographic conditions. These tests have typically consisted of using a fixed receiver (one to four elements, spaced for diversity) and a transmitter drifting out to ranges beyond the minimum detectable level. During these tests, waveforms to probe the channel in the 8 to 16 kHz band were transmitted at regular intervals. These probe signals were used to test source localization algorithms at these high frequencies. Model-based source localization at these very high frequencies requires

either very accurate modeling or algorithms inherently robust against model mismatch. Although the first couple of multipath arrivals can be stabilized from ping to ping, the later arrivals exhibit rapidly fluctuating amplitudes and times of arrival, due to the motion of the ocean surface, water column variability, and the varying bathymetry as the transmitter drifts in range. Measurements of the time-varying channel response and source localization results using the 8–16 kHz band SignalEx channel probes will be presented from sites at the New England Front and the Coronado Bank off San Diego.

11:00

2aUW4. Horizontal-array beamforming using the waveguide invariant. Daniel Rouseff (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105) and Altan Turgut (Naval Res. Lab., Washington, DC 20375)

It has been shown that with a horizontal array and incoherent processing, the measured acoustic intensity produced by a distant source exhibits striations when plotted versus range and frequency. The striations are a consequence of interference between pairs of propagating acoustic modes, and the waveguide invariant describes their slope. The present work extends this concept to coherent processing, such as beamforming with a horizontal array. Equations are derived that describe the trajectories of the striations observed in LOFARgrams. (LOFARgrams display the time-evolving spectrum of a beamformer for a particular look direction.) The validity of the equations is first tested in numerical simulations and then tested using horizontal-array data from the ONR Asian Seas International Acoustics Experiment (ASIAEx) conducted in the South China Sea. The effect of mismatch between the beamformer look direction and the true bearing of the source is also discussed. [Work supported by ONR.]

11:15

2aUW5. Multipath compression by multiple convolutions. George B. Smith (Code 7185, Ocean Acoust. Div., Naval Res. Lab., Stennis Space Center, MS 39529-5004)

Blind deconvolution algorithms can be useful as pre-processors for signal classification algorithms in shallow water. These algorithms remove the distortion of the signal caused by multipath propagation when no knowledge of the environment is available. A very simple blind algorithm is presented here which utilizes convolutions of multiple data channels for convolutional smoothing of multipath. Computer simulation studies show significant performance improvement when propagation effects are mitigated by the use of this blind technique. [Work supported by ONR and the Naval Research Laboratory.]

11:30

2aUW6. Performance bounds on wave-front curvature ranging.

Joseph E. Bondaryk (Titan Corp., 470 Totten Pond Rd., Waltham, MA 02451), Phillip Abbot, Charles Gedney, and Edward Sullivan (Ocean Acoust. Services and Instrumentation Systems, Inc., Lexington, MA 02421)

Acoustic wave-front curvature ranging (WCR) is a localization method that exploits the curvature of the arriving wave-front to determine the range of an acoustic emitter. By assuming that the wave-front is circular, it provides an estimate of the range as the radius of the circle. It is based on the measurement of time of arrival differences for correlated transient signals between pairs of spatially separated receivers. Performance bounds are determined by: (1) source and receiver geometry, (2) signal characteristics, and (3) spatial coherence limits due to oceanic effects. Both theoretical and Monte Carlo studies have been done which define the performance limitations as a function of aperture size, signal structure and spatial coherence. Results show that for a given aperture and fixed range accuracy, the low-frequency limit is bounded by the signal characteristics (SNR and time-bandwidth product), whereas the high-frequency limit is determined by the coherence length allowed by the ocean. This high-frequency limit is presently estimated by extrapolation of lower frequency models and is imprecisely known due to the unavailability of high-frequency, short time spatial coherence data.

11:45

2aUW7. Source bearing estimation using ice-mounted geophones.

Stan Dosso (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, sdosso@uvic.ca), Michael Vinnins (Defence Res. and Development Canada—Ottawa, Ottawa, ON K1A 0Z4, Canada), and Garry Heard (Defence Res. and Development Canada—Atlantic, Dartmouth, NS B2Y 3Z7, Canada)

This paper presents the results of Arctic field trials carried out to estimate the bearing to acoustic sources in the water column using seismo-acoustic particle motion measured at a tri-axial geophone mounted on the sea ice surface. Source bearings are estimated by applying polarization filters to suppress seismic wave types with transverse particle motion, and computing the incident power rotated into radial look angles from 0–360 deg. The 180-deg ambiguity inherent in this rotational analysis is resolved by requiring prograde particle motion in the radial-vertical plane. Results are presented for smooth and rough annual ice and multiyear ice for source ranges from 200 m to 50 km. The results indicate good bearing estimation to long ranges with little dependence on ice type.

Session 2pAAa**Architectural Acoustics: Student Design Competition**

Robert C. Coffeen, Cochair

School of Architecture and Urban Design, University of Kansas, Marvin Hall, Lawrence, Kansas 66045

Lily M. Wang, Cochair

Architectural Engineering, University of Nebraska–Lincoln, 200B Peter Kiewit Institute, 1110 South 67th Street, Omaha, Nebraska 68182-0681

Robin Glosemeyer, Cochair

Jaffe Holden Acoustics, 114A Washington Street, Norwalk, Connecticut 06854

The Technical Committee on Architectural Acoustics of the Acoustical Society of America and the National Council of Acoustical Consultants are sponsoring this Student Design Competition that will be professionally judged at this meeting.

The purpose of this design competition is to encourage students enrolled in architecture, architectural engineering, and other university curriculums that involve building design and/or acoustics to express their knowledge of architectural acoustics and building noise control in the schematic design of a building where acoustical considerations are of primary importance. This competition is open to undergraduate and graduate students from all nations.

The submitted designs will be displayed in this session and they will be judged by a panel of professional architects and acoustical consultants. Up to five entries will be selected for awards, one “First Honors” award and four “Commendation” awards. An award of \$1,000 will be given to the entry judged “First Honors.” An award of \$500 will be given to each of the entries judged “Commendation.”

TUESDAY AFTERNOON, 29 APRIL 2003

ROOMS 108/109, 1:30 TO 4:20 P.M.

Session 2pAAb**Architectural Acoustics and Musical Acoustics: Relationships of Synthesis and Processing to “Acoustical” Music**

K. Anthony Hoover, Chair

*Cavanaugh Tocci Associates, Inc., 327F Boston Post Road, Sudbury, Massachusetts 01776***Chair’s Introduction—1:30*****Invited Papers*****1:35**

2pAAb1. Virtual acoustics for music practice rooms. Ron Freiheit (Wenger Corp., 555 Park Dr., Owatonna, MN 55050, ron.freiheit@wengercorp.com)

The use of virtual acoustics has provided a new level of practice experience for the musician. By integrating the sound isolation of music practice rooms with the signal processing of an active acoustic system (with time variant-gain before feedback) musicians can now benefit from the experience of practicing in multiple acoustic environments. Musicians select from various acoustics environments from a typical small practice room to that of a large space such as a sports arena. The variability of the acoustic environment allows the musician to hear clearly their intonation and articulation, which may be difficult to discern in a small practice room. To effectively communicate the various acoustics environments, the musicians must be immersed in the sound field of the active acoustics without being able to discern source locations of the speakers. The system must also be able to support the dynamic range of the musicians without presenting artifacts of its own such as system noise or audible distortion. This paper deals with the design constraints needed to meet these requirements as well the antidotal responses from musicians who have used these environments for practice.

1:55

2pAAb2. Using miniature signal processing equipment in real-time brass performance. Thomas J. Plsek (Berklee College of Music, 1140 Boylston St., Boston, MA 02215)

Real time signal processing for brass instrument performance has been in use for more than 20 years now. It has been fraught with many problems not the least of which is the complexity, size, and expense of the equipment as well as the acoustical output of the instrument itself. One device which addresses these issues is the new Yamaha ST5: Personal Studio for brass instruments. By

combining a Yamaha Pickup Mute, which very effectively minimizes the acoustical output of the instrument, with a battery powered unit small enough to be worn on a belt (ca. 5 in.×3 in.×1 in.), this system enables the performer to use such effects as reverb, delay, chorus, equalizer, pitch shifter, etc., that can be used in a wide variety of situations such as private practice, live concert performances, and recordings. By creatively managing the acoustic instrument and the miniature electronic equipment, a reasonably large array of musical resources become available to the performer enabling him/her to enhance existing performance environments, as well as find and develop new ones.

2:15

2pAAb3. The use of delay in multitrack production. Alexander U. Case (Fermata Audio + Acoust., P.O. Box 1161, Portsmouth, NH 03802)

Delay, inevitable whenever sound propagates through space, is too often the bane of the acoustician's practice. An audible echo generally relegates a music performance hall—no matter how beautiful it otherwise might sound—to the lowest status. Multitrack music production on the other hand, with its aggressive use of overdubbing, editing, and signal processing, is not bound by those rules of time and space which determine the sound of a hall. In the recording studio, where music is synthesized for playback over loudspeakers, the delay is employed as a powerful, multipurpose tool. It is not avoided. It is in fact embraced. Echoes are used on purpose, strategically, to enhance the loudspeaker listening experience. Moreover, the humble delay is the basis for many nonecho effects. Flanging, chorus, and pitch shifting are delay-based effects regularly used in audio engineering practice. This paper discusses some of the more common delay-based effects, reviewing their technical structure, the psychoacoustic motivation behind them, and the musical value they create.

2:35

2pAAb4. Music 4C, a multi-voiced synthesis program with instruments defined in C. James W. Beauchamp (School of Music and Dept. of Elec. and Computer Eng., Urbana, IL 61801, j-beauch@uiuc.edu)

Music 4C is a program which runs under Unix (including Linux) and provides a means for the synthesis of arbitrary signals as defined by the C code. The program is actually a loose translation of an earlier program, Music 4BF [H. S. Howe, Jr., *Electronic Music Synthesis* (Norton, 1975)]. A set of instrument definitions are driven by a numerical score which consists of a series of "events." Each event gives an instrument name, start time and duration, and a number of parameters (e.g., pitch) which describe the event. Each instrument definition consists of event parameters, performance variables, initializations, and a synthesis algorithmic code. Thus, the synthetic signal, no matter how complex, is precisely defined. Moreover, the resulting sounds can be overlaid in any arbitrary pattern. The program serves as a mixer of algorithmically produced sounds or recorded sounds taken from sample files or synthesized from spectrum files. A score file can be entered by hand, generated from a program, translated from a MIDI file, or generated from an alpha-numeric score using an auxiliary program, Notepro. Output sample files are in wav, snd, or aiff format. The program is provided in the C source code for download.

2:55–3:05 Break

3:05

2pAAb5. Inside-in, alternative paradigms for sound spatialization. Curtis Bahn and Stephan Moore (Rensselaer Polytechnic Inst., iEAR Studios, Rochester, NY, crb@rpi.edu)

Arrays of widely spaced mono-directional loudspeakers (P.A.-style stereo configurations or "outside-in" surround-sound systems) have long provided the dominant paradigms for electronic sound diffusion. So prevalent are these models that alternatives have largely been ignored and electronic sound, regardless of musical aesthetic, has come to be inseparably associated with single-channel speakers, or headphones. We recognize the value of these familiar paradigms, but believe that electronic sound can and should have many alternative, idiosyncratic voices. Through the design and construction of unique sound diffusion structures, one can reinvent the nature of electronic sound; when allied with new sensor technologies, these structures offer alternative modes of interaction with techniques of sonic computation. This paper describes several recent applications of spherical speakers (multichannel, outward-radiating geodesic speaker arrays) and Sensor-Speaker-Arrays (SenSAs: combinations of various sensor devices with outward-radiating multi-channel speaker arrays). This presentation introduces the development of four generations of spherical speakers—over a hundred individual speakers of various configurations—and their use in many different musical situations including live performance, recording, and sound installation. We describe the design and construction of these systems, and, more generally, the new "voices" they give to electronic sound.

3:25

2pAAb6. The loudspeaker as musical instrument: An examination of the issues surrounding loudspeaker performance of music in typical rooms. David Moulton (Sausalito Audio Works, 61C Galli Dr., Novato, CA, dave@sawonline.com)

The loudspeaker is the most important and one of the most variable elements in the electroacoustic music performance process. Nonetheless, its performance is subject to a "willing suspension of disbelief" by listeners and its behavior and variability are usually not accounted for in assessments of the quality of music reproduction or music instrument synthesis, especially as they occur in small rooms. This paper will examine the aesthetic assumptions underlying loudspeaker usage, the general timbral qualities and sonic characteristics of loudspeakers and some of the issues and problems inherent in loudspeakers interactions with small rooms and listeners.

2p TUE. PM

2pAAb7. A new loudspeaker design: A case study of an effort to more fully integrate the loudspeaker into the playback room in a musical way. David Moulton (Sausalito Audio Works, 61C Galli Dr., Novato, CA 94949, dave@sawonline.com) and Poul Praestgaard (Bang & Olufsen, DK 7600 Struer, Denmark)

The authors have been members of a design and development team that has created a new loudspeaker that attempts to resolve several of the primary problems presented by the loudspeaker/room/listener interface, as described in one of the author's previous papers. This paper will describe that new loudspeaker, its various new approaches to the interactions between the loudspeaker, the room and the listeners, and a brief review of the research, findings and assumptions underlying its design. The authors hope to have examples of the loudspeaker available for demonstration.

Contributed Paper

4:05

2pAAb8. Physical modeling synthesis of recorder sound. Hiroko Shiraiwa (CCRMA, Dept. of Music, Stanford Univ., Stanford, CA 94305-8180), Kenshi Kishi (Univ. of Electro-Communications, Chofugaoka 1-5-1 Chofu, Tokyo 182-8585, Japan), and Isao Nakamura (Athena Co. Ltd., 1-33-25 Kokuryo-cho, Chofu, Tokyo 182-0022, Japan)

A time-domain simulation of the soprano baroque recorder based on the digital waveguide model (DWM) and an air reed model is introduced. The air reed model is developed upon the negative acoustic displacement model (NADM), which was proposed for the organ flue-pipe simulation

[Adachi, Proc. of ISMA 1997, pp. 251–260], based on the semiempirical model by Fletcher [Fletcher and Rossing, *The Physics of Musical Instruments*, 2nd ed. (Springer, Berlin, 2000)]. Two models are proposed to couple DWM and NADM. The jet amplification coefficient is remodeled for the application of NADM for the recorder, regarding the recent experimental reports [Yoshikawa and Arimoto, Proc. of ISMA 2001, pp. 309–312]. The simulation results are presented in terms of the mode transient characteristics and the spectral characteristics of the synthesized sounds. They indicate that the NADM is not sufficient to describe the realistic mode transient of the recorder, while the synthesized sounds maintained almost resemble timbre to the recorder sounds.

TUESDAY AFTERNOON, 29 APRIL 2003

ROOMS 103/104, 1:00 TO 5:15 P.M.

Session 2pAO

Acoustical Oceanography, Underwater Acoustics and Signal Processing in Acoustics: Geoacoustic Inversion III

Peter Gerstoft, Chair

Scripps Institution of Oceanography, Marine Physical Laboratory, University of California, San Diego, 9500 Gillman Drive, La Jolla, California 92093-0238

Invited Paper

1:00

2pAO1. Developments in self-noise towed-array inversion. David J. Battle (Marine Physical Lab, Univ. of California, San Diego, La Jolla, CA 92093-0238, davidb@mpl.ucsd.edu)

In preliminary experiments, self-noise geoacoustic inversion from towed-array data has been demonstrated as a feasible concept, whereby sonar performance predictions in shallow water can be based on locally measured geoacoustic parameters while a tow-ship is underway. In this approach, replica fields generated using a near-field propagation model are compared to measured source fields via a conventional matched-field processor (MFP). With parameter selection controlled by a combination of local and global search algorithms, MFP output power is used as a feedback signal to optimize the match between the modeled and actual acoustic environments. In addition to the desired geoacoustic parameters, the search space includes nuisance parameters such as unknown or poorly known depths, ranges and array perturbations. This paper discusses further results and practicalities of self-noise inversion using towed-arrays, including aspects of environmental sensitivity, propagation model selection, source signature acquisition and imaging.

Contributed Papers

1:20

2pAO2. Rapid geoacoustic inversion with a curved horizontal array. Laurie T. Fialkowski, Dalcio K. Dacol, Joseph F. Lingevitch, and Elisabeth Kim (Naval Res. Lab., Washington, DC 20375)

Real-time geoacoustic inversions with a towed array are of interest for rapidly characterizing the sediment properties over changing regions, and require an efficient and accurate forward propagation model. A wavenumber integration solution using the computationally efficient method of un-

determined coefficients and the Fast-Field Program is implemented with a simulated annealing optimization method and a coordinate rotation technique. The source function of interest may be broadband ship self-noise or a controlled source. The parameters sought in the inversion include, but are not limited to, geoacoustic sediment properties and receiver array geometry; the water column is assumed known. The forward propagation method is efficient and accurate for very short ranges. Results are presented of simulations with a controlled source, as well as preliminary MAPEX2000 ship self-noise. [Work supported by ONR.]

2pAO3. Forward modeling requirement for short-range geoacoustic inversion on a towed array. T. C. Yang and K. Yoo (Naval Res. Lab., Washington, DC 20375)

For geoacoustic inversion using the tow-ship noise received on a towed horizontal line array (HLA), one of the issues is speed and accuracy of the forward model. Is the normal mode good enough or is ray propagation adequate for modeling short-range forward propagation? Earlier work [Kuperman *et al.*, IEEE J. Oceanic Eng. **10** (1985)] showed that the continuum contribution arrives at beams far away from the forward directions. For practical applications, such contributions are often neglected to minimize interference from undesired sources. We examine in this paper the same issue (the importance of high angle arrivals at short range) in both the time and beam domain in the context of several bottom models (templates). Note that the high angle arrival contributions depend on the bottom type, the source–receiver range, the towed array aperture/spacing and the acoustic frequencies. The results of this analysis will shed light on the requirement of forward models at short ranges on a HLA. [Work supported by ONR.]

1:50

2pAO4. Sensitivity to array tilt and bow for broadband geoacoustic inversion using a towed array. K. Yoo, T. C. Yang, L. Fialkowski, D. Dacol, John Perkins (Naval Res. Lab., Washington, DC 20375), M. Fallat, P. Nielsen (SACLANT Undersea Res. Ctr., Italy), and M. Siderius (SAIC, La Jolla, CA 92037)

During the MAPEX 2000 experiment, broadband acoustic data from an acoustic source were received on a towed array, both towed by the same ship with the source–receiver range kept nearly constant. The data were used to invert for the geoacoustic properties of the bottom along the ship track. The experiment was conducted by the SACLANTCEN in the Sicily Strait where the bottom layer profile was also surveyed using standard seismic methods. The geoacoustic inversion results from the towed array data showed good agreement with the seismic data [M. Siderius *et al.*, J. Acoust. Soc. Am. **112**, 1523 (2002); M. Fallat *et al.*, SACLANTCEN Report No. SM-402, 2002] using a matched field correlation cost function that summed over frequencies coherently (for each phone) and summed over the phones incoherently. The authors reported that an approach that sums the matched-field correlation coherently over phones and incoherently over frequencies did not yield as good a result as the method mentioned above. The array tilt and bow may be one factor that impacts correlation functions in different ways. We study in this paper the matched field inversion sensitivity to the array tilt and bow as applied to the MAPEX 2000 data. [Work supported by ONR.]

2:05

2pAO5. Geoacoustic characterization of a range-dependent environment using towed array data. Mark Fallat, Peter Nielsen (SACLANT Undersea Res. Ctr., Viale S. Bartolomeo 400, 19138 La Spezia, Italy, fallat@saclantc.nato.int), and Martin Siderius (Sci. Applications Intl. Corp., La Jolla, CA 92037)

This paper describes geoacoustic characterization of a range-dependent environment using towed horizontal array data. Data from multiple points along a track are inverted using a short-range, range-independent scheme and the results are combined to produce a range-dependent model. Data from the MAPEX 2000 experiment, conducted by SACLANT Centre in the Mediterranean Sea, are used to determine seabed properties for a range-dependent environment. Inversion results are compared for both normal-mode and ray theory forward models. The layering structure of the range-dependent model compares favorably with a high-resolution seismic profile.

2:35

2pAO6. Geoacoustic inversion and source localization from multisource broadband HLA data. Tracianne B. Neilsen (Appl. Res. Lab., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, neilsen@arlut.utexas.edu), Craig S. MacInnes (Pontificia Universidad Catolica del Peru, Lima, Peru), and David P. Knobles (Univ. of Texas, Austin, TX 78713-8029)

In underwater acoustics research, source localization efforts are often hampered by incorrect environmental information, and geoacoustic inversion results can be limited by errors in the source description. In addition, any data set can be contaminated by sound from other acoustic sources in the ocean. To overcome these problems, an iterative rotated coordinates inversion method [T. B. Neilsen, “An iterative implementation of rotated coordinates for inverse problems,” J. Acoust. Soc. Am. (submitted)] is employed in conjunction with a broadband bearing estimator [J. Krolik and D. Swingler, IEEE Trans. Acoust., Speech and Signal Process. **37**, (1989)] and matrix filters [C. S. MacInnes, IEEE J. Ocean. Eng. (submitted)] to obtain estimates of the source locations and the sensitive geoacoustic parameters from multisource broadband data received on a horizontal line array (HLA). In the iterative inversion method, subsets of rotated coordinates are used to vary both the source and the sensitive environmental parameters in a series of simulated annealing inversions. The information contained in the rotated coordinates and corresponding eigenvalues determine which parameters are varied significantly in each simulated annealing inversion. [Work supported by ONR.]

2:50

2pAO7. Unified inversion using isolated moving sources. Steven A. Stotts, Robert A. Koch, Traci B. Neilsen, and Craig S. MacInnes (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, stotts@arlut.utexas.edu)

Geoacoustic inversion using beamformed data from a ship of opportunity has been demonstrated with a bottom mounted array [Koch and Knobles, J. Acoust. Soc. Am. **112**, 2282 (2002)]. An alternative approach transforms element level data into “beam space,” applies a bearing filter [MacInnes, IEEE J. Ocean. Eng. (submitted)] and transforms back to element level data prior to performing inversions [T. B. Neilsen and C. S. MacInnes (unpublished)]. Automation of this filtering approach is facilitated for broadband applications by restricting the inverse transform back to element data corresponding to the degrees of freedom of the array, i.e., the effective number of elements. Examples that demonstrate this filtering technique with simulated data are presented along with comparisons to inversion results using beamformed data. Vertical and horizontal ambiguity surfaces are compared for the two approaches. Examinations of cost functions calculated within a simulated annealing algorithm reveal the efficacy of the approach. Filter performance with real data will be discussed.

3:05

2pAO8. Geoacoustic inversion by using the broadband signals. Zhenglin Li, Renhe Zhang, Jianjun Liu, Fenghua Li, and Ling Xiao (Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, lzhl@fared.ioa.ac.cn)

Three different methods are used to invert the sea bottom parameters (sound speed, density, and attenuation) based on the fact that each parameter has different effect on the sound fields. First, the vertical reflection coefficients of sea bottom and the bottom acoustic impedance are inverted from the light-bulb signals. Second, the matched-field processing method is used to the propagation bomb signal to split the density and sound speed from the impedance with an assumed attenuation. In the end, the attenuation coefficients are gotten from the transmission loss data at the different frequency band. [Work supported by the National Natural Science Foundation of China.]

3:20

2pAO9. Inversion of range-dependent geoacoustic properties in South China Sea ASIAEx01 experimental site. Altan Turgut, Bruce Pasewark, Marshall Orr (Naval Res. Lab., Acoust. Div., Washington, DC 20375), James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Ching Sang Chiu (Naval Postgraduate School, Monterey, CA 93943)

Matched-field inversion of range-dependent geoacoustic properties is studied using broadband (50–200 Hz, 240–260 Hz, and 550–600 Hz) acoustic data collected in the South China Sea during the ASIAEx01 experiment. Range-dependent sediment sound-speed and attenuation profiles are inverted using a global optimization scheme based on genetic algorithms to minimize an objective function defined by the Bartlett processor output. For the forward model, an efficient coupled normal mode model is used to calculate broadband acoustic fields incorporating range-dependent bathymetry and sediment layers thickness obtained by chirp sonar surveys. Inversions were performed starting from a range-independent region with a relatively short source/receiver distance. Additional geoacoustic profiles were inverted by incorporating the previously inverted profiles as the source/receiver distance was increased. The results obtained at three different frequency bands are in good agreement, especially when the range-dependency of bathymetry and sediment layers thickness is included in the inversion. [Work supported by ONR.]

3:35

2pAO10. Geoacoustic inversion for elastic bottoms: Matching model predictions and backscatter data from littoral limestone. Robert F. Gragg, Raymond J. Soukup, and Roger C. Gauss (Naval Res. Lab., Code 7140, Washington, DC 20375)

The scattering strength of the ocean bottom as a function of angle and frequency is of fundamental importance in predicting the performance of active sonar systems, particularly in littoral waters. In this work, we apply an elastic-bottom scattering model [the small-slope model of Gragg *et al.*,

“Rough interface scattering,” J. Acoust. Soc. Am. **110** (2001)] to the problem of matching 2–3.5 kHz acoustic backscatter from rough limestone sea floors off the Carolina Coast and in the Straits of Sicily. We use a simplex/annealing algorithm to determine model parameters that optimally fit the data. Analysis of an ensemble of such inversion runs addresses the following questions. How well can the elastic theory match data measured at sea? How sensitive is the theoretical prediction to the values of its inputs—frequency and the (seven) parameters that characterize the geoacoustics of the bottom material and the roughness of the surface? Results to date indicate more sensitivity to roughness than to geoacoustic parameters. This supports the importance of estimating in situ bottom roughness and the feasibility of using a physics-based model for that purpose. [Work supported by ONR.]

3:50

2pAO11. Time-reversal using ambient noise as a probe source. Philippe Roux, Hee Chun Song, and W. A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California San Diego, La Jolla, CA 92093)

The Green’s function between receivers in the ocean can be obtained from the temporal cross correlation of ambient noise measured simultaneously at those receivers [Ph. Roux, S. Lynch, W. A. Kuperman, and the NPAL Group, J. Acoust. Soc. Am. **112**, 2421 (2002)]. Constructing the Green’s function similarly between a point receiver and an (source/receive) array from ambient noise creates, in effect, a surrogate probe source that provides data on a time reversal mirror. We show with theory and simulation that this noise-based time-reversal procedure results in a focus at the surrogate probe source. Finally, data are used to illustrate the feasibility of this process.

4:05–4:15 Break

4:15–5:15

Panel Discussion

TUESDAY AFTERNOON, 29 APRIL 2003

ROOMS 110/111, 1:00 TO 2:15 P.M.

Session 2pEA

Engineering Acoustics: Microphones and Sources

Stephen C. Thompson, Chair

Knowles Electronics, Inc., 1151 Maplewood Drive, Itasca, Illinois 60143

Contributed Papers

1:00

2pEA1. Beamforming beyond the $\lambda/2$ limit with microphone arrays. Philippe Moquin, Stéphane Dedieu (Mitel Networks, 350 Legget Dr., Kanata, ON K2K 1X3, Canada), and Rafik A. Goubran (Carleton Univ., Ottawa, ON K1S 5B6, Canada)

One limitation of microphone arrays is that the inter-microphone spacing is restricted to $\lambda/2$ of the shortest wavelength (highest frequency) of interest. For an increase in frequency range, the array must either be made smaller (thereby losing low-frequency directivity) or the number of microphones must be increased (thereby increasing cost). The other problem is

that the beamwidth decreases with increasing frequency and sidelobes become more problematic. This results in significant off-axis “coloration” of the signals. The extension of the working frequency range for an existing narrow-band (300–3 kHz) telephony microphone array to wide-band telephony (up to 7 kHz), without modifying its geometry and the number of microphones will be shown. Microphones are embedded in a diffraction structure that provides the desired directivity at high frequencies. To provide the desired directionality at lower frequencies, beamforming of the microphones is performed using digital signal processing techniques. The combination of beamforming and embedding the microphones in a diffraction structure that provides the desired directivity at high frequencies ad-

dresses the two weaknesses that arise in previous approaches: low-frequency directivity with small arrays and high-frequency difficulties that arise in conventional sensor arrays. [Work supported in part by Carleton University.]

1:15

2pEA2. Practical considerations for a second-order directional hearing aid microphone system. Stephen C. Thompson (Knowles Electron., LLC, 1151 Maplewood Dr., Itasca, IL 60143, steve.thompson@knowles.com)

First-order directional microphone systems for hearing aids have been available for several years. Such a system uses two microphones and has a theoretical maximum free-field directivity index (DI) of 6.0 dB. A second-order microphone system using three microphones could provide a theoretical increase in free-field DI to 9.5 dB. These theoretical maximum DI values assume that the microphones have exactly matched sensitivities at all frequencies of interest. In practice, the individual microphones in the hearing aid always have slightly different sensitivities. For the small microphone separation necessary to fit in a hearing aid, these sensitivity matching errors degrade the directivity from the theoretical values, especially at low frequencies. This paper shows that, for first-order systems the directivity degradation due to sensitivity errors is relatively small. However, for second-order systems with practical microphone sensitivity matching specifications, the directivity degradation below 1 kHz is not tolerable. A hybrid order directive system is proposed that uses first-order processing at low frequencies and second-order directive processing at higher frequencies. This hybrid system is suggested as an alternative that could provide improved directivity index in the frequency regions that are important to speech intelligibility.

1:30

2pEA3. Microphone matching for hybrid-order directional arrays in hearing aid applications. Daniel M. Warren and Steve C. Thompson (Knowles Electron., LLC, 1151 Maplewood Dr., Itasca, IL 60143, daniel.warren@knowles.com)

The ability of a hearing aid user to distinguish a single speech source amidst general background noise (for example, dinner table or cocktail party conversation) may be improved by a directional array of microphones in the hearing instrument. The theoretical maximum directivity index (DI) of a first-order pairing of microphones is 6 dB, and a second-order array of three microphones is 9.5 dB, assuming all three microphones have identical frequency responses. The close spacing of microphone ports in a hearing aid body means that directivity degrades rapidly with differences in microphone sensitivities. A hybrid of first- and second-order arrays can mitigate this effect, although close microphone matching is still necessary for high directivity. This paper explores the effect of microphone mismatch on the directivity of such arrays, and describes

practical criteria for selecting matched microphones out of production batches to maximize a speech intelligibility weighted directivity index. [Work supported by Knowles Electronics, LLC.]

1:45

2pEA4. Calculation of uncertainty in calibration of microphones by the pressure reciprocity technique. Peter Hanes, Lixue Wu, Won-Suk Ohm, and George S. K. Wong (Inst. for Natl. Measurement Standards, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada, peter.hanes@nrc-cnrc.gc.ca)

At the primary level, acoustical measurement standards are realized through calibration of the sensitivity level of Laboratory Standard microphones by the reciprocity technique. The technique is described in International Standard IEC 61094-2, which allows for various implementations of the measurement method. The pressure sensitivity levels of a set of three microphones are determined from the electrical and acoustical transfer impedances of pairs of the microphones. The transfer impedances in turn depend on the design and performance of the measurement apparatus, the dimensions and acoustical properties of the microphones and the cavity that acts as an acoustical coupler between the microphones, and the prevailing environmental conditions. The uncertainty in the pressure sensitivity level depends on the uncertainties in these input quantities and on how the sensitivity level varies with changes in the input quantities. The ISO/IEC Guide Express: *1995 Guide to the Expression of Uncertainties in Measurement* provides internationally agreed models and guidance for evaluating the expanded uncertainty of a measurement. The uncertainty model, the nature of the input variables, and the steps involved in the calculation of the expanded uncertainty are described for the realization of a particular implementation of the reciprocity technique.

2:00

2pEA5. High-frequency monopole sound source for anechoic chamber qualification. Patrick Saussus and Kenneth A. Cunefare (Georgia Inst. of Technol., 771 Ferst Dr., Atlanta, GA 30332, gte036z@prism.gatech.edu)

Anechoic chamber qualification procedures require the use of an omnidirectional monopole sound source. Required characteristics for these monopole sources are explicitly listed in ISO 3745. Building a high-frequency monopole source that meets these characteristics has proved difficult due to the size limitations imposed by small wavelengths at high frequency. A prototype design developed for use in hemianechoic chambers employs telescoping tubes, which act as an inverse horn. This same design can be used in anechoic chambers, with minor adaptations. A series of gradually decreasing brass telescoping tubes is attached to the throat of a well-insulated high-frequency compression driver. Therefore, all of the sound emitted from the driver travels through the horn and exits through an opening of approximately 2.5 mm. Directivity test data show that this design meets all of the requirements set forth by ISO 3745.

2p TUE. PM

Session 2pED

Education in Acoustics, Psychological and Physiological Acoustics, Noise, Speech Communication and ASA Committee on Standards: An Educated Consumer's Guide to Hearing Loss and Hearing Aids

Amy M. Donahue, Chair

NIDCD/NIH, Division of Human Communication, 6120 Executive Boulevard, Rockville, Maryland 20852

Chair's Introduction—1:00

Invited Papers

1:05

2pED1. Hearing and hearing loss: Causes, effects, and treatments. Richard A. Schmiedt (Dept. of Otolaryngol. and Head-Neck Surgery, Medical Univ. of South Carolina, P.O. Box 250550, Charleston, SC 29425, schmiera@muscc.edu)

Hearing loss can have multiple causes. The outer and middle ears are conductive pathways for acoustic energy to the inner ear (cochlea) and help shape our spectral sensitivity. Conductive hearing loss is mechanical in nature such that the energy transfer to the cochlea is impeded, often from eardrum perforations or middle ear fluid buildup. Beyond the middle ear, the cochlea comprises three interdependent systems necessary for normal hearing. The first is that of basilar-membrane micromechanics including the outer hair cells. This system forms the basis of the cochlear amplifier and is the most vulnerable to noise and drug exposure. The second system comprises the ion pumps in the lateral wall tissues of the cochlea. These highly metabolic cells provide energy to the cochlear amplifier in the form of electrochemical potentials. This second system is particularly vulnerable to the effects of aging. The third system comprises the inner hair cells and their associated sensory nerve fibers. This system is the transduction stage, changing mechanical vibrations to nerve impulses. New treatments for hearing loss are on the horizon; however, at present the best strategy is avoidance of cochlear trauma and the proper use of hearing aids. [Work supported by NIA and MUSC.]

1:30

2pED2. Hearing loss and the central auditory system: Implications for hearing aids. Robert D. Frisina (Otolaryngol. Div., Univ. of Rochester Med. School, 601 Elmwood Ave., Rochester, NY 14642-8629, rdf@q.ent.rochester.edu)

Hearing loss can result from disorders or damage to the ear (peripheral auditory system) or the brain (central auditory system). Here, the basic structure and function of the central auditory system will be highlighted as relevant to cases of permanent hearing loss where assistive devices (hearing aids) are called for. The parts of the brain used for hearing are altered in two basic ways in instances of hearing loss: (1) Damage to the ear can reduce the number and nature of input channels that the brainstem receives from the ear, causing plasticity of the central auditory system. This plasticity may partially compensate for the peripheral loss, or add new abnormalities such as distorted speech processing or tinnitus. (2) In some situations, damage to the brain can occur independently of the ear, as may occur in cases of head trauma, tumors or aging. Implications of deficits to the central auditory system for speech perception in noise, hearing aid use and future innovative circuit designs will be provided to set the stage for subsequent presentations in this special educational session. [Work supported by NIA-NIH Grant P01 AG09524 and the International Center for Hearing & Speech Research, Rochester, NY.]

1:55

2pED3. Introduction to auditory perception in listeners with hearing losses. Mary Florentine (Inst. of Hearing, Speech, & Lang. and SLPA Dept. (151A FR), Northeastern Univ., 360 Huntington Ave., Boston, MA 02115-5000, florentin@neu.edu) and Søren Buus (Inst. of Hear., Speech, & Lang. and Commun. & Dig. Sig. Proc. Ctr., Northeastern Univ., Boston, MA 02115-5000)

Listeners with hearing losses cannot hear low-level sounds. In addition, they often complain that audible sounds do not have a comfortable loudness, lack clarity, and are difficult to hear in the presence of other sounds. In particular, they have difficulty understanding speech in background noise. The mechanisms underlying these complaints are not completely understood, but hearing losses are known to alter many aspects of auditory processing. This presentation highlights alterations in audibility, loudness, pitch, spectral and temporal processes, and binaural hearing that may result from hearing losses. The changes in these auditory processes can vary widely across individuals with seemingly similar amounts of hearing loss. For example, two listeners with nearly identical thresholds can differ in their ability to process spectral and temporal features of sounds. Such individual differences make rehabilitation of hearing losses complex. [Work supported by NIH/NIDCD.]

2:35

2pED4. An introduction to hearing aids. Ole Dyrland (GNResound, Maarkaervej 2A, DK-2630 Taastrup, Denmark, odyrlund@gnresound.dk)

This presentation reviews hearing-aid development from analog to advanced digital technology. A basic hearing aid consists of a microphone, an amplification circuit that provides a gain that varies with frequency to accommodate variations in hearing loss with frequency, and a small earphone. In recent years, hearing aid technology has developed rapidly. Digital hearing aids have become commonplace and their share of the marketplace is increasing rapidly. Therefore, the main focus of this talk is signal-processing schemes in advanced digital hearing aids, including microphones with digitally controlled directional characteristics, wide-dynamic-range compression in multiple channels that allow the compression characteristics to vary with frequency, noise reduction, and feedback cancellation. Each of these signal-processing functions help address the needs of individuals with hearing losses.

3:00

2pED5. Hearing aids: Do they help and, if so, how does one know? Larry E. Humes (Dept. of Speech & Hearing Sci., Indiana Univ., Bloomington, IN 47405-7002, humes@indiana.edu)

For those individuals with sensorineural hearing loss, ranging from mild to severe in degree, the conventional hearing aid is the most appropriate rehabilitative device available. Despite the fact that such devices have been available commercially for over 60 years, until recently, relatively little research has been directed at evaluating the effectiveness of these rehabilitative devices. How does one evaluate the effectiveness of a hearing aid as a rehabilitative device? Should effectiveness be based on the relative improvement in communication with and without the hearing aid, typically referred to as hearing-aid benefit, the satisfaction of the consumer with the device, or simply whether and how much the hearing aid is used? How are these aspects of hearing-aid effectiveness or outcome measured? Are the measures of hearing-aid outcome related to one another? What evidence is there regarding the effectiveness of contemporary hearing aids? Recent research regarding these and other related questions will be reviewed in this presentation. [Work supported, in part, by NIA.]

3:25

2pED6. Beyond the hearing aid: Assistive listening devices. Alice E. Holmes (Dept. of Communicative Disord., Univ. of Florida, P.O. Box 100174, Gainesville, FL 32610, aholmes@hp.ufl.edu)

Persons with hearing loss can obtain great benefit from hearing aids but there are many situations that traditional amplification devices will not provide enough help to ensure optimal communication. Assistive listening and signaling devices are designed to improve the communication of the hearing impaired in instances where traditional hearing aids are not sufficient. These devices are designed to help with problems created by listening in noise or against a competing message, improve distance listening, facilitate group conversation (help with problems created by rapidly changing speakers), and allow independence from friends and family. With the passage of the Americans with Disabilities Act in 1990, assistive listening devices (ALDs) are becoming more accessible to the public with hearing loss. Employers and public facilities must provide auxiliary aids and services when necessary to ensure effective communication for persons who are deaf or hard of hearing. However many professionals and persons with hearing loss are unaware of the various types and availability of ALDs. An overview of ALDs along with a discussion of their advantages and disadvantages will be given.

Contributed Paper

3:30

2pED7. Teaching hearing science to undergraduate nonscientists.

Ernest M. Weiler (CSD, ML #394, College of Allied Health Sci., Univ. of Cincinnati, Cincinnati, OH 45267-0394, ernest.weiler@uc.edu), Suzanne Boyce, and Joseph Steger (Univ. of Cincinnati, Cincinnati, OH)

For those students interested in potential clinical careers in Speech Pathology, or Audiology, a knowledge of some of the scientific bases is important, but should not create a distaste for science. The authors have addressed themselves to these goals: (1) calculation of period, Hz, summation of two sine waves, phase and dB; (2) anticipating undergraduate Speech Science; (3) simple examples of hearing pathology; and (4) basic psycho-acoustical issues. The classic material of Harry Helson was used to

elucidate issues of context in experimental science, and that of S.S. Stevens was used to exemplify psycho-acoustical formulas of common use. Four texts that have been tried on approximately 200 students were evaluated. Surprisingly, the best provided the fewest formulas, short study questions with answers, good examples, and a list of common terms. The next best was aimed at slightly more advanced students, but each chapter contained introductory material, examples, and definitions suitable for naïve undergraduates. The least satisfactory text provided excerpts of technical material with abrupt transitions, no examples, and only part of the definitions needed for the naïve student. Perhaps the most difficult teaching issue is to avoid demanding graduate-level science from those undergraduates with clinical aspirations.

Session 2pPA**Physical Acoustics: Sono(con)-fusion II: Evaluating the Chances and Claims of Bubble Fusion**

D. Felipe Gaitan, Cochair

Impulse Devices, Inc., 12731-A Loma Rica Drive, Grass Valley, California 95945

R. Glynn Holt, Cochair

Aerospace and Mechanical Engineering, Boston University, 110 Cummings Street, Boston, Massachusetts 02215

Thomas J. Matula, Cochair

*Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105***Invited Papers****1:45**

2pPA1. Sonoluminescence and multi-bubble cavitation phenomena for selected research and industrial applications. Larry Greenwood, Khri Olsen, Morris Good, Leonard Bond (PNNL, P.O. Box 999, P7-22, Richland, WA 99352), Gerald Posakony, Timothy Peters, David Baldwin, Dennis Wester, and Salahuddin Ahmed (PNNL, Richland, WA 99352)

Single bubble sonoluminescence (SBSL), multi-bubble sonoluminescence (MBSL), multi-bubble sonochemiluminescence (MB-SCL) and other high power ultrasound cavitation and noncavitating ultrasound process stream interaction phenomena are known to produce a wide range of both physical and chemical effects that depend upon the system and operating conditions employed. Three interacting regimes are under investigation (a) high power and high frequency (including noncavitating systems), (b) single bubble resonance/sonoluminescence and (c) multi-bubble high power sonochemical processing. In all cases these involve various reactors, including possible schemes for continuous material feeding and processing for selected chemical, nonaqueous fluids and biological research and industrial applications. High power sonochemical and noncavitating ultrasound processing applications and a review of literature pertaining to the potential of high power processing, including fusion are discussed. Work includes the investigation of acoustic fields in reactors, characterization of sonoluminescence spectra, the investigation of system parameters to control maximum bubble temperature and pressure, and acoustic energy partition into light and acoustic emission/shock waves. Effects of various chemical systems on multi-bubble luminescence are being investigated and will be reported. Work to date has emphasized the evaluation of both single and multi-bubble sonoluminescence, spectral measurements, acoustic emission measurements and the observation of a continuous bubble feed phenomenon.

2:00

2pPA2. Basic physics boundary conditions of acoustically driven, inertial confinement fusion. Lawrence Forsley (JWK Intl. Corp., Ste. 800, 7617 Little River Turnpike, Annandale, VA 22003, lforsley@jwk.com), Robert August, Robert Whitlock (Naval Res. Lab., Washington, DC), and Jacques Deletraz (LLE, Rochester, NY 14623)

This paper defines boundary conditions derived from the basic physics of fusion and applies them to both laser driven and acoustically driven Inertial Confinement Fusion (ICF). Several experimental and theoretical papers, in addition to some patents, hold open the promise of acoustically driven ICF. There are several factors common to both drivers that must be taken into account. In particular, it has been observed in laser driven ICF plasmas that criteria on the ion temperature, the confinement time, the core density, and the minimum core radius must be satisfied to achieve fusion. The relationship of these criteria to acoustically driven inertial confinement will be discussed.

2:25

2pPA3. Observing sonoluminescent UV photons below the water cutoff. Robert August (NRL, Washington, DC 20375-5321, robert.august@nrl.navy.mil), Lawrence Forsley (JWK Intl. Corp., Annandale, VA 22003), and Robert Whitlock (NRL, Washington, DC 20375-5321)

This paper presents sonoluminescent UV photon data observed in water below the water cut-off (200 nm) with hermetically sealed VUV photodiodes. Carbon and thin film metal filters provide both the encapsulation against water as well as the band-pass filtering of the diodes allowing spectra to be inferred. No previous data exists in this spectral region due to an effectively infinite absorption of photons after a few microns of water. The difficulty of these measurements is amplified by the requirement to use a detector normally intended for use in a high vacuum system. These measurements are important because the spectral shape in this region is unknown. We expect from published measurements at longer wavelengths that the photon fluence in this region should be high. These measurements may allow a differentiation between several competing sonoluminescence theories.

2pPA4. Update and clarifications on experimental studies for nuclear emissions during acoustic cavitation. Rusi P. Taleyarkhan, C. D. West, J. S. Cho (Oak Ridge Natl. Lab., Oak Ridge, TN 37831), R. T. Lahey, R. C. Block (Rensselaer Polytechnic Inst., Troy, NY), and R. I. Nigmatulin (Russian Acad. of Sci., Russia)

A seminal discovery related to detection of nuclear emissions during acoustic inertial confinement fusion with deuterated acetone has been reported in *Science* (3/8/2002 issue). Nuclear emissions we measured included 2.5-MeV neutrons and tritium as would be expected from deuterium–deuterium nuclear fusion. These unmistakable statistically significant signatures were measured under conditions commensurate with degassed rapid condensation-induced implosion conditions only with the test fluid deuterated acetone. In these experiments bubble clusters are nucleated in tensioned degassed liquids with neutrons at the nanoscale level and are then made to grow by a factor of $\sim 100\,000$ in size to the mm scale prior to implosive collapse. Similarly conducted control experiments with natural acetone did not result in any statistically significant nuclear emissions. Shock code simulations (discussed in a companion talk) corroborated these observations and provided insights into the physics of the overall process. Since the recent announcement of this discovery several world-wide researchers have contacted the authors for further clarifications in a variety of areas. The presentation will discuss these issues and questions, and will provide relevant explanations with supporting evidence.

3:30

2pPA5. Nuclear fusion in collapsing bubbles—Is it there? An attempt to repeat an experiment that reported d–d fusion in bubble collapse induced by cavitation in deuterated acetone. Dan Shapira and Mike Saltmarsh (Phys. Div., Oak Ridge Natl. Lab., P.O. Box 2008, M.S. 6368, Oak Ridge, TN 37831-6368)

The experiment of Taleyarkhan *et al.* [*Science* **295**, 1686 (2002)] has been repeated [D. Shapira and M. Saltmarsh, *Phys. Rev. Lett.* **89**, 104302 (2002)] in an attempt to detect the emission of neutrons from d–d fusion during bubble collapse in deuterated acetone. Using the same apparatus and method for bubble seeding and cavitation but a more sophisticated data acquisition system, and a large liquid scintillator detector we find no evidence for 2.5 MeV neutron emission correlated with sonoluminescence from the collapsing bubbles. Any neutron emission that might occur is at least four orders of magnitude smaller than that necessary to explain the tritium production reported in Taleyarkhan *et al.* as being due to d–d fusion. We demonstrate that the proper allowance for random coincidence rates in such experiments requires the simultaneous measurement of the complex time-varying singles rates.

4:15–4:30 Break

4:30–5:30

Panel Discussion

TUESDAY AFTERNOON, 29 APRIL 2003

ROOM 206, 1:00 TO 4:00 P.M.

Session 2pPP

Psychological and Physiological Acoustics: Pitch, Temporal Effects and Hearing Impairment (Poster Session)

Jennifer Lentz, Chair

Department of Speech and Hearing Science, Indiana University, 200 South Jordan, Bloomington, Indiana 47405

Contributed Papers

All posters will be on display from 8:00 a.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. To allow for extended viewing time, posters will be on display beginning at 8:00 a.m.

2pPP1. Simulations of cochlear implant hearing using filtered harmonic complexes: Implications for concurrent sound segregation and pitch perception. John M. Deeks and Robert P. Carlyon (MRC Cognition & Brain Sci. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK)

We studied concurrent sound segregation using cues similar to those available to cochlear implant listeners. Sixteen normally hearing subjects listened to mixtures of IHR sentences, processed using a simulation of implant hearing. Target sentences were bandpass filtered into six frequency bands between 1020 and 5000 Hz. The amplitude envelope in each band modulated a filtered, alternating-phase harmonic complex with $F_0 = 40$ or 70 Hz. Each complex resembled a pulse train with pulse rate

$= 2F_0$ and was filtered in the same way as the speech band that modulated it. The complexes consisted only of unresolved harmonics, whose pitch is processed in a way similar to that of electric pulse trains in implant hearing. The interferer was 3.5 s of time-reversed speech, processed in the same way as the target. After processing, targets were added to interferers at a 1.2-s delay, with SNR = 9 dB. The target and interferer used complexes with either the same or different F_0 . Using a different F_0 for the interferer benefited performance when the target $F_0 = 70$ Hz, but had no effect when the target $F_0 = 40$ Hz. Modifying the scheme by processing one sentence on only the odd-numbered channels and the other on the even-numbered channels impaired performance in all conditions.

2pPP2. The effects of simulated cochlear-implant processing on F_0 discrimination. Michael K. Qin and Andrew J. Oxenham (MIT Res. Lab. of Electron. and Harvard-MIT Div. of Health Sci. and Technol., SHBT Prog., 3 Ames St., Cambridge, MA 02139-4307)

Fundamental frequency (F_0) information has long been thought to play an important role in perceptually segregating simultaneous and non-simultaneous sources. The ability to discriminate different F_0 's is thought to depend primarily on fine-structure information, in particular the information carried in peripherally resolved, lower-order harmonics. Current cochlear-implant users do not have access to such cues. While implant-processed stimuli can carry some periodicity information in the stimulus envelopes, the F_0 discriminability associated with envelope periodicity is rather weak. This study investigated F_0 discrimination of complexes with F_0 's of 130 and 220 Hz in normal-hearing listeners using noise-excited vocoders to simulate cochlear-implant processing. F_0 difference limens of tone complexes were measured as a function of processing condition (1, 4, 8, 24, 40 channels, and unprocessed) and environment (tone complex alone and in reverberation). Performance with vocoder processing was poorer than in unprocessed conditions, despite the relatively good frequency resolution with the highest number of channels. Vocoder processing was particularly detrimental to F_0 discrimination in the reverberant conditions. The detrimental effects of vocoder processing may be due to the elimination of temporal fine-structure information or the poorer spectral representation of the lower-order harmonics. [Work supported by NIDCD Grant No. R01 DC05216.]

2pPP3. Viability of spectral enhancement with harmonic stimuli. Jeffrey J. DiGiovanni (W235 Grover Ctr., School of Hearing Speech and Lang. Sci., Ohio Univ., Athens, OH 45701)

Loss of spectral resolution is an established consequence of sensorineural hearing loss. Traditional hearing aid design includes amplification and compression. These do not, however, account for the loss in frequency resolution. Recently, spectral enhancement processing has been designed to at least partially restore aspects of frequency resolution. The critical feature of this design is to increase the peak to trough ratio of the speech spectrum. These have been implemented with mixed success [e.g., Miller *et al.* (1999); Franck *et al.* (1999)]. More recently, DiGiovanni *et al.* (2002) showed promising results for normal and hearing-impaired subjects with psychophysical noise stimuli. The goal of this study was to expand these results to harmonic stimuli while adding peaks at fixed formant places within the spectrum. In that regard, subjects listened in two psychophysical experiments: detecting an F_2 -like spectral increment in a broadband harmonic complex and detecting the increment with an additional fixed formant peak added at an appropriate F_1 place. Preliminary results show that normally hearing subjects have an improved ability to detect a narrowband tone complex when there is a spectral decrement at frequencies adjacent to the increment. These results are further support that the idea of spectral enhancement is viable.

2pPP4. Neuroimaging of speech recognition under conditions of spectral reduction and frequency upshift. C.-Y. Peter Chiu (Dept. of Psych. & Dept. of Commun. Sci. and Disord., 401A Dyer Hall, Univ. of Cincinnati, Cincinnati, OH 45229-0376, peter.chiu@uc.edu)

In the current study explored the cortical dynamics of speech recognition, given spectral reduction and frequency upshifts, using functional MRI. Subjects with normal hearing either rested or listened to speech under different conditions. In the 8-channel condition, natural speech was processed by an 8-channel sinewave vocoder to remove its fine spectral details [Shannon *et al.*, J. Acoust. Soc. Am. **104**, 2467 (1998)]. In the upshifted condition, the carrier center frequency of each of the 8 channels

was further shifted upward in frequency from the corresponding analysis band by "6 mm" in cochlear frequency space [Fu and Shannon, J. Acoust. Soc. Am. **105**, 1889 (1999)]. All subjects received a brief practice session with the speech stimuli prior to scanning. In Experiment 1, subjects listened to nonmonosyllabic words and pressed a key whenever they heard a concrete noun. In Experiment 2, subjects listened to high context sentences (SPIN) and pressed a key whenever they recognized all the words in a particular sentence. Preliminary data suggested that, compared to rest, all speech conditions evoked comparable activities in largely similar sets of bilateral superior temporal regions, with relatively minor differences between words and sentences. Activation appeared to be least diffuse in the natural speech condition.

2pPP5. Learning to recognize speech that is spectrally reduced and frequency upshifted. Marie E. McCabe (Dept. of Psych., Dyer Hall, ML 0376, Univ. of Cincinnati, Cincinnati, OH 45229-0376, mariemccabe@yahoo.com) and Peter Chiu (Univ. of Cincinnati, Cincinnati, OH 45229-0376)

The current study explored to what extent training could ameliorate the deleterious effect of large frequency upshifts in spectrally reduced speech. During each training session, subjects attempted recognition of IEEE sentences spoken by a single talker once and received feedback for their responses. Training sentences were processed to simulate an 8-channel CIS cochlear implant processor with a "6 mm frequency upshift" [Fu and Shannon, J. Acoust. Soc. Am. **105**, 1889 (1999)]. Three test sessions were administered to all subjects to assess recognition of sentences (IEEE and HINT), consonants (/aCa/), and vowels (in /hVd/ and /bVt/ contexts) pre-, post-, as well as at the mid-point of training. Four processing conditions (i.e., unprocessed, 8-channel-unshifted, 8-channel-upshifted, and 8-channel-upshifted-and-compressed) were tested for each type of materials. Preliminary data suggest that performance improved for most subjects during training, but there were substantial individual differences in learning rates and asymptotic performance levels. Vowels were more difficult to recognize and showed smaller training-related gains when compared to consonants and sentences. The rank ordering of recognition performance was consistent among the four processing conditions (unshifted best; upshifted-and-compressed intermediate; upshifted worst) for all measures. Data comparing the efficacy of an alternative training method will also be presented.

2pPP6. Estimates of pitch strength for musicians and nonmusicians. Marsha G. Clarkson, Cynthia M. Zettler, Michelle J. Follmer, Margaret Faulk (Dept. of Psych., Georgia State Univ., Atlanta, GA 30303), and Michael J. Takagi (Monash Univ., Melbourne, Australia)

To measure the strength of the pitch of iterated rippled noise (IRN), 19 adults were tested in an operant conditioning procedure. Seven adults had music training and currently played an instrument; 12 adults had no training and did not currently play an instrument. To generate IRN, a 500-ms Gaussian noise stimulus was delayed by 5 or 6 ms (pitches of 200 or 166 Hz) and added to the original for 16 iterations. IRN stimuli having one delay were presented repeatedly. On signal trials the delay changed for 6 s. Stimulus level roved from 63-67 dBA (background of 28 dBA). Adults learned to press a button when the stimulus changed. Testing started with IRN stimuli having 0-dB attenuation (i.e., maximal pitch strength). Stimuli having weaker pitches (i.e., progressively greater attenuation applied to the delayed noise) followed. Strength of pitch was quantified as the maximum attenuation for which pitch was 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 similar for both nonmusically trained (18.70 dB) and musically trained adults (18.73 dB).

2pPP7. Further tests of the “two pitch mechanisms” hypothesis. Christophe Micheyl and Andrew Oxenham (Res. Lab. of Electronics, MIT, Bldg. 36-797, Cambridge, MA 02139-4307)

To further investigate the hypothesis that F_0 is encoded via different mechanisms for resolved and unresolved harmonics, we measured F_0 difference limens (DLF0's) between two groups of harmonics that were both resolved, both unresolved, or differed by resolvability (one resolved, one unresolved). In a first experiment, the complexes were filtered in the same or a different frequency region and presented sequentially; listeners had to indicate which had the higher pitch. In the second experiment, two pairs of simultaneous complexes filtered in a different region were presented sequentially; subjects had to indicate the interval containing the pair of complexes with different F_0 's. To reduce possible pitch-pulse-asynchrony cues, the starting phase of the entire complex was randomized across spectral regions in. The results of the first experiment showed that for tracks where the higher F_0 was in the higher spectral region (consistent pitch-timbre cues), DLF0's were consistently smaller for same-resolvability than for different-resolvability complexes. On the other hand, the preliminary results of the second experiment, using simultaneous complexes, do not provide clear support for the hypothesis that DLF0's for resolved versus unresolved comparisons are impeded by “translation” noise between the output of the two pitch mechanisms. [Work supported by NIDCD R01DC05216.]

2pPP8. A neural network model of harmonic detection. Clifford F. Lewis (Dept. of Psych., Kent State Univ., P.O. Box 5190, Kent, OH 44242-0001, clewis@kent.edu)

Harmonic detection theories postulate that a virtual pitch is perceived when a sufficient number of harmonics is present. The harmonics need not be consecutive, but higher harmonics contribute less than lower harmonics [J. Raatgever and F. A. Bilsen, in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford, 1992), pp. 215–222; M. K. McBeath and J. F. Wayand, *Abstracts of the Psychonom. Soc.* **3**, 55 (1998)]. A neural network model is presented that has the potential to simulate this operation. Harmonics are first passed through a bank of rounded exponential filters with lateral inhibition. The results are used as inputs for an autoassociator neural network. The model is trained using harmonic data for symphonic musical instruments, in order to test whether it can self-organize by learning associations between co-occurring harmonics. It is shown that the trained model can complete the pattern for missing-fundamental sounds. The Performance of the model in harmonic detection will be compared with experimental results for humans.

2pPP9. Effects of modulation phase on profile analysis in normal-hearing and hearing-impaired listeners. Deanna Rogers and Jennifer Lentz (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47408, jilentz@indiana.edu)

The ability to discriminate between sounds with different spectral shapes in the presence of amplitude modulation was measured in normal-hearing and hearing-impaired listeners. The standard stimulus was the sum of equal-amplitude modulated tones, and the signal stimulus was generated by increasing the level of half the tones (up components) and decreasing the level of half the tones (down components). The down components had the same modulation phase, and a phase shift was applied to the up

components to encourage segregation from the down tones. The same phase shift was used in both standard and signal stimuli. Profile-analysis thresholds were measured as a function of the phase shift between up and down components. The phase shifts were 0, 30, 45, 60, 90, and 180 deg. As expected, thresholds were lowest when all tones had the same modulation phase and increased somewhat with increasing phase disparity. This small increase in thresholds was similar for both groups. These results suggest that hearing-impaired listeners are able to use modulation phase to group sounds in a manner similar to that of normal listeners. [Work supported by NIH (DC 05835).]

2pPP10. Perceptual learning in frequency discrimination and amplitude-modulation rate discrimination, and generalization to fundamental frequency discrimination. Nicolas Grimault, Christophe Micheyl (UMR, CNRS 5020 UCBL1, 50 av Tony Garnier, 69366 Lyon, Cedex 07, France, nicolas.grimault@olfac.univ-lyon1.fr), Robert P. Carlyon (MRC Cognition and Brain Sci. Unit, Cambridge CB2 3EF, UK), Sid P. Bacon (Arizona State Univ., Tempe, AZ 85287-1908), and Lionel Collet (CNRS 5020 UCBL1, 69366 Lyon, Cedex 07, France)

Fifteen subjects were trained during twelve 2-h sessions in either frequency discrimination with pure tones, or amplitude-modulation rate discrimination of noise bands. Thresholds for the discrimination of pure-tone frequency, harmonic complex tone fundamental frequency, and amplitude-modulation rate were measured before, during, and after training. Comparison of pre- and post-training thresholds revealed significant improvements in all conditions in both subjects trained in frequency discrimination and subjects trained in modulation rate discrimination. Training in frequency discrimination resulted in larger improvements in fundamental frequency discrimination when the test complexes contained resolved harmonics than when they were composed of unresolved harmonics. Training in modulation rate discrimination did not result in larger fundamental frequency discrimination improvements for unresolved than for resolved harmonics. The implications for models of pitch perception are discussed.

2pPP11. The effects of three temporal cues on the detection of increments and decrements in intensity. Yang-soo Yoon and David M. Gooler (Dept. of Speech and Hearing Sci., Univ. of Illinois at Urbana-Champaign, Champaign, IL 61820, yyoons5@uiuc.edu)

This study investigated the effect of three temporal cues: interonset interval, interpulse interval, and pulse duration, on auditory intensity detection. Eight adults with normal hearing served as subjects. The level detection (target increment or decrement) ability of listeners was measured in a sequential oddball paradigm where targets occurred randomly. The experiment used spectrally identical, broadband noise bursts as standard (30 dB above threshold to a train of standard pulses) and target pulses (1, 2, and 3 dB *re*: standard pulse level). At each target level performance was measured for different pulse interonset intervals (20 ms to 180 ms), interpulse intervals (10 ms to 90 ms), and pulse duration (10 ms to 90 ms, 0.5 ms rise/fall time). Overall, the results showed that the performance of the intensity detection was more influenced by the interpulse interval than the interonset interval. Also, performance was dependent on pulse duration since a poorer performance with longer interpulse intervals could be improved by increasing pulse duration. Performance at different inter-pulse intervals may reflect different perceptual strategies for comparing acoustic events that occur in a sequence.

2pPP12. Effect of duration and inter-stimulus interval on auditory temporal order discrimination in young normal-hearing and elderly hearing-impaired listeners. Mini M. Narendran and Larry E. Humes (Dept. of Speech and Hearing Sci., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405, mnarendr@indiana.edu)

Increasing the rate of presentation can have a deleterious effect on auditory processing, especially among the elderly. Rate can be manipulated by changing the duration of individual components of a sequence of sounds, by changing the inter-stimulus interval (ISI) between components, or both. Consequently, when age-related deficits in performance appear to be attributable to rate of stimulus presentation, it is often the case that alternative explanations in terms of the effects of stimulus duration or ISI are also possible. In this study, the independent effects of duration and ISI on the discrimination of temporal order for four-tone sequences were investigated in a group of young normal-hearing and elderly hearing-impaired listeners. It was found that discrimination performance was driven by the rate of presentation, rather than stimulus duration or ISI alone, for both groups of listeners. The performance of the two groups of listeners differed significantly for the fastest presentation rates, but was similar for the slower rates. Slowing the rate of presentation seemed to improve performance, regardless of whether this was done by increasing stimulus duration or increasing ISI, and this was observed for both groups of listeners. [Work supported, in part, by NIA.]

2pPP13. Difference limen for perception of aspiration noise. Rahul Shrivastav and Christine M. Sapienza (Dept. of Commun. Sci. & Disord., Univ. of Florida, Dauer Hall, Gainesville, FL 32611)

The relationship between the vocal acoustic signal and the perception of voice quality has been the subject of much research. An increase in the intensity of aspiration noise is associated with a perception of greater breathiness in voices [Klatt and Klatt (1990); Childers (1993)]. Shrivastav (2001) found that subjective ratings of breathiness could be better predicted using measures calculated from an auditory spectrum of the vocal acoustic signal. These measures were found to vary with changes in the spectral slope as well as the intensity of aspiration noise in voice. The aim of the present experiment was to determine the smallest change in the intensity of aspiration noise that could be perceived by listeners. Six voice continua, with increasing intensity of aspiration noise, were generated using the Klatt synthesizer. Five listeners participated in an adaptive listening test, where these stimuli were presented in pairs and listeners were asked to identify the stimuli in each pair as being same or different. Listener responses were used to determine the difference limen for changes in intensity of aspiration noise in voice. These findings will help understand the perception of breathy voice quality and to develop objective measures for its quantification.

2pPP14. The aging middle ear: Wideband energy reflectance measurements. M. Patrick Feeney (Dept. of Otolaryngol., Head and Neck Surgery, V. M. Bloedel Hearing Res. Ctr., Univ. of Washington, Box 357923, Seattle, WA 98195) and Chris A. Sanford (Univ. of Washington, Seattle, WA 98195)

Several anatomical studies have documented aging effects in the human middle ear. However, efforts to study the effect of aging using both low-frequency and multifrequency tympanometry to 2000 Hz have been inconclusive. This study examined energy reflectance at ambient pressure from 250 to 10,080 Hz in 40 young ($M=22$ years) and 34 elderly adults ($M=72$ years). All subjects had normal 226 Hz tympanometry and audio-

metric air-bone gaps of 10 dB or less. Reflectance measurements were obtained in a sound-treated booth using a digitally-generated wideband chirp as the probe stimulus delivered by a receiver in an ER-10C microphone. Each reflectance measurement consisted of a time-waveform average of the microphone response to 8 chirps. Three such one-third-octave reflectance responses were averaged to obtain an estimate of middle ear reflectance for one ear of each subject. The average reflectance for the elderly subjects was significantly lower from 794 to 2000 Hz with a maximum difference of 15% at 1260 Hz. A reflectance minimum occurred at 4000 Hz for both groups, but was about 15% lower for the young group. Results will be compared to published adult data using similar systems. [Work supported by the NIDCD Grant No. DC04129.]

2pPP15. Evidence of peripheral nonlinearity in psychometric function slopes of forward-masked tones at 250 and 4000 Hz. Kim S. Schairer and Walt Jesteadt (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

Results from this laboratory suggest that the compressive nonlinearity of the basilar membrane is reflected in slopes of psychometric functions (PFs) for forward-masked tones of 4000 Hz. Briefly, as the signal level at the threshold increases, slopes of PFs decrease. Many other behavioral measures have suggested significant nonlinearity at high frequencies, but results are less consistent at low frequencies. The purpose of the current study was to use a PF slope to investigate nonlinearity using low- and high-frequency stimuli in the same normal-hearing adults. Signals were 250 or 4000 Hz, 10-ms (5-ms rise/fall) duration, presented with a 10-ms delay. The on-frequency forward-maskers were 200-ms (2-ms rise/fall), presented at levels of 30-, 50-, 70-, and 90-dB SPL in separate conditions. A two-track, two-interval-forced-choice adaptive procedure was used, with decision rules to estimate 71% correct on one track, and 87% on the other track. Data from both tracks were combined to estimate PF slopes. PF slopes were steeper in low-threshold conditions and shallower in high-threshold conditions for both frequencies. These results suggest a significant nonlinearity at 250 Hz. [Work supported by NIDCD.]

2pPP16. Confidence ratings and awareness measures in word recognition testing. Edward L. Goshorn and Jennifer D. Goshorn (Speech Dept., P.O. Box 3165 Tech Station, Louisiana Tech Univ., Ruston, LA 71272)

Word recognition tests primarily use percent correct to measure performance. Additional information may be gained by analyzing awareness of accurate perceptions (AA), awareness of errant perceptions (AE), composite awareness (AC), and awareness symmetry (AS). Awareness measures were derived from subjects' assignment of confidence ratings to a two-item multiple choice response test by designating "YES" or "NO" that their chosen response is accurate. Each response/confidence rating was categorized as a hit, miss, false alarm, or correct rejection. Awareness equations were: $AA = \text{hits}/(\text{hits} + \text{misses})$; $AE = \text{correct rejections}/(\text{correct rejections} + \text{false alarms})$; $AC = \text{SQRT}(AA^2 + AE^2)$; $AS = 0.707(AA - AE)$. Thus, AC is the vector to Cartesian coordinates AA, AE; AS is the distance of this point from a diagonal representing symmetrical awareness. Word recognition and awareness was investigated under two signal-to-noise ratios (3 and 6 dB). The Diagnostic Rhyme Test was presented at 50 dBHL to eight normal-hearing adults. Six replicates were obtained. Awareness measures provided additional performance information. Percent correct increased significantly as signal-to-noise ratio improved, but AE decreased and AS did not change significantly.

Session 2pSA

Structural Acoustics and Vibration and Noise: Damping and Absorption

Kenneth A. Cunefare, Chair

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

Contributed Papers

2:30

2pSA1. Optimization of a state-switched absorber applied to a vibrating continuous system. Mark Holdhusen and Kenneth Cunefare (The George W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

A state-switched device is a device that can instantaneously change one or more of its dynamic properties. A vibration absorber that can switch between discrete stiffnesses is termed a state-switched absorber. Using a maximum work extraction switching rule, the state-switched absorber has been shown to increase performance in lumped mass systems as compared to a classical tuned vibration absorber. Because the state-switched absorber can retune its resonance frequency during the response of the system, the bandwidth which the state-switched absorber is effective is larger than that of classical passive devices. The work at hand considers the effectiveness of vibration control using state-switched absorber attached to a continuous system subjected to a multifrequency excitation. The switching rule, tuning frequencies, and location of attachment on the continuous system are optimized for the state-switched absorber to achieve the lowest kinetic energy in the insulated system. Tuning and location optimization is also done for a classical tuned vibration absorber. The optimized performance of the state-switched absorber is compared to the optimized performance of the tuned vibration absorber for a number of different forcing cases.

2:45

2pSA2. A proposed definition for the induced noise control parameter. G. Maidanik and K. J. Becker (NSWCCD (DTMB), Code 7030, 9500 MacArthur Blvd., West Bethesda, MD 20817)

The response of an externally force-driven master dynamic system is examined under two conditions. In the first, the master dynamic system is in isolation and the (quadratic) response is stated in terms of its stored energy $E_0^0(\omega)$, where (ω) is the (angular) frequency. In the second, the master dynamic system is passively coupled to an adjunct dynamic system that is not externally driven. The response of the master dynamic system, that is subjected to the same external force drive, is changed due to the couplings. The change is from $E_0^0(\omega)$, in isolation, to $E_0(\omega)$, when coupled. The induced noise control parameter $\xi(\omega)$ is then defined: $\xi(\omega) = [E_0(\omega)/E_0^0(\omega)]$. The smaller $\xi(\omega)$ is, the more benefit, by the couplings, is accrued to the noise control of the master dynamic system. When $\xi(\omega)$ exceeds unity, by definition, an induced noise control reversal occurs. Illustrations of an induced noise control are provided in terms of a master dynamic system represented by an (harmonic) oscillator and an adjunct dynamic system represented by a set of satellite (harmonic) oscillators. The satellite oscillators are not coupled to each other; they are individually coupled to the master oscillator only. A coupling may include a combination of mass, stiffness, and gyroscopic elements.

3:00

2pSA3. Effect of burst parameters on automotive brake squeal suppression. Jeff Badertscher and Kenneth A. Cunefare (Georgia Inst. of Technol., 771 Ferst Dr., Atlanta, GA 30332, gtg437d@prism.gatech.edu)

Implementing a dither control signal with a 100% duty cycle is an effective means of suppressing automotive brake squeal. Dither control is a method by which high-frequency control efforts are introduced into a system to suppress a lower frequency disturbance. Dither is introduced to a brake by placing a piezoelectric stack actuator in the piston of a floating caliper brake. Burst mode dither control is characterized by duty cycles of less than 100%. A burst control signal of a specific duty cycle is also specified by the burst count and burst rate. Burst mode signals are shown to suppress brake squeal. This paper examines the nature of suppression and the effectiveness of burst mode dither control signals with varied burst parameters. An examination of the squeal response and dither control signal is used to examine the nature of suppression during bursting and dwell time. The amplitude of the control signal that is necessary to obtain full control of the system is used to assess control signal effectiveness.

3:15

2pSA4. Modeling the effect of bearing properties on the eigenvalues of a rotordynamic system. Ben B. Wagner and Jerry H. Ginsberg (Georgia Inst. of Technol., Woodruff School of Mech. Eng., 801 Ferst Dr. NW, Atlanta, GA 30332-0405)

Experimental data indicate that changes in bearing properties alter the natural frequency and damping ratios of rotordynamic systems [R. M. Baldwin, ASME 92-GT-53]. The work presented here examines the development of a mathematical model that will be used to investigate how natural frequency, modal damping ratio, shaft rotation rate, bearing clearance, and lubricant viscosity are related. The modeled system consists of a uniform, elastic, rotating shaft, which is supported by two plain journal bearings, and a single, rigid disk, which is concentrically welded to the shaft away from midspan. Standard lubrication theory is used to generate the stiffness and damping matrices for the bearings. A Ritz series expansion is used to generate the mass, shaft stiffness, and gyroscopic matrices describing the shaft and disk. The combined action of the bearings and shaft/disk system is described by matrices representing the effects of stationary and rotational inertia, stiffness, and internal and external damping. The radial clearance and lubricant viscosity are independently adjustable at each bearing, and the range of shaft speed for analysis is user-specified. A nonsymmetric generalized eigenvalue problem solver is used to calculate eigenvalues, whose real part is proportional to the modal damping ratio, and whose imaginary part is the natural frequency. For Structural Acoustics and Vibrations Best Student Paper Award. Submitted for Structural Acoustics and Vibrations Young Presenter Award.

2p TUE. PM

2pSA5. What is the best method for structural damping identification? Part 1: A survey of methods. Srikantha Phani and J. Woodhouse (Dept. of Eng., Univ. of Cambridge, Trumpington St., Cambridge CB2 1PZ, UK)

Identification of damping matrices from measured frequency response functions is a difficult and error-prone procedure. In practice, the assumption of proportional damping is frequently made without physical justification. Several researchers have suggested identification algorithms for the full damping matrix. A critical review is presented of algorithms which work in the frequency domain. Some novel approaches are included: one new algorithm is based on matrix perturbation theory, while others are hybrid combinations of earlier methods. The sensitivity of the algorithms to measurement noise and other commonly encountered difficulties in realistic measurements is addressed. Representative results from a simulation study comparing all the methods will be presented, leading to recommendations for the choice of identification algorithms in practice.

2pSA6. Boundary damping of flexural vibrations in beams and plates. Joel Garrelick (Applied Physical Sciences, Inc., 2 State St., Ste. 300, New London, CT 06320)

The damping of resonantly enhanced flexurally vibrating beams and plates is typically accomplished with surface treatments, viz. constrained or unconstrained damping layers. Alternatively, damping may be achieved at boundaries. In some sense this is a more fundamental approach in that for homogeneous plates, it is the boundaries that are solely responsible for resonant behavior. The theoretical performance of boundary treatments is unbounded, provided the treatment is full, that is continuous along the extent of the boundary. This is not the case however with partial coverage, where a portion of the boundary is left bare. The bounds for such treatments are explored in this paper. The treatment itself is defined in terms of a single bounce reflection coefficient (RC). It is found that for boundary damping in one dimension, viz., beams, the effective loss factor for individual modes is frequently invariant with either one or both ends treated, assuming RC constant. This is in contrast to the two-dimensional case, viz., thin plates, where analogous loss factor values are frequency dependent. Illustrative examples are presented and analyzed. [Work partially performed at CAA/Anteon Corp. and supported by NSWCCD and NSSC, Code 93R.]

2pSA7. What is the best method for structural damping identification? Part 2: Comparison of performance. Srikantha Phani and J. Woodhouse (Dept. of Eng., Univ. of Cambridge, Trumpington St., Cambridge CB2 1PZ, UK)

Damping identification in structural vibration from measured frequency response functions is a difficult task. Several methods have been proposed in the literature to identify the damping matrix. Some employ modal parameters such as natural frequencies, damping factors, and mode shapes, while others work directly from the frequency response function matrix. In practice, in the presence of noise and incompleteness of data, it is hard to guess how well any of these methods will perform. This paper reports a systematic simulation study to test the relative benefits/disadvantages of a wide range of approaches, existing and novel. Representative model systems have been chosen to test performance under various conditions of measurement noise, system complexity, damping type, modal overlap factor, and modal truncation. A very wide range of results is summarized by computing various norms of performance for all the methods studied. Recommendations are made for the best methods to extract reliable damping information from practical measurements.

2pSA8. Active noise control using damped resonant filters. Jesse B. Bisnette, Jeffrey S. Viperman, and Daniel D. Budny (Dept. of Mech. Eng., Univ. of Pittsburgh, 648 Benedum Hall, Pittsburgh, PA 15261)

Active Noise Control (ANC) has been found to work well at low frequencies, thereby complementing passive noise control techniques. There are two distinct topologies for ANC: feedforward and feedback. Feedforward implementations work through the superposition of a primary and secondary sound field while feedback systems work by augmenting the dynamics of an enclosed sound field to add damping. Model-based feedback control systems are cumbersome to implement due to typical high modal density and simple output feedback systems are complicated for two reasons. First, perfectly collocated sensor (microphone) and actuator (loudspeaker) pairs are difficult to achieve, and second, the dynamics of the loudspeaker are destabilizing. Lane and Clark (1998) have reported compensating loudspeaker dynamics over a bandwidth to produce an approximate volume velocity source and demonstrated rate feedback acoustic control. Here, an acoustic control method using second order filters, which is analogous to positive position feedback (PPF) control in structural systems, is presented. The phase response of the loudspeaker is accounted for by incorporating negative feedback with either band-pass or high-pass filters, and the technique appears to have the same tolerance for actuator dynamics as for the structural case. Alternatively, the approximate volume velocity source could also be used.

2pSA9. An active flow control scheme for the reduction of cavity pressure in flow-excited Helmholtz resonators. Jong Beom Park and Luc Mongeau (Ray W. Herrick Labs., School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907-2031)

Flows over cavities acting as flow-excited Helmholtz resonators feature the formation of strong vortices over the cavity orifice, which strength regulates the amplitude of the cavity pressure oscillations. Much is already known about this phenomenon, and models for predicting the magnitude and the frequency of the pressure oscillations are available. Various control devices and schemes have been developed to suppress the oscillations. In particular, active flow control using oscillated leading edge spoilers has been proposed as an alternative to passive spoilers, or flow injection systems. Little is known, however, about the detailed effects of such an active flow control system on the flow field. In the present study, a simple control loop was implemented to adjust the phase between the leading edge spoiler motion and the cavity pressure. The sound pressure inside a resonator was measured over a range of phase differences and spoiler motion amplitudes. It was found using flow visualization that optimally synchronized spoiler motion caused the vortices over the cavity orifice to be less concentrated. The suppression mechanisms were explained using a model based on vortex sound theory.

2pSA10. When things go crunch: Gap length effects on a magneto-rheological tunable vibration absorber. Anne-Marie Albanese and Kenneth Cunefare (Georgia Inst. of Technol., Atlanta, GA 30332)

A tunable vibration absorber (TVA), where the tunable element is a magneto-rheological (MR) elastomer spring, has demonstrated up to a 460% change on tuning frequency. A frequency increase is associated with a significant decrease in the springs static equilibrium length, referred hereto as a crunch. The crunch is caused by magnetic attractive forces across the spring. The spring, an iron-doped silicone gel, is placed between two halves of a low-carbon steel loop. One loop half behaves as an

absorber mass, and has approximately 200 turns of magnet wire around it. Driving current through this wire generates a magnetic flux around the steel path and through the MR spring. Beyond a threshold, the magnetic field induces a large frequency shift in the absorber, with a crunch observed across the elastomer. The crunch can occur when the static equilibrium length is unconstrained by geometry, thus a magnetic attractive

force shortens the spring length. The relationship between the magnitude of the crunch and the frequency shift will be presented. Additionally, the impact of different initial MR spring lengths on the frequency behavior will be considered. Finally, the frequency variability achievable by MR-spring-based TVAs with and without the crunch will be assessed.

TUESDAY AFTERNOON, 29 APRIL 2003

ROOM 205, 1:00 TO 5:20 P.M.

Session 2pSC

Speech Communication: Man and Machine in Cocktail Parties

Pierre L. Divenyi, Cochair

Speech and Hearing Research, VA Medical Center, 150 Muir Road, Martinez, California 94553

Daniel P. W. Ellis, Cochair

Department of Electrical Engineering, Columbia University, Mail C4712, 500 West 120th Street, Room 1312, New York, New York 10027

Chair's Introduction—1:00

Invited Papers

1:05

2pSC1. Some components of the “cocktail-party effect,” as revealed when it fails. Pierre L. Divenyi and Brian Gygi (Speech and Hearing Res., Veterans Affairs Northern California Health Care Systems and East Bay Inst. for Res. and Education, Martinez, CA 94553)

The precise way listeners cope with cocktail-party situations, i.e., understand speech in the midst of other, simultaneously ongoing conversations, has by-and-large remained a puzzle, despite research committed to studying the problem over the past half century. In contrast, it is widely acknowledged that the cocktail-party effect (CPE) deteriorates in aging. Our investigations during the last decade have assessed the deterioration of the CPE in elderly listeners and attempted to uncover specific auditory tasks, on which the performance of the same listeners will also exhibit a deficit. Correlated performance on CPE and such auditory tasks arguably signify that the tasks in question are necessary for perceptual segregation of the target speech and the background babble. We will present results on three tasks correlated with CPE performance. All three tasks require temporal processing-based perceptual segregation of specific non-speech stimuli (amplitude- and/or frequency-modulated sinusoidal complexes): discrimination of formant transition patterns, segregation of streams with different syllabic rhythms, and selective attention to AM or FM features in the designated stream. [Work supported by a grant from the National Institute on Aging and by the V.A. Medical Research.]

1:35

2pSC2. Monaural speech segregation. DeLiang Wang (Dept. of Computer & Information Sci. and Ctr. for Cognit. Sci., The Ohio State Univ., Columbus, OH 43210, dwang@cis.ohio-state.edu) and Guoning Hu (The Ohio State Univ. Biophys. Prog., Columbus, OH 43210)

Speech segregation from a monaural recording is a primary task of auditory scene analysis, and has proven to be very challenging. We present a multistage model for the task. The model starts with simulated auditory periphery. A subsequent stage computes midlevel auditory representations, including correlograms and cross-channel correlations. The core of the system performs segmentation and grouping in a two-dimensional time-frequency representation that encodes proximity in frequency and time, periodicity, and amplitude modulation (AM). Motivated by psychoacoustic observations, our system employs different mechanisms for handling resolved and unresolved harmonics. For resolved harmonics, our system generates segments—basic components of an auditory scene—based on temporal continuity and cross-channel correlation, and groups them according to periodicity. For unresolved harmonics, the system generates segments based on AM in addition to temporal continuity and groups them according to AM repetition rates. We derive AM repetition rates using sinusoidal modeling and gradient descent. Underlying the segregation process is a pitch contour that is first estimated from speech segregated according to global pitch and then adjusted according to psychoacoustic constraints. The model has been systematically evaluated, and it yields substantially better performance than previous systems.

2p TUE. PM

2pSC3. How does spatial hearing affect cocktail party conversations? Barbara G. Shinn-Cunningham (Boston Univ., 677 Beacon St., Boston, MA 02215, shinn@cns.bu.edu)

Although spatial cues are not a dominant factor in auditory scene analysis by humans, spatial separation of a target talker and an interfering sound source often increases the intelligibility of the target. A large portion of this improvement arises from a simple acoustic effect: separating the target and interferer generally increases the target-to-interferer energy ratio at one ear. When the target is near threshold, binaural processing provides an additional small, but important improvement. Finally, in conditions where both target and interferer are audible but are difficult to segregate, spatial separation can improve the ability to stream the sources and interpret the target message. The echoes and reverberation present in everyday environments influence all of these factors, altering the target-to-interferer energy ratio, the effectiveness of binaural processing, and the ability to stream the sources. A number of studies will be reviewed to demonstrate how these different factors influence speech intelligibility and affect cocktail party conversations in everyday environments. [Work supported by the Air Force Office of Scientific Research and the Alfred P. Sloan Foundation.]

2pSC4. Active audition for humanoid robots that can listen to three simultaneous talkers. Hiroshi G. Okuno (Dept. of Intelligence Sci. and Technol., Grad. School of Informatics, Kyoto Univ., Sakyo, Kyoto 606-8501, Japan, okuno@i.kyoto-u.ac.jp) and Kazuhiro Nakadai (Kitano Symbiotic Systems Project, JST, M-31 6-31-15 Jingumae, Shibuya, Tokyo 150-0001, Japan)

The direction-pass filter (DPF) separates sounds originating from a particular direction by using a pair of microphones embedded in each ear of humanoid robot. DPF first extracts harmonic structures from each channel, finds a corresponding pair on right and left channels, and then calculates their interaural phase difference (IPD) and interaural intensity difference (IID). These IPD and IID are matched with reference data obtained by HRTF or by the geometrical relation to determine the sound source direction. The direction obtained by face detection may be used as a candidate for the direction. Finally, all subbands from the direction are collected to synthesize a wave form by inverse FFT. The allowance of collection depends on the direction; narrow (10 deg) at center, while wide (30 deg) at the periphery. This property is called "auditory fovea" and is exploited by DPF actively to improve performance of sound source separation. In addition, a humanoid actively turns its head toward the speaker to listen better. Real-time DPF is implemented by distributed processing with five PCs. Preliminary experiments of active audition in speech recognition of three simultaneous utterances of digits in a normal room is also reported. [Work supported by JSPS.]

2pSC5. Machine recognition of sounds in mixtures. Daniel Ellis (Columbia Univ., 500 W. 120th St., Rm. 1312, New York, NY 10027) and Jon Barker (Sheffield Univ., Sheffield S1 4DP, UK)

A machine simulation of human auditory perception must be able to recognize and classify individual sound sources. The most successful technique for sound classification is the statistical pattern recognition approach employed in speech recognizers; however, in most practical cases, this approach assumes that the entire (monaural) signal represents the source to be classified. Realistic "cocktail-party" scenes, composed of multiple, overlapping sources with comparable energies, do not come close to meeting this assumption. A more workable assumption is to treat each time-frequency cell as representing a single source, and to use missing-data techniques to perform recognition using only a subset of the cells. This precludes the use of cepstral features (which depend on every frequency component), but is otherwise practical. The problem then becomes finding the "present data mask" that indicates which cells are to be considered during classification of a particular source. We will present a system based on these principles, with applications both to speech recognition in dynamic, noisy backgrounds, and also to nonspeech sounds such as alarms that can occur at very poor signal-to-noise ratios.

2pSC6. How is harmonicity used in grouping speech sounds? Chris Darwin (Exp. Psych., Univ. of Sussex, Brighton BN1 9QG, UK, cjd@biols.susx.ac.uk)

This paper asks how a common property of voiced speech sounds harmonicity is used by the auditory system to improve the perception of speech in the presence of simultaneous competing sounds. We present data from three different experimental paradigms concerned, respectively, with the combination of sounds across different ears, different frequency regions, and different times. The first set of experiments qualify the conclusion that sounds from the same harmonic series fuse into a single object when presented to different ears. The second impose limits on the ability of harmonicity to combine information across different frequency regions. The third demonstrate the utility of continuity of pitch (compared with a vocal-tract size) in maintaining attention to a single sound source. Elucidating the mechanisms by which we segregate speech from background sounds requires proper consideration both of the structure of the speech signal and of the auditory system through which it passes.

2pSC7. Scene analysis without spectral analysis? Alain de Cheveigne (Ircam-CNRS, 1 place Igor Stravinsky, Paris 75004, France, cheveign@ircam.fr)

Auditory scene analysis is often described in terms of grouping stimulus components. Components, once grouped, are assigned to one source or another [A. S. Bregman, *Auditory Scene Analysis* (MIT, Cambridge, MA, 2002)]. The actual grouping must operate on whatever representation is available within the auditory nervous system. An obvious hypothesis is that correlates of individual stimulus components are created by peripheral spectral analysis. However, peripheral frequency resolution is limited. The number of resolved partials is between 5 and 8 for a harmonic complex in isolation, but resolution must necessarily be less good for the interleaved components of concurrent sources. Source amplitudes are rarely equal, and partials of a weaker source must be particularly hard to resolve. The question is thus: given the paucity of resolved elements to group, how does the auditory system perform the grouping? A number of possibilities will be reviewed. One is that partials not resolved peripherally are somehow resolved centrally (a modern version of the "second filter" hypothesis). Another is that scene analysis does not operate by grouping resolved elements, but instead by modifying directly unresolved entities, for example by time-domain processing.

2pSC8. Issues in the use of acoustic cues for auditory scene analysis. Albert S. Bregman (Psych. Dept., McGill Univ., 1205 Doctor Penfield Ave., Montreal, QC H3A 1B1, Canada)

Issues concerning auditory scene analysis (ASA) raised by the previous speakers will be discussed: (1) Disorders of ASA in humans can tell us about the weighting of cues in ASA. (2) The apparent weakness of spatial cues for ASA may simply show that they interact strongly with other ASA cues (c.f., recent research in the author's lab). (3) The power of harmonic relations among partials as a grouping cue is not guaranteed, but depends on many other factors. (4) Abstract models of ASA may require the peripheral auditory system to carry out analyses that are questionable, based on current psychophysical and physiological findings. Is this where psychologists and computational ASA (CASA) modelers part company? (5) The "old-plus-new heuristic," one of the most potent ASA mechanisms, is neglected by existing CASA models. (6) The different roles of bottom-up and top-down processes (e.g., in "exclusive allocation" of sensory evidence) should be reflected in models. (7) Should the output of a CASA system be the reconstructed signal of a single source, as a front end to a recognition system, or should grouping mechanisms merely form an interacting part of a larger system that outputs a higher-level description (e.g., a series of words)?

TUESDAY AFTERNOON, 29 APRIL 2003

ROOM 203, 1:00 TO 2:35 P.M.

Session 2pSP

Signal Processing in Acoustics and Physical Acoustics: Subspace Methods for Acoustical Imaging II

Sean K. Lehman, Cochair

Lawrence Livermore National Laboratory, L-154, 7000 East Avenue, Livermore, California 94550

David H. Chambers, Cochair

Lawrence Livermore National Laboratory, L-154, P.O. Box 808, Livermore, California 94551-5508

Chair's Introduction—1:00

Invited Papers

2pSP1. A model subspace method for matched field source localization in a shallow-water acoustic waveguide. Peter M. Daly (SAIC Ocean Systems Operation, 1710 SAIC Dr., M.S. 1-11-15, McLean, VA 22102, peter.m.daly@saic.com)

Model subspace methods for parameter estimation can yield promising results when temporal or spatial variability dominate the model inputs. Application of Gaussian signal detection theory, summarized by Van Trees, is applied to the localization problem. Model inputs, such as the sound velocity profile in water, bottom properties, and array position, rather than being deterministic, are characterized using a Gaussian random vector. This, in turn, yields a random signal model which is represented as a vector space. Detection and localization via a maximum-likelihood estimator becomes a question of whether the received signal fits the subspace defined by the random signal model.

2pSP2. Interference cancellation in a waveguide using signal subspace decomposition. Paul Hursky, Michael B. Porter (Sci. Applications Intl. Corp., 10260 Campus Point Dr., San Diego, CA 92121), and John P. Ianniello (Sci. Applications Intl. Corp., Mystic, CT 06355)

When the presence of a strong interferer hides a quieter target, an eigenanalysis of the data covariance matrix (also called a principal component analysis) may separate the quieter target from the strong interferer. However, the strong interferer may occupy more than a single principal component when the data covariance matrix is estimated over a time interval during which the strong interferer has moved between multiple spatial resolution cells of the receive array. Also, when the sources to be distinguished are of comparable strength, the principal components fail to isolate the sources, forming mixtures of the sources instead. Results will be presented of applying this technique to the SWellEx-96 experiment data, recorded on both vertical and horizontal arrays, in a shallow water environment off the coast of San Diego. Projections of the principal eigenvectors onto beams and matched-field processing range-depth cells will be shown to indicate to what degree a principal component analysis isolates the different sources. Strategies for coping with source motion will be discussed.

Contributed Papers

2:05

2pSP3. Application of Hilbert space decomposition to acoustical inverse problems. Sean K. Lehman (Lawrence Livermore Natl. Lab., L-154, 7000 East Ave., Livermore, CA 94550)

In a recently developed theory, the forward integral acoustic scattering operator is cast into the formalism of a Hilbert space operator which projects the continuous space scattering object into the discrete measurement space. By determining the singular value decomposition (SVD) of the forward scattering operator, one obtains optimal, orthonormal bases for each of these spaces in the form of the singular vectors. In formulating the inverse scattering problem, it is best to express the unknown object distribution in terms of an expansion of the singular vectors. Using this expansion, the best reconstruction, in a LMS sense, can be obtained from measured scattered field data. We present reconstruction results using this new theory. [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.]

2:20

2pSP4. Acoustical imaging of spheres above a reflecting surface. David Chambers and James Berryman (Lawrence Livermore Natl. Lab., P.O. Box 808, L-154, Livermore, CA 94551)

An analytical study using the MUSIC method of subspace imaging is presented for the case of spheres above a reflecting boundary. The field scattered from the spheres and the reflecting boundary is calculated analytically, neglecting interactions between spheres. The singular value decomposition of the response matrix is calculated and the singular vectors divided into signal and noise subspaces. Images showing the estimated sphere locations are obtained by backpropagating the noise vectors using either the free space Green's function or the Green's function that incorporates reflections from the boundary. We show that the latter Green's function improves imaging performance after applying a normalization that compensates for the interference between direct and reflected fields. We also show that the best images are attained in some cases when the number of singular vectors in the signal subspace exceeds the number of spheres. This is consistent with previous analysis showing multiple eigenvalues of the time reversal operator for spherical scatterers [Chambers and Gautesen, *J. Acoust. Soc. Am.* **109** (2001)]. [Work performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.]

Session 3aAA

Architectural Acoustics, Noise, Physical Acoustics, Signal Processing in Acoustics and Underwater Acoustics: A Celebration of the Life in Acoustics of Richard H. Bolt

William J. Cavanaugh, Chair

Cavanaugh Tocci Associates, Inc., 327F Boston Post Road, Sudbury, Massachusetts 01776

Chair's Introduction—8:30

Invited Papers

8:35

3aAA1. Richard H. Bolt: Scientist, educator and friend. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138-5755)

The author will discuss his close relation to Dick Bolt, starting with their first meeting in 1938. This will include Bolt's doctoral years at the University of California, his National Research Fellowship in Physics at MIT, his contributions to the World War II effort, his tenured appointment, guidance of students and formation of the Acoustical Laboratory at MIT, and his co-founding of and subsequent participation in Bolt Beranek and Newman. Dick's principal contributions to the field of acoustics, to education and to the Acoustical Society of America will be presented up until 1957, after which his talents were largely taken over by national missions, which are mentioned. Richard H. Bolt was an inspiring educator, researcher and organizer. He truly was a national treasure.

9:00

3aAA2. Richard H. Bolt introduced me to acoustics, and I am still engaged. Ira Dyer (Dept. of Ocean Eng., MIT., Cambridge, MA 02139)

I arrived at the MIT Acoustics Laboratory in 1948, and quickly learned from Dick Bolt that U.S. submarines needed coatings to reduce vulnerability via active sonar detection. All who remember Dick will be familiar with his next step: He brought me to Robert Blizzard and A. Wilson Nolle, both then at the Lab, from whom I learned in detail how to measure and then how to understand the dynamic moduli of rubber-like materials. My path, of course, widened to include medical acoustics (under the leadership of Theodore Hueter) and scattering (Phillip Morse and Uno Ingard). Fast-forward to the present. Rubber-like materials are not only in use as submarine coatings to avoid detection by active sonars, but also to reduce noise radiated by submarines. Further, such materials cover submarine acoustic arrays to reduce flow-noise interference with signals received by such arrays. The problem that most engages me today is the stochastic nature of sound propagation in the ocean, a problem that has the practical consequence of degraded sonar performance. In the spirit of Dick, I will present snippets of my current work, with the hope of painting the big picture that he always did so well.

9:25

3aAA3. Richard H. Bolt's work on theoretical aspects of room acoustics. Richard H. Lyon (RH Lyon Corp, Cambridge, MA 02138, rhlyon@lyoncorp.com)

From about 1940 to 1950 Dick Bolt undertook a number of studies of the transmission of sound in a room. These were mostly theoretical and appear to have had two sources of inspiration: the work of Wente at Bell Laboratories in the 1930s, and Philip Morse's work on the eigenfunctions and eigenvalues of idealized rooms in the early 1940s. In 1944, the *Reviews of Modern Physics* published in its April issue a single paper—that of Dick Bolt and Phil Morse, titled “Sound waves in rooms.” This *tour de force* took the viewpoint that a deterministic calculation of the modes of an idealized room could reveal important features of the acoustics of less ideal spaces. The relations between wall impedance and modal damping for axial, tangential, and oblique modes and the construction of direct fields from modal expansions are worthy of special note. Perturbation analysis is used to couple idealized modes as a first attempt to deal with “real” rooms and the transition to ergodic behavior. This work then led Bolt to a number of studies of the irregularity of spacing of modal resonance frequencies, which had some influence, in turn, on models of modal distributions in Statistical Energy Analysis.

9:50–10:10 Break

10:10

3aAA4. Acoustics and its relation to language: The influence of Dick Bolt. Kenneth N. Stevens (Res. Lab. of Electronics and Dept. of Elec. Eng. and Computer Sci., MIT, Cambridge, MA 02139)

Under the mentoring of Dick Bolt, and the stimulation he provided in the Acoustics Laboratory at MIT, many students were exposed to a range of topics in acoustics, including mechanisms of sound generation, radiation, and propagation, noise control, acoustics of resonators and rooms, human responses to sound, and speech perception under various adverse conditions. As someone

who became interested in speech communication, I have recognized that this kind of quantitative background in acoustics is an important requirement for developing models of how humans produce speech, how they perceive and understand speech, and how children acquire these skills. Speech production involves sound sources produced by a nonlinear mechanical system and by noise arising from turbulent airflow. Sound is propagated in a vocal tract with yielding walls, and acoustic coupling is introduced by lossy resonators attached to the vocal tract, including the trachea and the nasal cavity. These acoustic principles of sound generation create an inventory of sound types that give rise to distinctive responses in the ears and brains of listeners. The solid grounding in acoustics provided by Dick Bolt and his leadership have helped in the formation of this linkage between acoustics, speech physiology, linguistics, and human perception.

10:35

3aAA5. Richard H. Bolt—Mentor and colleague. Ewart A. Wetherill (Shen, Milsom & Wilke/Paoletti, 649 Mission, San Francisco, CA 94105, rwetherill@sf.smwinc.com)

Of his many accomplishments in acoustics, perhaps one of Richard H. Bolt's greatest legacies will prove to be his contribution, both directly and through his teaching, to everyday hearing conditions in buildings. In a discipline that attempts to bridge the technical and cultural gap between a pure science and the pragmatic and often-haphazard process of building design, he combined a deep understanding of both professions with an ability to communicate complex ideas that is reminiscent of Wallace Clement Sabine. His welcoming enthusiasm and humility enabled him to attract and to work well with people of complementary talents, in both theoretical research and the gritty details of a consulting practice, as well as to envision the potential of still-unexplored subjects. A logical outcome of this combination was the profoundly influential technical group known as BBN, whose pioneering integration of acoustics with building technology is echoed by many consulting groups throughout the world.

11:00–11:40

Panel Discussion

WEDNESDAY MORNING, 30 APRIL 2003

ROOM 201, 7:55 A.M. TO 12:05 P.M.

Session 3aAO

Acoustical Oceanography, Animal Bioacoustics and Underwater Acoustics: Bioacoustic Resonance Spectroscopy

Orest I. Diachok, Chair

Naval Research Laboratory, Code 7420, 4555 Overlook Avenue Southwest, Washington, D.C. 20375-0002

Chair's Introduction—7:55

Invited Papers

8:00

3aAO1. Acoustic swimbladder resonance spectroscopy: Fundamentals in scattering theory. David T. I. Francis (Dept. of Electron., Elec. and Computer Eng., Univ. of Birmingham, Birmingham B15 2TT, UK) and Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

A history of the physics of acoustic resonance is given. The primary, low-frequency, resonant scattering model for air bubbles in water [Minnaert (1933)] is reviewed. Subsequent applications to swimbladder fish, including models by Andreeva (1964), Love (1978), and Feuille and Nero (1998), among others, are developed. Reference is made to exemplary measurements of backscattering by Holliday (1972) and Loevik and Hovem (1979), and of forward scattering, or absorption, by Weston (1967) and Diachok (2000), among others. High-frequency resonances are also described, with presentation of both analytical and numerical results for the immersed air bubble. Comparison of these validates the numerical, boundary-element method (BEM). The BEM allows high-frequency resonances to be studied for swimbladders of realistic shapes under pressure and for typical wave-number-swimbladder length products of order 10–40. Implications of high-frequency swimbladder resonance for auditory function in fish are mentioned. [Work supported by ONR.]

9:00

3aAO2. High-frequency acoustic scattering from gas-bearing zooplankton. Andone C. Lavery, Timothy K. Stanton, Peter H. Wiebe (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Mark C. Benfield (Louisiana State Univ., Baton Rouge, LA 70803), and Charles H. Greene (Cornell Univ., Ithaca, NY 14853)

High-frequency acoustic scattering is well-suited to the synoptic investigation of marine organisms that inhabit the water-column, such as zooplankton and fish. However, the scattering characteristics of the organisms can be highly complex, and one must look for ways to distinguish one type of organism from another when interpreting the received echoes. From an acoustic scattering perspective,

zooplankton generally fall into one of three categories: gas-bearing, fluid-like, or elastic-shelled. Scattering models, of different levels of sophistication are available for all three types of zooplankton. Gas-bearing zooplankton, unlike the other zooplankton categories, have a strong scattering resonance, which can be exploited for their identification. Scattering from gas-bearing zooplankton at frequencies close to the resonance frequency is much stronger than scattering from other zooplankton. Thus, in this frequency region, acoustic scatter from a small number of gas-bearing zooplankton can overwhelm that from a far larger number of fluid-like or elastic-shelled zooplankton. A series of zooplankton surveys of the Gulf of Maine has recently been completed in which a multi-frequency towed instrument platform, BIOMAPER-II, was employed. The strong scattering resonance of gas-bearing zooplankton, together with scattering models and ground-truthing, was exploited in order to determine regions with high densities of gas-bearing zooplankton.

9:20

3aAO3. Potential effect of resonant scattering from multiple swimbladders on audition in juvenile fish. Mardi C. Hastings (ONR 342, BCT-1, 800 N. Quincy St., Arlington, VA 22217, hastinm@onr.navy.mil)

The swimbladder, a gas-filled chamber in the abdominal cavity of most bony fishes, is a hydrostatic organ that enables fish to maintain neutral buoyancy; however, it also responds to acoustic pressure and radiates a secondary acoustic field that enhances detection capability of the inner ear. Recent experiments have indicated that resonant response of the swimbladder may control the auditory bandwidth in at least four species of fish [Hastings *et al.*, J. Acoust. Soc. Am. **110**, 2640 (2001)]. The auditory bandwidths of these fishes, however, do not change appreciably while they grow even though the resonance frequency of the swimbladder decreases with increasing body length. Results of an analysis inspired by Feiullade *et al.* [J. Acoust. Soc. Am. **112**, 2206 (2002)] show that the downward shift and broadening associated with resonance of the aggregate scattered field from multiple fish is perhaps sufficient enough to account for this discrepancy. Effects of resonant characteristics of a single swimbladder, fish length, and number of fish on the changes in the collective scattered field are presented. Thus the resonant scattered field created by relatively large schools of juvenile fish may enhance their auditory capability.

9:40

3aAO4. Experimental measurements of lung resonant frequencies in a bottlenose dolphin (*Tursiops truncatus*) and white whale (*Delphinapterus leucas*). James J. Finneran (U.S. Navy Marine Mammal Prog., Space and Naval Warfare Systems Ctr., San Diego, Code 2351, 53560 Hull St., San Diego, CA 92152)

An acoustic backscatter technique was used to estimate *in vivo* whole-lung resonant frequencies in a bottlenose dolphin (*Tursiops truncatus*) and a white whale (*Delphinapterus leucas*). Subjects were trained to submerge and position themselves near an underwater sound projector and a receiving hydrophone. Acoustic pressure measurements were made near the subjects' lungs while insonified with pure tones at frequencies from 16 to 100 Hz. Whole-lung resonant frequencies were estimated by comparing pressures measured near the subjects' lungs to those measured from the same location without the subject present. Experimentally measured resonant frequencies and damping ratios were much higher than those predicted using equivalent volume spherical air bubble models. The experimental technique, data analysis method, and discrepancy between the observed and predicted values will be discussed. The potential effects of depth on the resonance frequencies will also be discussed.

10:00–10:10 Break

10:10

3aAO5. The influence of adjacent structure on the response of insonified fish and marine mammal resonant cavities. Joel Garrelick (Applied Physical Sciences, Inc., 2 State St., Ste. 300, New London, CT 06320)

It has long been known that the acoustic response of, and scattering from, finite elastic structures and cavities may be enhanced at their submerged natural frequencies, that is, at resonance. Fish and marine mammal cavities, viz., air bladders and lungs, respectively, are prime examples. In either case, classic analyses represent the cavity as being in an otherwise free field and spherical in geometry, initially as well as in its deformed state. Such simplifying assumptions may be invalidated by the presence of adjacent structure, whether or not such structure is transparent to the incident wave. This is especially the case for local response functions, namely stresses and strains in surrounding tissue, as opposed to say the acoustically scattered return. A strong sensitivity of these response functions to the assumed geometry compounds the problem. These issues are analytically examined in this paper for the example of a plane acoustic wave incident on a membrane-sheathed cavity located adjacent to structure of prescribed high or low impedance. Predictions for the scattered acoustic field and cavity stress/strains are compared to classic free field values. Enhanced levels are computed and interpreted in terms of mode coupling to higher order membrane/cavity vibration modes.

10:30

3aAO6. Measurements and bioeffects of resonant gas bodies *in vivo*. Diane Dalecki (Dept. of Biomed. Eng., Univ. of Rochester, Rochester, NY 14627, dalecki@bme.rochester.edu)

The response of resonant gas bodies *in vivo* (such as the gas filled lung and intestine of mammals) to exposure to low frequency underwater sound was characterized through a series of investigations. Measurements were performed in the laboratory using a specially designed acoustic exposure system, capable of generating maximum acoustic fields of ~ 200 dB *re* 1 μ Pa for frequencies

spanning 100–2500 Hz. Both acoustic scattering and pulse-echo ultrasound techniques were used to characterize the response of gas bodies to underwater sound exposure and to determine the resonance frequencies of gas bodies in mammalian tissues *in vivo*. A series of investigations has demonstrated that the effects of low frequency (~ 100 –2500 Hz) underwater sound can be significant in and near tissues that contain resonant gas bodies. For example, mice exposed to underwater sound at the resonance frequency of their lung exhibited lung damage and the extent of tissue damage increased with increasing pressure amplitude. Similar types of investigations were performed with mammalian lungs of various sizes and with intestinal gas *in vivo*.

Contributed Papers

10:50

3aAO7. Observations of frequency shift associated with schooling fish.

Orest Diachok (Naval Res. Lab., Washington, DC 20375, orest@wave.nrl.navy.mil)

The number of sardines per school, N , is nominally 10 000 and the separation between sardines in school, s , is nominally 1 fish length, L . s is much smaller than the wavelength at f (the resonance frequency of individuals), which suggests that schools may exhibit acoustic properties of bubble clouds. Long-term, broadband transmission loss measurements at a shallow-water site in the Gulf of Lion revealed absorption lines due to dispersed sardines at 1.3 kHz at 20 m at night and 2.7 kHz at 65 m at dawn. Temporal changes in observed values of f were consistent with concurrent echo sounder observations of the vertical migration of sardines, and theoretical computations based on laboratory measurements of swim bladder dimensions. The measured resonance frequency of sardines in schools during daytime, 1.7 kHz at 65 m, was $0.6f$ at the same depth at dawn. The observed frequency shift is consistent with a hybrid model of the fundamental resonance frequency of a bubble cloud, which is based on theories developed by Feuilleade, Nero, and Love (1996) and d'Agostino and Brennan (1988), and $s=0.8L$ and $N=5000$ fish per school. [Work supported by ONR.]

11:05

3aAO8. Backscattering and extinction cross sections of two swimbladdered fishes at the lowest resonance, as modeled by the boundary-element method.

Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543) and David T. I. Francis (Univ. of Birmingham, Birmingham B15 2TT, UK)

The boundary-element method has been applied to backscattering and extinction of sound by swimbladdered fish at the lowest, breathing-mode resonance. Corresponding cross sections have been computed for specimens of two representative kinds of swimbladder-bearing fish, namely physostomes and physoclists, which, respectively, possess and lack an external duct. The respective fishes are herring (*Clupea harengus*) and pollack (*Pollachius pollachius*), for which swimbladder morphometric data are available. The depth dependences of the cross sections are computed over the range 0–500 m. Comparisons are made with measurements and other modeled results for a number of species. [Work supported by ONR.]

11:20

3aAO9. Target strength and density structure of Hawaiian mesopelagic boundary community patches.

Kelly Benoit-Bird and Whitlow Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, benoit@hawaii.edu)

A broadband sonar system and digital camera with strobe lights were mounted on a vertically profiling frame with a depth sensor. The target strengths and densities of animals within individual mesopelagic boundary community patches were investigated as a function of depth. Simultaneous echosounder surveys permitted comparison of density estimates from echo-energy integration and echo-highlight counting. Target strength values suggest nearshore boundary community animals are primarily myctophid fishes which was confirmed by preliminary photographic evidence. Target strength varied significantly as a function of distance from the shoreline and time. Echo-energy integration estimates of density made with these revised target strengths compare well with those made with echo highlight counting. These density measures show that previous density estimates were too low but do not change the conclusions of these

studies. Vertical microstructure in density was apparent but animal size and compositional structure was not evident within a patch. Patch edges were abrupt, with no differences in the density or target strength from patch interiors. These edges were generally straight, with a sharp drop in density to the background density of zero. Estimates of animal size as a function of time provide information about the diel migration patterns of these mesopelagic animals.

11:35

3aAO10. Inversion of the depth, thickness, and absorption coefficient of a layer of fish (anchovies) from transmission loss measurements in the Yellow Sea.

Orest Diachok and Stephen Wales (Naval Res. Lab., Washington, DC 20375, orest@wave.nrl.navy.mil)

Estimates of bioacoustic parameters of fish (anchovies) and geoacoustic parameters of the bottom were simultaneously inverted from the TL measurements of Qiu *et al.* (1999) in the Yellow Sea. Replica fields were calculated with BIO-C-SNAP, a C-SNAP based normal mode model, which permits inclusion of bio-absorption layers. The inversion was based on minimizing the rms difference, D , between measured and calculated values of TL at multiple source and receiver depths and ranges, and involved a simultaneous search for bio-layer depth (d), bio-layer thickness, bio-alpha (a), geo-sound speed, and geo-alpha. The resultant values of D were extremely small (1.9 dB at 1.35 kHz). By contrast, inversion calculations, which assumed that all excess attenuation at this site was due to the bottom, resulted in unacceptably large values of D (9.5 dB at 1.35 kHz). The inverted value of d , 6.8 m, was consistent with laboratory measurements of the resonance frequencies of 10-cm anchovies (1.35 kHz at 6 m), the dominant species in the Yellow Sea. The inverted value of a was consistent with previously reported number densities of anchovies in the Yellow Sea. Inverted geoacoustic parameters were within previously reported bounds. [Work supported by ONR.]

11:50

3aAO11. Studies on geometrical backscattering models of marine bodies.

Anil Kumar C. Parameswaran, Sajith N. Pai, N. Soniraj, P. R. Saseendran Pillai, James Kurian, Supriya M. Hariharan, C. Madhavan, and T. K. Mani (Cochin Univ. of Sci. and Technol., Cochin 682022, India)

The target strength of marine bodies depends on two components—one of them is easy to measure and a good relationship can be established with the target strength value, while the other presents a higher variability that no method can help to reduce, which include the orientation of the fish and a range of environmental conditions and a set of biological facts. Studies on physical models and aspects are expected to provide a clear insight into the issues relating to the target strength variability. Such physical models are developed by converting the physiological shape of the fish into standard and simple geometrical shapes. Data obtained from some of the commercially important species, individually positioned at the center of the acoustic beam, 3 m from the transducer in a test facility were used for the computation of target strength. The target strength value obtained from these reference targets is an indicator of the model performance. Mathematical description of the scattering by some of the species and subsequent comparison with laboratory data have demonstrated that the scattered level by an individual due to a single ping, strongly depends upon size, shape, frequency, material properties, and orientation. Perhaps one of the most notable peculiarities of this work is the simplicity of the approximation and the close agreement between the real world value and the model solution.

Session 3aBB**Biomedical Ultrasound/Bioresponse to Vibration: Audible-Frequency Medical Diagnostic Methods**

Thomas J. Royston, Chair

*Department of Mechanical Engineering, University of Illinois at Chicago, Mail Code 251, 842 West Taylor Street, Chicago, Illinois 60607-7022***Chair's Introduction—8:00*****Invited Papers*****8:05****3aBB1. Audible-frequency methods in assessment of soft tissues.** Armen P. Sarvazyan (Artann Labs., 1753 Linvale-Harbourton Rd., Lambertville, NJ 08530)

One of the reasons why MHz-range ultrasonic compressional waves have become a powerful tool for medical diagnostics is that their wavelength is of the order of a millimeter, adequate for imaging an anatomical structural. Other types of acoustic waves that can be generated in soft tissue, such as shear and surface waves, are two–three orders of magnitude slower and, respectively, have a millimeter range wavelength in the audible-frequency range. These waves can be effectively used for imaging and soft tissue characterization in terms of the shear elasticity modulus. This paper is a summary of the studies on applications of the surface and shear waves for diagnostic assessment of soft tissue elasticity and dynamic viscosity performed by the author and his coworkers over three decades. A number of techniques for measuring shear elastic moduli of soft tissue using both wave propagation and the acoustic impedance methods are described. Audible-frequency methods and devices for assessment of skin in dermatological applications, for the detection of breast cancer, for imaging of brain and muscle tissues are analyzed. Various embodiments of the Shear Wave Imaging modality are considered. The reasons for a limited success of the audible-frequency methods despite their great potential in ultrasonic medical imaging and diagnostics are discussed.

8:30**3aBB2. Vibro-acoustography: The sound of tissue.** James Greenleaf and Mostafa Fatemi (Mayo Clinic, 200 First St. SW, Rochester, MN 55905)

Palpation is routinely used for detecting pathology using the stiffness of the tissue and is more than 2000 years old. Palpation is subjective and limited to individual experience and to the accessibility of the tissue region to touch. Noninvasive images of elastic modulus may be useful to distinguish tissues and pathologic processes based on mechanical properties such as elastic modulus. Static, quasistatic, or cyclic stresses have been applied while strains have been measured using ultrasound or MRI. The related elastic modulus has been estimated from visco-elastic models of tissue mechanics. Recently we have developed a new ultrasound technique that produces speckle free images related to both tissue stiffness and reflectivity. This method, termed Ultrasound Stimulated Vibro-acoustography [Science **280**, 83–85 (1998); PNAS **96**, 6603–6608 (1999)], uses ultrasound radiation pressure to produce sound vibrations from a small region of the tissue that depend on the scattering and elastic characteristics of the tissue. The method can detect microcalcification within breasts, and promises to provide high-quality images of calcification within the arteries. Vibro-acoustography can detect mechanical defects in prostheses such as artificial mitral and aortic valves. The method may also be used in nondestructive evaluation. [Work supported by HL 61451, DBI 7980-4.]

8:55**3aBB3. Supersonic imaging of elasticity.** Mathias Fink, Michael Tanter, and Jeremy Bercoff (Laboratoire Ondes et Acoustique, 10 rue Vauquelin, 75005 Paris, France)

This paper shows the new possibilities of supersonic imaging of elasticity. It combines ultra-high speed ultrasonic imaging of tissues with low frequency supersonic shear sources induced in the body by radiation pressure. We have built an Ultrafast Scanner providing up to 5000 frames/s and using conventional arrays linked to a new beamformer. We will show how this scanner is able to follow in real time tissue displacements induced by the transient propagation of LF shear wave in the human body. From the spatio-temporal evolution of the displacements, a shear elasticity map is constructed using a local inversion algorithm. In order to obtain unbiased shear elasticity map, different configurations of shear sources induced by radiation pressure of focused transducer

arrays are used. A very interesting configuration that induces quasi plane shear waves will be described. It used a shear source that moves at supersonic velocities, and that is created by using a very peculiar beam forming in the transmit mode. *In vitro* and *in vivo* results will be presented that demonstrate the interest of this new transient elastographic technique.

9:20

3aBB4. Vibration amplitude sonoelastography lesion imaging using low-frequency audible vibration. Lawrence Taylor (Biomed. Eng. Dept., Univ. of Rochester, P.O. Box 270168, Rochester, NY 14627) and Kevin Parker (Univ. of Rochester, Rochester, NY 14627)

Sonoelastography or vibration amplitude imaging is an ultrasound imaging technique in which low-amplitude, low-frequency shear waves, less than 0.1-mm displacement and 1-kHz frequency, are propagated deep into tissue, while real time Doppler techniques are used to image the resulting vibration pattern. Finite-element studies and experiments on tissue-mimicking phantoms verify that a discrete hard inhomogeneity present within a larger region of soft tissue will cause a decrease in the vibration field at its location. This forms the basis for tumor detection using sonoelastography. Real time relative imaging of the vibration field is possible because a vibrating particle will phase modulate an ultrasound signal. The particle's amplitude is directly proportional to the spectral spread of the reflected Doppler echo. Real time estimation of the variance of the Doppler power spectrum at each pixel allows the vibration field to be imaged. Results are shown for phantom lesions, thermal lesions, and 3-D *in vitro* and 2-D *in vivo* prostate cancer. MRI and whole mount histology is used to validate the system accuracy.

9:45

3aBB5. Acoustic detection of pneumothorax. Hansen A. Mansy (Biomed. Acoust. Res. Group, Rush Medical College, 1725 W. Harrison St., Ste. 946, Chicago, IL 60612, hmansy@rush.edu), Thomas J. Royston (Univ. of Illinois at Chicago, Chicago, IL), Robert A. Balk, and Richard H. Sandler (Biomed. Acoust. Res. Group, Rush Medical College, Chicago, IL 60612)

This study aims at investigating the feasibility of using low-frequency (<2000 Hz) acoustic methods for medical diagnosis. Several candidate methods of pneumothorax detection were tested in dogs. In the first approach, broadband acoustic signals were introduced into the trachea during end-expiration and transmitted waves were measured at the chest surface. Pneumothorax was found to consistently decrease pulmonary acoustic transmission in the 200–1200-Hz frequency band, while less change was observed at lower frequencies ($p < 0.0001$). The ratio of acoustic energy between low (<220 Hz) and mid (550–770 Hz) frequency bands was significantly different in the control (healthy) and pneumothorax states ($p < 0.0001$). The second approach measured breath sounds in the absence of an external acoustic input. Pneumothorax was found to be associated with a preferential reduction of sound amplitude in the 200- to 700-Hz range, and a decrease of sound amplitude variation (in the 300 to 600-Hz band) during the respiration cycle ($p < 0.01$ for each). Finally, chest percussion was implemented. Pneumothorax changed the frequency and decay rate of percussive sounds. These results imply that certain medical conditions may be reliably detected using appropriate acoustic measurements and analysis. [Work supported by NIH/NHLBI #R44HL61108.]

10:10–10:20 Break

10:20

3aBB6. Passive acoustic monitoring of human physiology during activity indicates health and performance of soldiers and firefighters. Michael V. Scanlon (US Army Res. Lab., AMSRL-SE-SA, 2800 Powder Mill Rd., Adelphi, MD 20783-1197, mscanlon@arl.army.mil)

The Army Research Laboratory has developed a unique gel-coupled acoustic physiological monitoring sensor that has acoustic impedance properties similar to the skin. This facilitates the transmission of body sounds into the sensor pad, yet significantly repels ambient airborne noises due to an impedance mismatch. The sensor's sensitivity and bandwidth produce excellent signatures for detection and spectral analysis of diverse physiological events. Acoustic signal processing detects heartbeats, breaths, wheezes, coughs, blood pressure, activity, motion, and voice for communication and automatic speech recognition. 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. Comfortable acoustic sensors can be in a helmet or in a strap around the neck, chest, and wrist. Noise-canceling sensor arrays help remove out-of-phase motion noise and enhance covariant physiology by using two acoustic sensors on the front sides of the neck and two additional acoustic sensors on each wrist. Pulse wave transit time between neck and wrist acoustic sensors will indicate systolic blood pressure. Larger torso-sized arrays can be used to acoustically inspect the lungs and heart, or built into beds for sleep monitoring. Acoustics is an excellent input for sensor fusion.

3aBB7. Bone and tissue conduction of high intensity acoustic energy to the human cochlea. Richard McKinley (Air Force Res. Lab., AFRL/HECB, 2610 Seventh St., Wright-Patterson Air Force Base, OH 45433-7901, richard.mckinley@wpafb.af.mil), Armand Dancer (French German Res. Inst., St. Louis, France), and Henning von Gierke (Air Force Res. Lab. (Emeritus), AFRL/HECB, OH 45433-7901)

Noise fields near operating high performance fighter aircraft range from 140 to 150 dB overall sound pressure level. Much of the acoustic energy which is transmitted to the cochlea in a well protected ear (earplugs and earmuffs) arrives by tissue and bone conduction pathways. A better understanding of this transmission of acoustic energy by these biological materials could lead to improved protection and/or methods of actively attenuating this acoustic energy in the cochlea. Associated issues to be discussed are the possible damaging levels of bone/tissue conducted acoustic energy and the actual bone/tissue attenuation. This paper will describe a series of experiments in a recent investigation of bone and tissue conduction of acoustic energy over a broad band of frequencies and a look at the initial results of the first of the studies. [Work supported in part by the Air Force Office of Scientific Research.]

Contributed Papers

11:10

3aBB8. A vibration model for frequency analysis of arterial tubes with tissue. Xiaoming Zhang, Mostafa Fatemi, and James F. Greenleaf (Ultrasound Res. Lab., Mayo Clinic, 200 First St. SW, Rochester, MN 55905, zhang.xiaoming@mayo.edu)

Vibro-acoustography is a new noncontact imaging method based on the radiation force of ultrasound. We extend this technique for tissue characterization of arterial tubes by vibration techniques. The arterial tube can be excited remotely by ultrasound at its resonant frequencies where the vibration and acoustic emission of the tube can be measurable. From these resonant frequencies, the material properties of the arterial tube can be found. A theory for a tube with tissue is formulated using first-order shear deformation theory to include the effects of transverse shear deformation and rotary inertia. A wave propagation approach is applied for easy handling of the boundary conditions. Experimental studies were carried out on a silicone tube embedded in a cylindrical gel phantom. A confocal transducer is used to produce the radiation force of ultrasound for exciting the tube-phantom structure. The vibration of the tube and the phantom are measured with a laser vibrometry system. The fundamental mode of a tube-phantom structure is well excited by the radiation force of ultrasound, and was measured to be 81.8 Hz, which is close to the theoretical prediction of 83.3 Hz. Both excitation and measurement are remote and noncontact, important attributes for future study of arteries.

11:25

3aBB9. A new imaging technique based on resonance for arterial vessels. Xiaoming Zhang, Mostafa Fatemi, and James F. Greenleaf (Ultrasound Res. Lab., Mayo Clinic, Rochester, MN 55905, zhang.xiaoming@mayo.edu)

Vibro-acoustography is a new noncontact imaging method based on the radiation force of ultrasound. We extend this technique for imaging of arterial vessels based on vibration resonance. The arterial vessel is excited remotely by ultrasound at a resonant frequency, at which the vibration of the vessel as well as its transmission to the body surface are large enough to be measured. By scanning the ultrasound beam across the vessel plane and measuring the vibration at one single point on the body or vessel surface, an image of the interior artery can be mapped. Theory is developed that predicts the measured velocity is proportional to the value of the mode shape at resonance. Experimental studies were carried out on a silicone tube embedded in a cylindrical gel phantom of large radius, which simulates a large artery and the surrounding body. The fundamental frequency was measured at which the ultrasound transducer scanned across the tube plane with velocity measurement at one single point on the tube or on the phantom by laser. The images obtained show clearly the interior tube and the modal shape of the tube. The present technique offers a new imaging method for arterial vessels.

11:40

3aBB10. Computerized analysis of bowel sounds in normal and small bowel obstructed rats. Richard Sandler, Hansen Mansy, Michael Uhing, Peter Meyer, and Robert Kimura (Biomed. Acoust. Res. Group, Rush Medical College, 1725 W. Harrison St., Ste. 946, Chicago, IL 60612, rsandler@rush.edu)

Small bowel obstruction (SBO) is a common surgical emergency which may be mimicked by ileus or other nonsurgical conditions. The aims of this work is to delineate gastrointestinal sound (GIS) correlates in a rat model. Seven rats were studied in paired SBO and control states. Computerized analysis of GIS was performed under continuous IV sedation. After adaptive filtering, every GIS event was isolated and analyzed for duration and dominant frequency. It was found that long duration sounds (greater than 100 ms) occurred in each of the obstructed, but in none of the nonobstructed cases ($p=0.02$). The overall mean event duration and dominant frequency in SBO compared to control states was both significantly longer and lower ($22.42.6$ vs 7.0 ± 2.6 ms, $p=0.001$ for duration; and 296 ± 34 vs 427 ± 33 , $p=0.001$ for frequency). Besides these mean differences, there was also a clear evolution with time in GIS characteristics, with lengthening of the duration ($+0.56$ ms/min, $p=0.001$) and lowering of the dominant frequency (-3.3 Hz/min, $p=0.01$). It is concluded that GIS analysis may prove useful in the noninvasive, rapid, and accurate diagnosis of SBO.

11:55

3aBB11. Response of a viscoelastic halfspace to subsurface distributed acoustic sources with application to medical diagnosis. Thomas J. Royston, Yigit Yazicioglu, and Francis Loth (Univ. of Illinois at Chicago, 842 W. Taylor St., MC 251, Chicago, IL 60607, troyston@uic.edu)

The response within and at the surface of an isotropic viscoelastic medium to subsurface distributed low audible frequency acoustic sources is considered. Spherically and cylindrically distributed sources are approximated as arrays of infinitesimal point sources. Analytical approximations for the acoustic field radiating from these sources are then obtained as a summation of tractable point source expressions. These theoretical approximations are compared to computational finite element predictions and experimental studies in selected cases. The objective is to better understand low audible frequency sound propagation in soft biological tissue caused by subsurface sources. Distributed acoustic sources could represent vibratory motion of the vascular wall caused by turbulent blood flow past a constriction (stenosis). Additionally focused vibratory stimulation using a dynamic radiation force caused by interfering ultrasound beams effectively creates a distributed subsurface acoustic source. A dynamic radiation force has been investigated as a means of probing subsurface tissue anomalies, including calcified vascular plaque and tumorous growths. In these cases, there is an interest in relating acoustic measurements at the skin surface and within the medium to the underlying flow/constriction environment or tissue anomaly. [Research supported by NIH NCRR 14250 and Whitaker Foundation BME RG 01-0198.]

Session 3aED**Education in Acoustics, Physical Acoustics and Noise: Demos for all Ages—2003**

James M. Sabatier, Cochair

National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Murray S. Korman, Cochair

Department of Physics, United States Naval Academy, Annapolis, Maryland 21402

Matthew E. Poese, Master of Ceremonies

Graduate Program in Acoustics, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16801

The Committee on Education in Acoustics is proud to present a showcase of demonstrations that we think will inspire and captivate a diversified audience as well as motivate further inquiries in the science of sound and associated areas. Please note that presentation times vary. The demonstrations are listed by presenter in alphabetical order but not necessarily in the order of presentation:

Invited Presentations

Modes in the pressure plate of a landmine and their changes with increasing load. W. C. Kirkpatrick Alberts II (Univ. of Mississippi, University, MS 38677)

A laser Doppler vibrometer will be used to measure the vibration of the pressure plate and its output will be shown on a signal analyzer. A load will be added to the pressure plate of the landmine which will demonstrate how the modes change with increasing load.

Spinning cups with sound. Anthony A. Atchley (Graduate Program in Acoustics, Pennsylvania State Univ., 217 Applied Science Bldg., University Park, PA 16802)

After creating a high amplitude two-dimensional standing wave field, first one then four upside-down styrofoam coffee cups can be made to spin individually about a balance point. By changing the relative phases of the field, the cups can be made to change their rotational direction.

Deflection of light by light, Cartesian diver and Super rebound. Lawrence A. Crum (Center for Industrial and Medical Ultrasound, Applied Physics Lab., Univ. of Washington, 1013 NE 40th Street, Seattle, WA 98105)

Three unique demos: 1. "Deflection of Light by Light" (actually a thermal gradient), and its relation to bending of sound waves in the ocean, 2. "Cartesian Diver," and its relation to some traveling gliders that APL is developing using the same principle, and 3. "Super Rebound," an interesting phenomenon that occurs when two elastic objects are combined.

A capacitive wave height detector. Jessica Drewery (National Center for Physical Acoustics, 1 Coliseum Dr., University, MS 38677)

The capacitive wave height detector converts low frequency fluid height into a voltage that can be seen on the oscilloscope. From there, you can see the relationship between frequency and wavelength as well as dispersion effects.

Acoustic levitation of droplets and styrofoam pellets in air. E. Carr Everbach (Engineering Dept., Swarthmore College, 500 College Ave., Swarthmore, PA 19081-1397)

Based on a demonstration popularized by Bob Apfel, a 22.4 kHz ultrasonic emulsifier, often used for sonification in biology labs, is used to create an intense standing wave field in air. Small styrofoam pellets from a bean-bag chair and liquid droplets can be levitated and manipulated by audience members with a piece of window screen or hookup wire.

Chladni patterns: Observing normal modes in two dimensional structures. Uwe J. Hansen (Dept. of Physics, Indiana State Univ., Terre Haute, IN 47803)

Sand accumulates along the nodal lines on a plate vibrating at one of the frequencies associated with a normal mode of vibration, thus providing a visual representation of the normal mode pattern. The phase difference of vibrations in adjacent sections separated by a nodal line is observed in the Lissajous figure on an oscilloscope screen when a monitoring microphone crosses a nodal line.

Phase locking. Robert M. Keolian (Applied Research Lab., Pennsylvania State Univ., University Park, PA 16802)

An organ pipe emits a tone whose pitch is normally determined by the pipe's length and the speed of sound. But the organ pipe can change its pitch to match that of another sound source of nearly the same pitch if the two sources are placed close enough together, locking the phases of the pressure swings of the two sounds together and creating one unified sound.

Acoustics at the breakfast table. Andrew Morrison (Physics Dept., Northern Illinois Univ., DeKalb, IL 60115)

The breakfast table is an excellent place to observe some interesting acoustical effects. An empty coffee cup, like an ancient Chinese two-tone bell, emits two distinctly different tones depending upon where it is tapped. When it is filled with hot water and some instant coffee is added, however, a whole new set of sounds are heard when the cup is tapped.

Effects of rooms on voice and music, room mode simulator. Ralph T. Muehleisen (Civil, Environmental and Architectural Engineering, Univ. of Colorado, 428 UCB, Boulder, CO 80309-0428)

Animations and auralizations developed using MATLAB, CATT Acoustics, and Cooledit are used to enhance the understanding of the effects of (1) room architecture on voice and music, and (2) let students simulate measurements of standing waves in a room.

Acoustic log starter. Matthew E. Poese (Graduate Program in Acoustics, Pennsylvania State Univ., P.O. Box 30, State College, PA 16804)

Like a fire starter for a fireplace, a steel tube with a line of small holes drilled along the top is filled with propane and the gas jets are set afire. When the tube is insonified by a loudspeaker placed at one end, the flames above the holes are seen to vary in height along the tube due to second order acoustic effects in the gas contained in the tube.

Demonstration of nonlinear acoustic detection of buried landmines. Waini Karen Tai and Murray S. Korman (Dept. of Physics, U. S. Naval Academy, Annapolis, MD 21402) and James M. Sabatier (Univ. of Mississippi, University, MS 38677)

Airborne sound at two primary frequencies f_1 and f_2 are chosen at plus and minus 5 Hz from the natural resonance frequency of an inert model landmine that is buried a few inches deep in a sifted soil. A geophone located on the soil surface (over the mine) detects a rich spectrum of nonlinearly generated tones and only very weak nonlinear signals “off the mine.”

Change of sound speed with temperature. Edward J. Tucholski (Dept. of Physics, U. S. Naval Academy, Annapolis, MD 21402)

The effect of changing the temperature of the medium on the speed of a sound pulse is investigated and displayed.

Rijke tube and singing flame. Ray Scott Wakeland (Graduate Program in Acoustics, Pennsylvania State Univ., P.O. Box 30, State College, PA 16804)

Two unique demonstrations: 1. “Rijke Tube,” a tube having a short stack of screens across its diameter, located between a pressure anti-node and node, is made to resonate when the screens are heated. These oscillations require a steady air current through the tube (provided in this case by natural convection) and continue until the screens cool, 2. “Singing Flame,” an empty tube can be made to resonate at its fundamental frequency when suspended over a flame from a Bunsen burner.

WEDNESDAY MORNING, 30 APRIL 2003

ROOMS 110/111, 8:30 TO 11:40 A.M.

Session 3aMU

Musical Acoustics and Speech Communication: Singing Voice Acoustics

Thomas Cleveland, Chair

Department of Otolaryngology, Vanderbilt University, 1161 21st Avenue Northeast, Nashville, Tennessee 37232

Chair's Introduction—8:30

Invited Papers

8:35

3aMU1. Can listeners hear who is singing? A comparison of three-note and six-note discrimination tasks. Molly Erickson (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, Knoxville, TN 37996) and Susan Perry (Univ. of Tennessee, Knoxville, TN 37996)

Timbre is typically investigated as a perceptual attribute that differentiates a sound source at one pitch and loudness; however, singers are believed to have one timbre and singers with similar timbres are thought to comprise a voice category. Yet, by the technical definition, each singer would have a set of timbres across pitch, vowel, and loudness combinations, so that each singer or singing voice category might have a timbre template which allows listeners to recognize one singer or category of singer. This paper investigated the ability of listeners to identify which pitch in an ascending or descending sequence of three or six stimuli was sung by a different singer within and across voice categories. For three-note sequences, the task was difficult and listeners chose the most dissimilarly pitched stimulus as coming from the oddball singer. For six-note sequences, the detection of the oddball singer was much improved, with cross-category comparisons being the easiest. These results support the idea that timbre should be understood as a transformation that connects the different sounds of one source, that a “rich” set of exemplars is necessary to discover the trajectory, and that singers of the same category have similar timbre transformations.

3aMU2. Analysis of levels of support and resonance demonstrated by an elite singing teacher. Ronald C. Scherer, Nandhakumar Radhakrishnan (Dept. of Commun. Disord., Bowling Green State Univ., 200 Health Ctr., Bowling Green, OH 43403), and Andreas Poulimenos (Indiana Univ., Bloomington, IN)

This was a study of levels of singing expertise demonstrated by an elite operatic singer and teacher. This approach may prove advantageous because the teacher demonstrates what he thinks is important, not what the nonsinging scientist thinks should be important. Two pedagogical sequences were studied: (1) the location of support—glottis (poor), chest (better), abdomen (best); (2) locations of resonance—hard palate/straight tone (poor), mouth (better), sinus/head (best). Measures were obtained for a single frequency (196 Hz), the vowel /ae/, and for mezzo-forte loudness using the /pae pae pae/ technique. Sequence differences: The support sequence was characterized by formant frequency lowering suggestive of vocal tract lengthening. The resonance sequence was characterized by flow (AC, mean flow) and abduction increases. Sequence similarities: The best locations had the widest $F2$ bandwidths. The better and best locations had the largest dB difference between $F2$ and $F3$. Although acoustic power increased through the sequences, the acoustic efficiency was not a discriminating factor. Open and speed quotients were not differentiating. The flow resistance was highest and aerodynamic power the lowest for the first of each sequence. Combined data: The maximum flow declination rate correlated highly with the AC flow ($r = -0.92$) and SPL ($r = 0.901$).

9:15

3aMU3. Sex and the singer: Gender categorization aspects of singing voice. Sten Ternström (Dept. of Speech, Music and Hearing, Kungliga Tekniska Högskolan, S-100 44 Stockholm, Sweden, sten@speech.kth.se)

The singing voice exhibits many systematic differences by gender and age. The physiological differences between the voice organs of males, females, and children are well known and give rise to several acoustic differences, including acoustic power, pitch range, and spectral distribution. Vocal artists often strive to widen their range of expression, and it is not uncommon for males to sing in a femalelike register, as in counter tenors and in some pop/rock genres. The opposite, however, is quite rare. While ambiguous or contradictory gender in speech is usually a social disadvantage, in singing it can be a desired effect. The physical differences in singing voice production between males and females are reviewed in detail. Some interesting borderline cases are examined from an acoustic standpoint.

9:35

3aMU4. Simulation of singing qualities governed by lower vocal tract adjustments. Ingo R. Titze (Dept. of Speech Pathol. & Audiol., Univ. of Iowa & Natl. Ctr. for Voice & Speech, Denver Ctr. of Performing Arts, 1245 Champa St., Denver, CO 80204, ingo-titze@dcpa.org)

In previous meetings, voice qualities such as pressed, ring, yawn, and twang were discussed in a speech context. It was shown that these qualities have unique spectral characteristics brought about by combinations of glottal and lower vocal tract adjustments (the epilarynx tube and the pharynx). Yawn has a wide glottis, a wide epilarynx tube, and a wide pharynx. On the contrary, twang has a general narrowing of all these airway sections. Ring has a wide pharynx and a relatively narrow epilarynx tube. A pressed voice is primary laryngeal, with a narrowed glottis. In this presentation, similar adjustments are made for singing with a voice simulator that controls vocal tract area functions and glottal flow pulses by rules. Results suggest that various singing styles, such as country-western, opera, or pop, may in part be characterized by these unique combinations of source and filter adjustments.

9:55–10:10 Break

10:10

3aMU5. The singing voice and country music. Wendy D. LeBorgne (The Blaine Block Inst. for Voice Anal. and Rehabilitation, 369 W. First St., Ste. 408, Dayton, OH 45402, wlvlar@aol.com)

Preliminary acoustic measures on the Broadway Belt voice suggest uniqueness in this type of vocal production. This study objectively compared the acoustic production of the Broadway Belt voice in four elite and four average belters. Three casting directors evaluated the vocal quality of 20 musical theater majors proficient in the singing style referred to as belting. Each belter sang two specified vocalizes as well as six short excerpts from the belting repertoire. The raters judged the belters on a set of seven perceptual parameters (loudness, vibrato, ring, timbre, focus, nasality, and registration breaks) and reported an overall score. Initially, Pearson product-moment correlation coefficients were calculated and reported for perceived loudness, vibrato, ring, timbre, focus, and nasality for the elite and average groups. Then, significant acoustic results related to vocal intensity, amplitude and magnitude of vibrato, increased spectral energy in the expected Singer's Formant area, and trends in $F1$ – $F2$ characteristics were assessed. Overall patterns of these results suggest the elite belters maintained a greater magnitude of vocal vibrato, a brighter vocal quality on some vowels, and different harmonic—formant relationships than average belters. Specific relevant data related to these acoustical events will be the focus of this presentation.

10:30

3aMU6. Acoustic and aerodynamic characteristics of Country-Western, Operatic and Broadway singing styles compared to speech. Robert E. Stone, Jr. (Retired, Dept. of Otolaryngol., Vanderbilt Univ. Medical Ctr., 5921 Woodland Hills Dr., Nashville, TN 37211, restone@comcast.net), Thomas F. Cleveland (Vanderbilt Univ. Medical Ctr., Nashville, TN 37212), and P. Johan Sundberg (Voice Res. Ctr., Stockholm, Sweden)

Acoustic and aerodynamic measures were used to objectively describe characteristics of Country-Western (C-W) singing in a group of six premier performers in a series of studies and of operatic and Broadway singing in a female subject with professional experience in both styles of singing. For comparison purposes the same measures also were applied to individuals while speaking the same material as sung. Changes in pitch and vocal loudness were investigated for various dependent variables, including subglottal pressure, closed quotient, glottal leakage, $H1$ – $H2$ difference [the level difference between the two lowest partials of the source

spectrum and glottal compliance (the ratio between the air volume displaced in a glottal pulse and the subglottal pressure)], formant frequencies, long-term-average spectrum and vibrato characteristics (in operatic versus Broadway singing). Data from C-W singers suggest they use higher sub-glottal pressures in singing than in speaking. Changes in vocal intensity for doubling sub-glottal pressure is less than reported for classical singers. Several measures were similar for both speaking and C-W singing. Whereas results provide objective specification of differences between operatic and Broadway styles of singing, the latter seems similar to features of conversational speaking style.

Contributed Papers

10:50

3aMU7. Analysis and enhancement of country singing. Matthew Lee and Mark J. T. Smith (Ctr. for Signal and Image Processing, School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0250)

The study of human singing has focused extensively on the analysis of voice characteristics. At the same time, a substantial body of work has been under study aimed at modeling and synthesizing the human voice. The work on which we report brings together some key analysis and synthesis principles to create a new model for digitally improving the perceived quality of an average singing voice. The model presented employs an analysis-by-synthesis overlap-add (ABS-OLA) sinusoidal model, which in the past has been used for the analysis and synthesis of speech, in combination with a spectral model of the vocal tract. The ABS-OLA sinusoidal model for speech has been shown to be a flexible, accurate, and computationally efficient representation capable of producing a natural-sounding singing voice [E. B. George and M. J. T. Smith, *Trans. Speech Audio Processing* **5**, 389–406 (1997)]. A spectral model infused in the ABS-OLA uses Generalized Gaussian functions to provide a simple framework which enables the precise modification of spectral characteristics while maintaining the quality and naturalness of the original voice.

Furthermore, it is shown that the parameters of the new ABS-OLA can accommodate pitch corrections and vocal quality enhancements while preserving naturalness and singer identity. Examples of enhanced country singing will be presented.

11:05

3aMU8. Comparison of hearing and voicing ranges in singing. Eric J. Hunter and Ingo R. Titze (Dept. of Speech Pathol. and Audiol., Natl. Ctr. for Voice and Speech, The Univ. of Iowa, Iowa City, IA 52242)

The spectral and dynamic ranges of the human voice of professional and nonprofessional vocalists were compared to the auditory hearing and feeling thresholds at a distance of one meter. In order to compare these, an analysis was done in true dB SPL, not just relative dB as is usually done in speech analysis. The methodology of converting the recorded acoustic signal to absolute pressure units was described. The human voice range of a professional vocalist appeared to match the dynamic range of the auditory system at some frequencies. In particular, it was demonstrated that professional vocalists were able to make use of the most sensitive part of the hearing thresholds (around 4 kHz) through the use of a learned vocal ring or singer's formant. [Work sponsored by NIDCD.]

11:20–11:40

Panel Discussion

WEDNESDAY MORNING, 30 APRIL 2003

ROOM 102, 8:30 TO 10:00 A.M.

Session 3aNSa

Noise: Flow Noise

Courtney B. Burroughs, Chair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

Contributed Papers

8:30

3aNSa1. Prediction of the interior pressure oscillations included by flow over a cavity. Jin-Seok Hong, Jong Beom Park, and Luc Mongeau (Ray W. Herrick Labs., School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907-1077, hong14@ecn.purdue.edu)

The flow-excited resonance of an open cavity exposed to a grazing flow was investigated experimentally. Flow visualization provided information about the position and the strength of the vortices formed over the orifice. A model based on vortex sound theory was used to predict the frequency and the relative magnitude of the self-sustained interior pressure oscillations. Values for the vortex circulation strength and the convection speed were estimated using a least mean square method. The analytical predictions were compared with experimental values and are in reasonably good agreement with the experimental data. The influence of spoilers and other control devices on the excitation mechanism was investigated.

8:45

3aNSa2. Fan broadband noise from the ingestion of inhomogeneous turbulence. D. A. Lynch, W. K. Blake (NSWC, Carderock Div., 9500 MacArthur Blvd., West Bethesda, MD 20817), and R. Martinez (Anteon Corp./CAA, Cambridge, MA 02140)

Calculations of broadband noise from pre- and post-swirl turbomachines are compared to recent wind-tunnel data. The inflow cases addressed are: (1) grid-generated turbulence for which near-homogeneous conditions prevailed (for the pre-swirl case); and (2) sharply inhomogeneous turbulence encapsulated within the wakes of the rotating upstream blades (for the post-swirl case). Computed results help explain the degree to which strip aerodynamics succeeds in reproducing the noise data, particularly upon comparison to more rigorous estimates that account for spanwise wave numbers in the flow-to-lift transfer function. The paper also reports on the merit of basing the calculation of broadband noise on a circumferentially averaged inflow for case (2), whose turbulence is

thereby rendered artificially homogeneous. The presentation ends with a discussion of the experimental and numerical means whereby the size of the predominant integral scale of the wake-encapsulated turbulent flow was determined, and the extent to which the two sets of values agreed.

9:00

3aNSa3. Acoustical analysis and modeling of reciprocating compressors, noise produced by gas pulsation, using four-pole method. I. Ali T. Herfat and Robert V. Seel (Emerson-Copeland Corp., 1675 W. Campbell Rd., Sidney, OH 45365)

Presented in Paper I are the fluid-structure interactions, structural dynamics, and thermodynamic analyses of reciprocating compressors (such as air conditioning and refrigeration reciprocating compressors). The compressor performance can be analyzed using the follows criteria: (1) thermodynamic model of the cylinder process, using the polytropic process model for thermodynamic model of cylinder; (2) suction and discharge valves dynamics analysis and modeling; (3) the valves modeling and the cylinder volume calculation; (4) Effective flow area and effective force area models.

9:15

3aNSa4. Sources of flow noise due to time-varying tip relief. R. Martinez (Cambridge Acoust. Assoc./Anteon Corp., 84 Sherman St., 3rd Level, Cambridge, MA 02140, rxmartinez@anteon.com)

This paper begins with a brief review of the conventional cause of shed vorticity in two-dimensional unsteady lifting flows around airfoils. It then poses the following three-dimensional theoretical problem in support of experimental work recently conducted at Penn State/ARL: Take a wing tip of arbitrary three-dimensional shape and place a wall near it such that its plane is normal to that of flight. Let the wall be sinusoidally corrugated in the direction of the flow and be convected by it, so that an observer stationed on the wing's tip experiences time-varying influences determined by the wall's passing serpentine pattern. Our presentation will describe a linearizing perturbation analysis whereby the ratio of the wall's sinusoidal amplitude to its wavelength becomes a small parameter. The zeroth-order problem becomes then one of steady three-dimensional aerodynamics barring the presence of external fluctuations. The first-order problem is fundamentally unsteady and driven, as usual, by the zeroth-order one. An unconventional source of flow noise emerges: the ultimately unacceptable axial changes in the strength of trailing vortices, as brought about by their irregular images behind the corrugated wall, lead to new neutralizing shed vortices and harmonic lift.

9:30

3aNSa5. Acoustical analysis and modeling of reciprocating compressors, noise produced by gas pulsation, using four-pole method. II. Ali T. Herfat (Emerson-Copeland Corp., 1675 W. Campbell Rd., Sidney, OH 45365) and Robert V. Seel (Emerson-Copeland Corp., Sidney, OH 45365)

Presented in Paper II is the noise analysis of reciprocating compressors (such as air conditioning and refrigeration reciprocating compressors) using the four-pole method. The gas pulsation noise inside compressor head cavities, mufflers, and through-valves can be analyzed by applying the FPM. This method formulates the characteristics of acoustic elements by establishing a relationship between their input and output gas pressures and volume flow rates. When the acoustic elements in the system (compressor) are connected at points between them, the FPM allows an easy assembly of element equations to obtain system acoustical model.

9:45

3aNSa6. Sound quality metric development for the Air Conditioning and Refrigeration Institute. Kathleen K. Hodgdon and Russell C. Burkhardt (Appl. Res. Lab., Penn State, University, Park, PA 16802)

The current method of assessing acoustic signatures from residential air conditioning units is defined in the Air Conditioning and Refrigeration Institute (ARI 270) 1995 Standard for the Sound Rating of Outdoor Unitary Equipment. This research project was designed to assess the efficacy of the metric to predict consumer preference. The original metric and modified versions of that metric were evaluated as tools for use in today's market. The ARI 270 metric was implemented in software with additional features and the flexibility to modify the features applied. Numerous product acoustic signatures were analyzed and compared using the original metric and various modifications of that metric. As a result of that analysis a set of synthesized signatures was generated that targets problem areas for an application of the metric to typical signature configurations. A subjective jury evaluation was conducted to establish the consumer preference for those synthesized signatures. A statistical correlation of the various configurations was conducted to assess the degree of the relationship between the subjective preferences and the various metric calculations. Recommendations were made for modifications to improve the current metric's ability to predict a subjective preference. [Programming completed by Jonathan Peters (ARL/PSU). Research supported by ARI.]

WEDNESDAY MORNING, 30 APRIL 2003

ROOM 102, 10:15 TO 11:30 A.M.

Session 3aNSb

Noise: Transportation and Community Noise

Bennett M. Brooks, Chair

Brooks Acoustics Corporation, 27 Hartford Turnpike, Vernon, Connecticut 06066

Contributed Papers

10:15

3aNSb1. Aircraft system noise prediction: Past, present, and future. Robert A. Golub and Joe W. Posey (Aeroacoustics Branch, NASA Langley Res. Ctr., M.S. 461, Hampton, VA 23681)

Aircraft system noise prediction is necessary to estimate the community noise impact of future aircraft and to estimate the noise impacts of changes in propulsion systems, airframes, or operations of current aircraft.

Aircraft system noise is the sum of noise generated by various components of the propulsion system and the various components of the airframe including the landing gear. Predicting noise on the ground from an aircraft flyover requires estimating the noise generated by the many contributing sources during the flyover as the flight conditions change, summing these sources as a function of time, and propagating the resultant combined source through the atmosphere to the observer location. NASA introduced the Aircraft Noise Prediction Program (ANOPP) about 30 years ago and continually upgraded and extended the code prediction capability. The

history of ANOPP will be reviewed along with current efforts to make it more useful as a design tool. A proposed new systems prediction program, AVATAR, will be less empirical and capable of predicting community noise from unconventional aircraft planforms. Innovative/unconventional aircraft configurations will be required to meet aggressive noise goals in the future.

10:30

3aNSb2. Prediction-based aircraft flyover noise synthesis. Stephen A. Rizzi and Brenda M. Sullivan (Structural Acoust. Branch, NASA Langley Res. Ctr., Hampton, VA 23681, s.a.rizzi@larc.nasa.gov)

Investigation of community noise impact of aircraft flyover noise through subjective testing requires time histories of the acoustic pressure at one or more listener positions. The use of recorded aircraft flyover noise for this purpose is often problematic. Factors contributing to this include extraneous noise in the recordings (natural or man-made), a finite number of fixed recording positions, and the cost of conducting the flight test. Perhaps the most limiting factor is the inability to examine proposed aircraft, engines, flight procedures, and other conditions or configurations for which recorded data are unavailable. Synthesis of aircraft flyover noise as an alternative to recordings is thus desirable. An approach to the synthesis of broadband and narrowband aircraft component noise is presented based upon output from community noise prediction codes. These codes typically calculate aircraft flyover noise at designated listener positions in the frequency domain. The necessity to convert these to the time domain presents two challenges both relating to the fact that the predictions are, in essence, time averaged. The first of these involves the reconstruction of phase. The second is the synthesis of temporal fluctuations known to exist in real data. A semiempirical approach to calculating these effects is also presented.

10:45

3aNSb3. Characterization of highway traffic noise generated by rigid pavement contraction joints. Lawrin T. Ellis, Christopher Niezrecki (Dept. of Mech. and Aerosp. Eng., Univ. of Florida, P.O. Box 116250, Gainesville, FL 32611), and David Bloomquist (Univ. of Florida, Gainesville, FL 32611)

Contraction joints in rigid (concrete) pavements are required to permit expansion of each monolithic section of roadway. At higher speeds, the major source of highway noise is attributed to vehicle tire/roadway interaction. Current concerns about limiting the impact of highway traffic noise has forced transportation agencies to consider strategies to control noise generated by tire/roadway interaction. Within this work the difference in noise generated by 1/4- vs 3/8-in. joint widths is conducted. The study focuses on passenger vehicles including a sedan and a light duty van/truck. Both vehicle in-cabin and roadside noise levels are measured for vehicle speeds of 50, 60, and 70 miles per hour. For the sedan, the minimum and maximum observed in-cabin differences were determined to be 1.08 and 1.82 dB(A), respectively. Minimum and maximum observed roadside differences are 1.19 and 2.58 dB(A), respectively. Van tests resulted in mini-

um and maximum observed in-cabin differences of 0.60 and 1.09 dB(A) and minimum and maximum observed roadside differences of 1.05 and 3.18 dB(A), respectively. This paper contains details of reference standards, test methods, and the results obtained.

11:00

3aNSb4. Chicago Transit Authority train passenger environmental noise study. Matt R. McDuffee and Chris Karner (Columbia College Chicago, 600 S. Michigan Ave., Chicago, IL 60605, Attn: Dominique Cheenne, dcheenne@popmail.colum.edu)

The Chicago Transit Authority (CTA) train system is referred to as the "L" because most of the track throughout the city is elevated. Passengers riding aboard the "L" are often subjected to high levels of noise due to the aging metal girder system that the tracks are perched on, as well as some train cars that are in disrepair. The environmental acoustics class of nine students at Columbia College Chicago decided to quantify exactly how much noise an "L" passenger is subjected to. Using a Quest 2900 integrating sound level meter the class split up and took L_{eq} measurements on all of the seven train lines. Each line was tested in both directions of travel twice, with the meter taking samples every 3 s, which added up to a total of approximately 65 000 samples. The data were then averaged and synthesized into a graph using ESRI ArcView software. The graph is actually a map of the CTA "L" system that is color coded according to the L_{eq} level that the passengers are subjected to between each station. It was interesting to see the difference in noise levels according to the type of track construction.

11:15

3aNSb5. A new approach to control noise from entertainment facilities: Active control and measurement of amplified community noise. Richard J. Peppin (Scantek, Inc., 7060-L Oakland Mills Rd., Columbia, MD 21046, PeppinR@ScantekInc.com) and Joan Casamajó (DICESVA, Villar 20, Barcelona 08041, Spain)

While traffic noise is perhaps the most pervasive of community noises, much of the contribution now comes from amplified sound: live music, discos, theme parks, and exercise studios. Those producing the sound or music want it loud and those not interested want to be protected against noise. Noise limits at the receiving or producing property line must be met for the minimum community acceptance. However the time-, and perhaps the spatially-, varying sound in entertainment facilities is often constantly modified (and maybe monitored) near the source of the sound. Hence it is hard to relate and to control the sound at the property line. This paper presents a unique noise control device. It is based on the octave band "transfer function" between the sound produced in the entertainment area and the noise received at the property line. The overall insulation can be measured and is input to the instrument. When a noise level limit is exceeded at the receiver, due to the amplified interior noise at the facility, the sound output of the device is automatically controlled to reduce the noise. The paper provides details of the design and possible abatement scenarios with examples.

3a WED. AM

Session 3aPA

**Physical Acoustics, Noise and Engineering Acoustics: Wind Noise in Outdoor Measurements—
Characterization and Reduction**

Richard Raspet, Chair

*National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677***Chair's Introduction—8:25*****Invited Papers*****8:30**

3aPA1. Turbulent pressure statistics in the atmospheric boundary layer from large-eddy simulation. Natasha L. Miles (Dept. of Meteorol., Pennsylvania State Univ., 503 Walker Bldg., University Park, PA 16802), John C. Wyngaard, and Martin J. Otte (Duke Univ., Durham, NC 27708)

Turbulent pressure fluctuations advected past a sensor contribute to wind noise and can thus significantly degrade acoustic signals. In this study, large-eddy simulation is used to calculate the turbulent pressure field in three types of atmospheric boundary layers. A Poisson equation is used to represent turbulent pressure as the sum of mean-shear, turbulence-turbulence, subfilter-scale, Coriolis, and buoyancy parts. At variance-containing scales in the free-convection case, turbulent-turbulent pressure dominates over the entire boundary layer. It dominates also up to midlayer in the forced-convection case; above that mean-shear pressure dominates. In the stable case mean-shear pressure dominates over the entire layer. Part of the inertial subrange of the pressure spectrum is resolved in the forced- and free-convection cases; it is dominated by the turbulence-turbulence pressure and has a three-dimensional spectral constant of 4.0. This agrees well with quasi-Gaussian predictions but is a factor of 2 less than results from direct numerical simulations at moderate Reynolds numbers. Although past measurements of turbulent pressure have been hampered by instrumental problems, such measurements that could be used to determine the inertial subrange pressure spectral constant would be most useful. [Work supported in part by the U.S. Army through the National Center for Physical Acoustics.]

8:50

3aPA2. Recent progress with atmospheric noise-reducing filters. Michael Hedlin (9500 Gilman Dr., La Jolla, CA 92093-0225)

Wind noise is one of the key issues facing the community of acousticians who monitor the atmosphere for long-period signals from natural atmospheric phenomena or from man-made sources, such as nuclear weapons tests. Presently, a network of infrasound observing sites is being constructed to provide global coverage of acoustic sources in the atmosphere. It is well known that spatial filters can be used to attenuate noise due to atmospheric turbulence near the recording point and increase the ratio of signal-to-noise. Rosette spatial filters are currently used at the new infrasound monitoring stations and have been found to provide a significant reduction of noise across the frequency band of interest to the acoustic monitoring community. Research continues into modifications to existing rosette filters that might improve their performance. Additional research is providing insight into entirely different filters, such as wind barriers. Some of the monitoring systems array sites are located on oceanic islands at locations where spatially extensive noise filters will not be practical. This practical limitation fuels our interest in effective, spatially compact filtering methods. This presentation gives an overview of the various filters currently being used and research into improvements for these filters and into entirely different filter designs.

9:10

3aPA3. The use of an infrasound microphone array to study wind noise spectra and correlation. F. Douglas Shields and Carrick Talmadge (Nat'l. Ctr. for Physical Acoust., Coliseum Dr., University, MS 38677)

A three dimensional array of infrasound sensors of original design has been constructed and used to study wind generated pressure signals in the frequency range from 0.1 to 100 Hz. The ten sensors in each arm of the array are 2 feet apart. An ultrasonic anemometer ten feet off the ground was used to make simultaneous measurements of the three components of the wind velocity. Several sets of data have been taken in open fields with different ground cover. The data have been spectrally analyzed and, over a limited frequency range, the velocity and pressure variations found to obey the 5/3 and 7/3 power law that is expected for the inertial range. A study has also been made of the dependence of the correlation between the pressure signals and the sensor separation. The coherence of the pressure signals indicates that the convection velocity is nearly independent of frequency, and the correlation has an exponentially decaying sinusoidal dependence on the sensor separation. The array has also been used successfully to localize infrasound sources. [Work supported by the U.S. Army Armament Research Development and Engineering Center.]

3aPA4. Influence of turbulence frequency and flow unsteadiness on noise reduction of windscreens. Z. Charlie Zheng and Ning Zhang (Mech. and Nuclear Eng. Dept., Kansas State Univ., Manhattan, KS 66506)

Windscreens are widely used in outdoor microphone measurement. However, in many of these applications, low-frequency wind noise interferes with the signals and the performance of measurement microphones significantly depends on the correct design of the windscreen that is used to maximize the signal-to-noise ratio of the sensing system. There are two possible noise sources around a windscreened microphone: one is from the turbulence carried in the incoming flow; the other is from the wake vortex shedding due to interaction between the windscreen and the wind. Recently, we investigated the effects of windscreens on low-frequency wind noise reduction using a steady-state computational fluid dynamic model. The justification of the use of a steady-state flow model was based on experimental data by Morgan and Raspet that showed that when the corresponding frequency was below 10 Hz, the wind noise reduction was almost constant. Under this low-frequency assumption, we found that the wind noise reduction increases with decrease of the Reynolds number. We are currently concentrated on higher frequencies where the wind noise reduction is no longer independent of frequencies and unsteady fluid dynamics is required to provide pressure fluctuation information on the windscreen surface.

3aPA5. Performance comparisons of microphone windscreens. Edward R. Maniet, Jr. (Textron Systems, 201 Lowell St., Rm. 3101, Wilmington, MA 01887, emaniet@systems.textron.com)

The effectiveness of microphone windscreens is evaluated using a unique low-noise wind tunnel facility that enables controlled, repeatable experiments to be performed [E. R. Maniet, Jr., J. Acoust. Soc. Am. **111**, 2373–2374 (2002)]. Wind noise levels are characterized as a function of frequency, mean wind velocity and turbulent velocity component for several standard windscreen types, including porous foam balls and hollow shell windscreens. The experimental results are used to develop a model for wind noise levels based upon Strasbergs [J. Acoust. Soc. Am. **83**, 544–548 (1988)] dimensional scaling analysis. Also presented is a comparison of microphone wind noise levels for omnidirectional and supercardioid microphone elements.

10:10–10:25 Break

3aPA6. An investigation of outdoor wind noise reduction by spherical windscreens. Jeremy Webster and Richard Raspet (Dept. of Phys. and Astron., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677)

Phelps [RCA Rev. **3**, 203–212 (1938)] proposed that wind noise pressure fluctuations measured at the center of a spherical wind screen are the area averages resultant of the pressure fluctuations at the surface of the screen. If this hypothesis is applied to the steady state pressure distribution around a spherical windscreen, good agreement between data and theory should be achieved for turbulence structures which are large compared to the windscreen. In this paper we report on experiments which extend the work done by Morgan (Ph.D. dissertation, University of Mississippi, 1992) on spherical windscreens. In this experiment, probe microphones placed within reticulated foam windscreens were used in conjunction with anemometers placed directly in the airflow near the windscreen to investigate the correlation between incident flow and the resulting pressure measurements throughout the foam.

Contributed Papers

3aPA7. Effect of atmospheric pressure fluctuations on low frequency and infrasound detection. Carrick L. Talmadge and Doug Shields (Univ. of Mississippi, NCPA, Oxford, MS 38677, clt@olemiss.edu)

A fundamental difficulty in low-frequency and infrasound detection is the so-called “wind noise” problem. At higher (>50 Hz), this wind noise is usually associated with the obstruction formed by the pressure probe body. In this case, most of this wind noise can be adequately removed by the use of an appropriate wind screen. At lower frequencies (<50 Hz), the magnitude of the intrinsic atmospheric fluctuations appears to be much larger than the fluctuations induced by the sensor body. A new large-element (>30 elements) 3-axis pressure sensor array is used to quantify the relative magnitude of the intrinsic and induced pressure fluctuations. Signal processing algorithms are discussed which enhance the signal-to-noise ratio over that obtained using standard beam forming algorithms, by utilizing the localized nature of many of the fluctuations across the array. The effect of the height of the sensor above ground on the magnitude of the observed pressure fluctuations is also examined using a 3-axis linear pressure sensor array. Finally, the performance of a single-element porous hose array is compared to that of a large-element array.

3aPA8. Measuring immeasurable sound pressure levels. Kenneth E. Gilbert, Carrick L. Talmadge (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Coliseum Dr., University, MS 38677), and Xiao Di (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

The AWSUM(K) processors developed by R. A. Wagstaff make possible the detection of narrowband signals deeply buried in wind noise. However, the signal-to-noise ratio one obtains with the AWSUM(K) processors is not the ratio of signal power to noise power, but a nonphysical ratio that depends on the coherence of the signal and the noise distribution. A recently developed theory, used in conjunction with the AWSUM(K) processors, allows the recovery of a physically meaningful signal-to-noise ratio, even when wind noise prevents a direct measurement of the signal. Thus, it is possible not only to detect a signal buried in wind noise, but also to estimate the physical sound pressure level. Such a capability can be important for identifying a source using, for example, the ratio of harmonic amplitudes produced by a particular engine in a ground vehicle. The theory for recovering the physical sound pressure level will be discussed and applications to field data will be presented.

3aPA9. Wind noise and the spectrum of atmospheric turbulence pressure fluctuations. D. Keith Wilson (U.S. Army Cold Regions Res. and Eng. Lab., 72 Lyme Rd., Hanover, NH 03755)

Previous research [S. Morgan and R. Raspet, J. Acoust. Soc. Am. **92**, 1180–1183 (1992)] has shown that wind noise is predominantly caused by pressure fluctuations intrinsic to the turbulent atmospheric flow. Therefore it should be possible to predict wind noise from models for turbulent pressure spectra in the atmosphere. Based on simple dimensional analysis and an application of Taylor's hypothesis, the inertial-subrange power spectrum for turbulent pressure fluctuations should be proportional to $f^{-7/3}$, where f is frequency. But more recent atmospheric observations and theoretical arguments [J. D. Albertson, G. G. Katul, M. B. Parlange, and W. E. Eichinger, Phys. Fluids **10**, 1725–1732 (1998)] suggest that the power spectrum actually goes as $f^{-3/2}$. In this paper, it is shown that the $f^{-3/2}$ dependence predicts a much slower decay in wind noise with increasing acoustic frequency than is typically observed. Possible reasons for this discrepancy are discussed.

3aPA10. A feasibility study of air-coupled ultrasonic vibrometers. Andi Petculescu and James Sabatier (Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., University, MS 38677, apetcule@olemiss.edu)

Several key issues arising when an ultrasonic field is used to probe small surface vibrations are presented. Sum and difference frequencies (sidebands) can be produced by the Doppler shift of the primary probe frequency incurred at the vibrating boundary and by the nonlinear parametric interaction with the radiation from the oscillating surface. Carrier and surface vibration amplitudes, the condition of the oscillating surface, and ambient parameters, influence the received signal. Wind critically affects the reliable operation of air-coupled acoustic sensors. As wind energy increases, the sensor detects a progressively richer turbulent spectrum that drowns the information-bearing sidebands. Experimental data of the scattering of ultrasonic energy by a turbulent environment is shown, aimed at searching for ways to actively recover the vibration sidebands. [Work supported by ONR.]

WEDNESDAY MORNING, 30 APRIL 2003

ROOM 204, 8:00 A.M. TO 12:00 NOON

Session 3aPP

Psychological and Physiological Acoustics and Animal Bioacoustics: Honoring the Contributions of Jozef Zwislocki

Robert D. Frisina, Chair

Otolaryngology Division, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, New York 14642-8629

Chair's Introduction—8:00

Invited Papers

8:10

3aPP1. Jozef Zwislocki: Impact on models of coding in the auditory nerve. Murray B. Sachs (Dept. of Biomed. Eng., Ross 720, Johns Hopkins Univ., Baltimore, MD 21205, msachs@bme.jhu.edu)

The auditory nerve has long been considered a window on the biophysical mechanisms of cochlear transduction and the most carefully characterized aspect of the responses of single auditory-nerve fibers has been the tuning curve. Perhaps the most intensively studied question in auditory theory is: What is the relationship between the shapes of these tuning curves and basilar membrane displacements? The basilar membrane measurements of Georg von Békésy stimulated a generation of basilar-membrane modelers, none more notable than Joe Zwislocki, who was awarded the first von Békésy Medal by the Acoustical Society in 1985. The impact of Zwislocki's basilar membrane models on our understanding of auditory nerve tuning will be reviewed. The properties of auditory-nerve discharge patterns are also shaped by the filtering properties of the hair cell/synapse complex. The major contributions of Joe and his students to our understanding of this filtering through their elegant experimental and modeling studies of adaptation in the auditory nerve will be presented. Throughout his career, Joe Zwislocki has maintained an active interest in loudness summation and his work in relating the input/output characteristics of auditory-nerve fibers to loudness will be highlighted.

8:35

3aPP2. Jozef Zwislocki in the post-von Békésy era of cochlear physiology. William S. Rhode (Univ. of Wisconsin, 1300 University Ave., Madison, WI 53706, rhode@physiology.wisc.edu)

Professor Joe Zwislocki is a rare individual who has a strong background in mathematics and hydrodynamics that he applied to the physiological characterization of cochlear function. Realizing the preeminent need for more accurate measurements of cochlear mechanical properties, he undertook the study of properties of the tectorial membrane (TM). His measurements are unique in that they remain the only *in vivo* measurements of stiffness of the TM. In 1979, he introduced longitudinal effects into cochlear modeling that today has been shown to result in realistic cochlear mechanical transfer curves. A difference between the filter properties seen in auditory nerve fibers and basilar membrane (BM) mechanics led to a leading theory of separate resonances for organ of Corti and TM. He developed a model of how this would affect the shear motion between the TM and reticular lamina. Additional studies were undertaken that substituted recordings in the cells of Hensen for the very difficult direct observation of BM mechanics that epitomize the ingenuity and resourcefulness of Joe Zwislocki.

9:00

3aPP3. Jozef Zwislowski's integrated approach to psychoacoustics. Rhona P. Hellman (Dept. of Psych., Northeastern Univ., 360 Huntington Ave., Boston, MA 02115, hellman@neu.edu)

Early in the psychophysics phase of his illustrious and varied scientific career, Joe Zwislowski recognized that loudness is of key importance for understanding signal processing by the auditory system. To better comprehend just how loudness can be incorporated into basic auditory theory, Joe sought to determine the loudness-intensity relation down to near threshold levels. Together with Hellman, this aim motivated the development of absolute magnitude-scaling procedures. Later, Joe demonstrated that absolute scaling yields results compatible with nonmetric measures of loudness additivity. His search for a comprehensive theory of auditory function led him to deduce that both the observed proportionality between loudness and sound intensity near threshold and the compressive nonlinearity in the loudness function at moderate-to-high levels are generated by the peripheral auditory system. These innovative concepts were incorporated into his theoretical analyses of temporal summation and central masking. They also provided the basis of a loudness model formulated to describe loudness growth in quiet and in noise. Not only are Joe Zwislowski's extraordinary insights compatible with recent loudness measures, his loudness model can be extended successfully to predict the growth of loudness characteristic of cochlear-impaired hearing. [Work supported by NEDO, Japan.]

9:25

3aPP4. Understanding cochlear function through auditory-nerve activity: A Zwislowski perspective. Richard A. Schmiedt (Dept. of Otolaryngol.-Head and Neck Surgery, P.O. Box 250550, Medical Univ. of South Carolina, Charleston, SC 29425, schmiera@musc.edu)

Dr. Joe Zwislowski served as my dissertation advisor during those tumultuous years in cochlear physiology when our understanding of outer hair cell (OHC) function was evolving from that of a sensory cell to that of a mechanical amplifier. Spoendlin had recently demonstrated that 90%–95% of auditory-nerve afferents originated from inner hair cells (IHCs), but the characteristics of IHC receptor potentials remained an enigma. Otoacoustic emissions and OHC electromotility were terms yet to be defined. Theories relating auditory-nerve activity to basilar-membrane mechanics included concepts of second filters, basilar-membrane nonlinearities, and phase opposition. It was a fertile time for theories and experiments attempting to describe a black-box system that did not yield its mysteries easily. Around 1977, IHC receptor potentials were found to be as sharply tuned as auditory-nerve responses, and the era of cochlear micromechanics began. Joe Zwislowski, as usual, has played a primary role in defining this new era, utilizing the relationships between the OHC stereocilia and the tectorial membrane as his modeling clay.

9:50

3aPP5. Adapting to Jozef Zwislowski. Robert L. Smith (Depts. of Bioengineering & Neurosci. and the Inst. of Sensory Res., Syracuse Univ., 621 Skytop Rd., Syracuse, NY 13244, bob_smith@isr.syr.edu)

Dr. Joe Zwislowski has been a source of excellent advice throughout my career. However, two very important things he told me were wrong as will be disclosed in the presentation. Nevertheless, among the many fundamental things I learned from him was how one must attempt to take the data, and turn it over and over again until the hidden patterns and quantitative inter-relationships emerge. It was because of Joe that I first began to study auditory-nerve adaptation. His challenge to me was to find a model of adaptation that was not only functionally useful but also physiologically realizable, and that quest is still continuing, on my part and that of many others in our field. When Joe suggested auditory-nerve adaptation as a possible source of psychophysical overshoot, our analysis showed that could be true, but only for an on-frequency overshoot of several dB. Consequently the sources of the large off-frequency overshoot remain enigmatic. New results showing overshoot in cochlear implant listeners will be presented and some implications for acoustic overshoot discussed. However, my overall conclusion is that I still have not adapted to Joe Zwislowski, and he continues to be a significant ongoing source of new insights and analysis.

10:15–10:30 Break

10:30

3aPP6. Understanding loudness: Testing physiological models of intensity discrimination and masking for their consistency with loudness functions. Laurel H. Carney (Depts. of Bioengineering & Neurosci. and the Inst. of Sensory Res., Syracuse Univ., 621 Skytop Rd., Syracuse, NY 13244, laurel_carney@isr.syr.edu)

Joe Zwislowski continues to contribute actively to ongoing scientific discourse at the Institute for Sensory Research (ISR). Naturally, these discussions include, amongst other topics, the relationship between loudness and intensity discrimination. Being in the ISR environment has provided the motivation to test models for intensity discrimination and masking in the context of loudness. This problem is timely in that it relates to Joe's current efforts to assemble a written summary of his contributions to the study of loudness. In this presentation, the ability of recent neural models for intensity discrimination and masking to explain some basic aspects of the loudness functions reported by Joe Zwislowski and his colleagues will be explored. [Work supported by NIDCD-NIH.]

10:55

3aPP7. Jozef Zwislowski's contribution to the understanding of cutaneous sensation. Stanley J. Bolanowski (Inst. for Sensory Res., Dept. of Bioengineering & Neurosci., 621 Skytop Rd., Syracuse Univ., Syracuse, NY 13244, Sandy_Bolanowski@isr.syr.edu)

Whereas Professor Zwislowski is well known for his theoretical and experimental activities that discovered many principles about the auditory system as outlined in this special session, his influence on research efforts and contributions to the knowledge base of the cutaneous sensory system has not been as widely appreciated. Philosophically, he believes that all of the sensory systems have common, as well as different capabilities, and it is this philosophy which led him to explore many of the underlying factors behind

3a WED. AM

somatosensation. This presentation will outline his scientific and philosophical input to the understanding of somatosensation from the level of receptor function to higher cognitive aspects. For example, he has influenced various views regarding tactile psychophysical thresholds and the relationships between sensation magnitude and the Differenz Limen. His theories on temporal summation and thoughts regarding independent tactile channels of communication originating in the periphery and passing on to the central nervous system will also be discussed. Physiologically he was a prominent player in determining transduction mechanisms of one of the prototypical mechanoreceptors found within the skin, the Pacinian corpuscle. Indeed, how somatosensation comes about has progressed greatly from his oftentimes unrealized influence.

11:20

3aPP8. The “Zwislocki effect” in my work with cochlear implants. Monita Chatterjee (House Ear Inst., Dept. of Auditory Implants & Percept., 2100 W. Third St., Los Angeles, CA 90057, monita@hei.org)

Joe Zwislocki’s uniqueness as a scientist lies in his ability to transcend methodology, sensory modality, or system—he has spanned the realms of mathematics, engineering, physics, physiology, and psychophysics to study hearing, touch, and vision. As a graduate student, I was amazed and inspired by his unerring ability to find unifying patterns underlying seemingly different phenomena. I became interested in cochlear implants while a graduate student in Joe’s lab. Cochlear implants present a unique opportunity to the sensory scientist: by bypassing the complex frequency analysis of the cochlea and directly stimulating the auditory nerve, they allow us to separate aspects of temporal and spectral (tonotopic) processing by higher centers of the auditory system. In a series of recent experiments, we have exploited this feature of cochlear implants to better understand the processing of complex signals by the auditory system. In this presentation, I will review these and other experiments on loudness, masking, and temporal processing in electrically stimulated hearing. Although Joe has not (as yet) worked on cochlear implants, his approach to sensory systems is undoubtedly the strongest scientific influence (the “Zwislocki effect”) in my work in this area.

Contributed Paper

11:45

3aPP9. A comparison of Levitt and Zwislocki decision rules for use with forced-choice adaptive procedures. Robert S. Schlauch and Edward J. Carney (Dept. of Commun. Disord., 164 Pillsbury Dr. SE, Minneapolis, MN 55455, Schla001@umn.edu)

Forced-choice adaptive procedures enjoy widespread use for the measurement of detection and discrimination thresholds. Zwislocki *et al.* [J. Acoust. Soc. Am. **30**, 254–262 (1958)] proposed an adaptive procedure with a decision rule that targets 75% correct several years before these procedures gained acceptance in psychophysics, and even today little is

known about the statistical properties of this decision rule. This paper evaluates, using computer simulations, the bias and efficiency of Zwislocki’s proposed decision rule in 2-alternative forced-choice (AFC) and 3-AFC procedures under conditions of full attention and inattention. The results for this decision rule compare favorably with two popular rules proposed by Levitt [J. Acoust. Soc. Am. **49**, 467–477 (1971)] that target 70.7% correct and 79.4% correct. In summary, the rule that targets 75% correct (Zwislocki’s rule) produces less biased threshold estimates than the rule that targets 70.7% correct and it is affected less by inattention than the rule that targets 79.4% correct. This research supports the addition of Zwislocki’s rule to the psychophysical toolbox.

WEDNESDAY MORNING, 30 APRIL 2003

ROOMS 105/106, 8:00 TO 9:45 A.M.

Session 3aSAa

Structural Acoustics and Vibration and Signal Processing in Acoustics: Interior Noise in Aircraft and Rocket Fairings

Robert L. Clark, Chair

*Department of Mechanical Engineering and Material Science, Duke University, Box 90300,
Durham, North Carolina 27708-0300*

Chair’s Introduction—8:00

Invited Papers

8:05

3aSAa1. Development of a device for reduction of low-frequency sound transmission reduction in small launch vehicles. Steven F. Griffin and Steven A. Lane (Boeing SVS, 4411 The 25 Way NE, Ste. 350, Albuquerque, NM 87109)

Launch acoustic loads have the potential to damage sensitive payloads within a payload fairing, often requiring more structural mass to withstand these loads than would otherwise be necessary to survive launch. A typical payload fairing will exhibit several cavity resonance modes related to its axial length. These modes are excited both by the vibrations transmitted through the structure during launch, and by acoustic ground reflections coupled into the system during the initial launch phase. Conventional acoustic blankets mitigate the acoustic environment within a payload fairing, but are generally only effective above 250 Hz. In this work we

present an Adaptive Vibro-Acoustic Device (AVAD), which is designed to actively and passively absorb acoustic energy in a payload fairing at frequencies below 250 Hz. To date, a prototype has been designed and tested for application to a sounding rocket experiment, Vibro-Acoustic Launch Protection Experiment (VALPE), being conducted by the Air Force Research Laboratory. Prototype test results as well as projected performance in flight will be presented.

8:30

3aSAa2. Limits on the performance of Distributed Vibration Absorbers for the control of broadband disturbances. Marty Johnson, Tony Harris, and Chris Fuller (Vib. and Acoust. Labs., Virginia Tech, VA 24061-0238)

Distributed Vibration Absorbers (DVAs) have been shown to be useful passive control devices for reducing the vibration levels on lightly damped aerospace structures subjected to broadband excitation. These devices work by coupling to the modes of the structure and efficiently adding damping. One advantage of the DVA over the classic vibration absorber (VA) is that it acts over a large area instead of at a single attachment point. This paper investigates how this spatial distribution allows the absorber mass to be re-used multiple times in order to control multiple modes. In order to quantify the benefit of using a DVA this paper compares the mass of the DVA under investigation to the mass of a set of VAs that produce the same vibration reduction. This mass ratio (R) is therefore a measure of the effectiveness of the DVA. The effect of the complexity of the DVA design and the spatial extent of the DVA on R , are both investigated and results are presented. Using a computational model of a cylinder it will be shown that large mass reductions can be achieved using DVAs but that they must act over a large area and are susceptible to variations in structural properties. [Work supported by Boeing.]

8:55

3aSAa3. Active control of payload fairing noise using distributed active vibration absorbers. Arnaud Charpentier (Vibro-Acoust. Sci., Inc., 12555 High Bluff Dr., Ste. 310, San Diego, CA 92130, arnaud@vasci.com), Marty E. Johnson, and Chris R. Fuller (Virginia Tech, Blacksburg, VA 24060)

High sound pressure inside a launch vehicle fairing during lift-off can damage the payload. Interior levels of up to 140 dB between 60 and 250 Hz are mostly due to exhaust plume noise combined with the limited transmission loss of lightweight composite fairings and little acoustic damping in the fairing volume. Past work using passive and hybrid passive/reactive noise control devices has shown that their limitations are mostly due to packaging volume and weight penalty. The objective of this work is to design a lightweight, compact, and low electrical power active noise control system to reduce the fairing interior sound level. Hybrid active/passive actuators such as Smart Foam (Couche and Fuller, *Proceedings of Active 1999*, Ft. Lauderdale, FL, pp. 609–620) and Distributed Active Vibration Absorbers (Marcotte, Fuller, and Johnson, *Proceedings of Active 2002*, ISVR, Southampton, England, pp. 535–546) are optimized for fairing noise control. The latter have been used to increase the transmission loss of the fairing. Active noise control test results on a sub-scale, sandwich composite fairing are presented. The global interior acoustic response due to airborne exterior excitation is minimized using an adaptive multiple-input, multiple-output feedforward controller. [Work supported by the Air Force Research Laboratory, Space Vehicles Directorate (AFRL).]

9:20

3aSAa4. The control of rocket fairing interior noise with a networked embedded system. Kenneth D. Frampton (Dept. of Mech. Eng., Vanderbilt Univ., VU Station B 351592, Nashville, TN 37235)

Numerous investigations have been conducted with the purpose of attenuating the acoustic environment within rocket payload fairings. These, to date, theoretical and experimental laboratory studies have demonstrated a great deal of success. However, practical applications to this, and other large-scale noise control problems, have been limited in their success. These limitations are due to non-scalable control systems, weight constraints and complexity. This work seeks to address these limitations by investigating the use of an array of networked embedded processors to control the interior acoustics of a rocket fairing. This networked embedded system consists of numerous computationally elements, paired with appropriate sensors and actuators, that communicate with each other over a wired or wireless network. The goal of the network is to minimize the interior acoustic level while expending a minimum amount of energy. Results from the simulation of such a control system will demonstrate the effectiveness of such an approach. These results will also be compared with those obtained by traditional, centralized control architectures.

Session 3aSAb

Structural Acoustics and Vibration: Computational Methods

Lonny L. Thompson, Chair

Department of Mechanical Engineering, Clemson University, 219 EIB, Box 340921, Clemson, South Carolina 29634-0921

Contributed Papers

10:00

3aSAb1. Adaptive time-discontinuous Galerkin finite element methods for acoustic scattering. Dantong He and Lonny L. Thompson (Dept. of Mech. Eng., Clemson Univ., Clemson, SC 29634-0921)

Comprehensive self-adaptive procedures with efficient sparse multi-level iterative solution algorithms for the time-discontinuous Galerkin space-time finite element method (DGFEM) including high-order accurate nonreflecting boundary conditions (NRBC) are developed for acoustic scattering problems. An h -adaptive space-time strategy is employed based on a superconvergent patch recovery (SPR) technique, together with a temporal error estimate. The use of sub-time steps with pre-integrated local space-time elements to efficiently track waves in space-time are also demonstrated. For accurate data transfer (projection) between meshes, new superconvergent interpolation (SI) procedures are developed. Numerical studies of transient acoustic scattering demonstrate the accuracy, reliability and efficiency gained from the adaptive strategy. [Work supported by NSF.]

10:15

3aSAb2. Parallel iterative solution of large-scale acoustic scattering problems using exact non reflecting conditions on distributed memory computer systems. Cristian Ianculescu and Lonny L. Thompson (Dept. of Mech. Eng., Clemson Univ., Clemson, SC 29634-0921)

Parallel iterative methods for fast solution of large-scale acoustic radiation and scattering problems are developed using exact Dirichlet-to-Neumann (DtN) nonreflecting boundaries. For elongated scatterers such as submarines, it is shown that the generalization of the DtN to elliptical/spheroidal artificial boundaries improves significantly the computational efficiency of accurate finite element methods for the solution of acoustic scattering problems. The outer-product structure of the DtN map is exploited as a low-rank update of the system matrix to efficiently compute the matrix-by-vector products found in Krylov subspace based iterative methods. For the complex non-Hermitian matrices resulting from the Helmholtz equation, a distributed-memory parallel BICG-STAB iterative method is used in conjunction with a hybrid parallel SSOR/Jacobi preconditioner. The domain decomposition with interface minimization was performed to ensure optimal inter-processor communication. For the distributed memory architectures tested, including Linux/Intel Beowulf clusters, when implemented as a low-rank update, the nonlocal character of the DtN map shows little impact on the scale up or parallel efficiency compared to approximate local boundary conditions. [Work supported by NSF.]

10:30

3aSAb3. Hybrid NAH formulations for an arbitrary object in a nonfree field. Sean Wu (Dept. of Mech. Eng., Wayne State Univ., Detroit, MI 48202)

Previous nearfield acoustical holography (NAH) formulations based on the Helmholtz integral theory and boundary element method (BEM) are effective for an arbitrary object in a free field, although a true free field is nonexistent. While a modified Helmholtz equation least squares (HELS) method using expansions of both outgoing and incoming spherical waves can provide an approximate solution for such a scenario [Wu and Zhao, J. Acoust. Soc. Am. **111**, 2409 (2002)], the reconstruction accuracy may be

unsatisfactory when the source is of arbitrary geometry. In this paper, a hybrid NAH formulation that combines a modified HELS method, the Helmholtz integral theory, BEM, and regularization techniques (e.g., modified Tikhonov regularization, generalized cross validation) is developed for reconstructing acoustic radiation from an arbitrary object in a nonfree field. This hybrid formulation has the advantages of the modified HELS method and BEM-based NAH, and represents a significant improvement over the combined HELS method [Wu and Zhao, J. Acoust. Soc. Am. **112**, 179–188 (2002)]. The input data are collected on a conformal surface at close range, so the evanescent waves can be captured and the reconstruction accuracy can be improved. Moreover, the majority of the input data are calculated but not measured, and the reconstruction efficiency is enhanced. [Work supported by NSF.]

10:45

3aSAb4. Further developments of a high-frequency broadband energy-intensity boundary element method. Jerry W. Rouse and Linda P. Franzoni (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708-0300, jwrouse@me1.egr.duke.edu)

The prediction of the spatial mean square pressure distribution within enclosed high-frequency broadband sound fields is computationally intensive if determined on a frequency-by-frequency basis. Recently an energy-intensity boundary element method has been developed which employs uncorrelated broadband directional energy sources to predict expeditiously such pressure distributions. The source strength and directivity associated with the energy sources, distributed over enclosure boundaries, were determined in an iterative manner. Here, further refinement of the method is presented including the direct determination of source strength and directivity without iteration. Additionally, recent improvements allow for specifying the degree to which the reflected field is modeled as diffuse relative to specular for a given element. Good agreement between the improved energy-intensity boundary element method and experimental measurements and/or exact analytical solutions are shown. [Work sponsored by NSF.]

11:00

3aSAb5. An expanded spherical wave expansion for arbitrary sound fields. Nassif E. Rayess (Univ. of Detroit Mercy, Detroit, MI 48219)

Spherical wave function expansions as means of describing the acoustic field radiated by arbitrarily shaped objects are very convenient and gaining popularity. The HELS method for acoustic holography is one methodology advocating the use of spherical wave functions as a basis for the solution of the inverse acoustic problem. Such methodologies only provide an approximation to the actual sound field and as such suffer from errors, albeit understood to be bounded and manageable in a large number of cases. Based on the sound radiation model of a volume distribution of monopole sources, these approximation errors are found to be due to the presence of monopole sources outside the largest inscribed virtual sphere. Monopole sources outside that virtual sphere are not accounted for in the expansion and thus come out as errors. These errors are greatly reduced if the spherical wave function expansion is modified to account for the largest number of monopole sources in the model. This is accomplished by employing the addition theorem for spherical harmonics. This modifica-

tion complicates the problem mathematically and might require a greater rate of physical sampling, however the reduction in the reconstruction errors might prove beneficial.

11:15

3aSab6. The use of a waveguide expansion to model landmine detection using acoustic to seismic coupling. Doru Velea (Planning Systems, Inc., 12030 Sunrise Valley Dr., Reston Plaza I, Ste. 400, Reston, VA 20191) and Roger Waxler (Natl. Ctr. for Physical Acoust., The Univ. of Mississippi, 1 Coliseum Dr., University, MS 38655)

The use of a waveguide of sufficiently large radius to simplify the modeling of the infinite space response of a buried landmine to airborne sound has been investigated. It was previously determined that if the ground is modeled as an effective fluid, an efficient and rapidly converging algorithm can be obtained. The effective fluid has been replaced with an elastic solid. For such a model, this technique fails to converge sufficiently rapidly to the infinite space limit. The surface wave supported by the shear sector gets excited by the landmine. The decay of this surface wave with distance is too slow, forcing one to use a waveguide of a prohibitively large radius. In conclusion, this technique at best gives qualitative results. [Work sponsored by the U.S. Army Communications–Electronics Command, Night Vision and Electronics Sensors Directorate.]

11:30

3aSab7. Acoustical wave propagator technique for time-domain analysis of dynamic stress in a step plate. Shuzhi Peng and Jie Pan (School of Mech. Eng., The Univ. of Western Australia, 35 Stirling Hwy., Crawley, WA 6009, Australia, speng@mech.uwa.edu)

In this paper, we introduce an explicit acoustical wave propagator technique to describe the time-domain evolution of acoustical waves in two-dimensional plates. This technique uses a combined scheme with

Chebyshev polynomial expansion and fast Fourier transformation for implementation of the operation of the acoustical wave propagator. We also apply the acoustical wave propagator for studying dynamic stress in a step plate in time-domain.

11:45

3aSab8. 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 reg. 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 quasi-periodic source, placed on the ground surface or at 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 last are included in 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 under ground 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 used approach a series of measurements was carried out near the experimental part of monorail road designed in Moscow. Both calculation and measurements results are presented in the paper.

WEDNESDAY MORNING, 30 APRIL 2003

ROOM 206, 8:00 A.M. TO 12:00 NOON

Session 3aSC

Speech Communication: Voices in the Neighborhood: Production, Perception and Anatomy (Poster Session)

Lori L. Holt, Chair

Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213

Contributed Papers

All posters will be on display from 8:00 a.m. to 3:00 p.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 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon. To allow for extended viewing time, posters will remain on display until 3:00 p.m.

3aSC1. Gestural overlap of stop-consonant sequences: Evidence from analysis and synthesis. Sherry Zhao and Kenneth N. Stevens (Res. Lab. of Electron. and Dept. of EECS, MIT, Cambridge, MA 02139, szhao@mit.edu)

This study uses an analysis-by-synthesis approach to discover possible principles governing the coordination of oral and laryngeal articulators in the production of English stop-consonant sequences. Individual recordings were made of two male and two female native American-English speakers reading phrases which include voiced and voiceless stop consonants in word-initial (V#CV) and word-final (VC#V) positions, as well as in VC#CV stop–stop consonant sequences. Articulatory timing estimates were made based on analyzing acoustic data including formant move-

ments, closure durations, release bursts, and spectrum shape at low frequencies. Based on the gestural estimates, the same consonant sequences were generated using HLsyn, a quasiarticulatory synthesizer. The synthetic utterances were acoustically and perceptually compared to the actual utterances in order to verify and refine the articulatory timing estimates from which possible principles could be derived. Preliminary results agree with earlier findings of more overlapping of oral gestures in sequences with front-to-back order of place of articulation than those with back-to-front order [Chitoran, Goldstein, and Byrd, *Lab. Phonology* 7, 419–448 (2002)]. Furthermore, overlapping of laryngeal gestures is suggested by the smaller acoustical loss at the glottis in vowels after voiced–voiceless sequences than voiceless–voiceless sequences.

3aSC2. Infants in cocktail parties. Rochelle S. Newman (Dept. of Hearing & Speech Sci., Univ. of Maryland, College Park, MD 20742, newman@hesp.umd.edu)

Most work on listeners' ability to separate streams of speech has focused on adults. Yet infants also find themselves in noisy environments. In order to learn from their caregivers' speech in these settings, they must first separate it from background noise such as that from television shows and siblings. Previous work has found that 7.5-month-old infants can separate streams of speech when the target voice is more intense than the distractor voice (Newman and Jusczyk, 1996), when the target voice is known to the infant (Barker and Newman, 2000) or when infants are presented with an audiovisual (rather than auditory-only) signal (Hollich, Jusczyk, and Newman, 2001). Unfortunately, the paradigm in these studies can only be used on infants at least 7.5 months of age, limiting the ability to investigate how stream segregation develops over time. The present work uses a new paradigm to explore younger infants' ability to separate streams of speech. Infants aged 4.5 months heard a female talker repeat either their own name or another infants' name, while several other voices spoke fluently in the background. We present data on infants' ability to recognize their own name in this cocktail party situation. [Work supported by NSF and NICHD.]

3aSC3. Cocktail party effect in infants: Visual information and speech segmentation in noise. George Hollich (Dept. of Psych. Sci., Purdue Univ., 703 Third St., West Lafayette, IN 47907-2004, ghollich@purdue.edu)

What are infants' abilities to use what they see to segment speech in a noisy environment? Infants often find themselves in situations far louder and more complex than the acoustic isolation chambers of traditional infant testing. The current series of studies used the headturn preference procedure (with video familiarization) to examine 7.5-month-old infants' abilities to use visual/auditory correlations to reliably attend to and segment a given speech stream in the face of a distracting voice. Results indicated that in contrast to seeing a static face, infants succeeded at segmentation when a dynamic visual display of the face of the talker matched the acoustic passage. That is, when two blended voices were of equal loudness, infants could use visual correspondences to reliably recognize words presented in the matching video. Furthermore, they did so even if the video display was changed to a synchronized oscilloscope pattern, rather than a face. They also succeeded when the video display was simply a synchronized flashing light. These results suggest that this ability is related to domain-general sensitivities for intermodal correspondences, rather than any face-specific effects, and suggest a mechanism whereby face-specific effects could develop.

3aSC4. Two reasons not to bring your children to cocktail parties. Susan Nittrouer, Melanie Wilhelmsen, Kathy Shapley, Kathi Bodily (Ctr. for Persons with Disabilities, Utah State Univ., 6840 Old Main Hill, Logan, UT 84322), and Tom Creutz (Boys Town Natl. Res. Hospital, Omaha, NE 68131)

Two kinds of noise can affect speech perception at cocktail parties: others' speech (sloping spectrum) and the environment (level spectrum). Recognition scores for phonetically balanced monosyllabic words in level noise at five SNRs for adults and children (4 to 8 years) were collected, and compared to those of Nittrouer and Boothroyd for speech-shaped noise [J. Acoust. Soc. Am. **87**, 2705–2715 (1990)]. All listeners showed similar effects of level noise, and children's results matched those for speech-shaped noise. However, adults showed a 22% advantage for speech-shaped over level noise, suggesting they use high-frequency speech elements not masked by the sloping long-term spectrum of speech. Next labeling of words differing in syllable-final voicing (using vocalic duration and offset transitions, neither high-frequency cues) was examined for children and adults in quiet and in level noise. All listeners performed similarly in quiet. Adults' results for quiet and noise matched, but children showed a decreased weighting of transitions in noise compared to quiet,

suggesting enhanced masking of this information. Thus children's speech perception in noise is impaired both because they fail to use high-frequency cues available against the background of others' speech and they experience enhanced masking of formant transitions, their preferred source of information. [Work supported by NIDCD Grant No. DC-00633.]

3aSC5. Effects of response format on speech intelligibility in noise: Results obtained from open-set, closed-set, and delayed response tasks. Cynthia G. Clopper, Adam T. Tierney, and David B. Pisoni (Speech Res. Lab., Dept. of Psych., Indiana Univ., Bloomington, IN 47405, cclopper@indiana.edu)

Many word recognition studies over the last 40 years have used forced-choice closed-set tasks, based on the assumption that closed-set and open-set tasks differ only in the level of chance performance. However, Sommers, Kirk, and Pisoni (1997) found that lexical competition and talker variability produce robust effects on performance only in open-set tasks, suggesting fundamental differences in the task demands and processing strategies between open- and closed-set tasks. In the present study, listeners were asked to recognize spoken words degraded by a bit-flipping algorithm in three response formats: open-set, closed-set "before," and closed-set "after." In the closed-set "before" condition, the six response alternatives were presented 1 s prior to the onset of the auditory signal. In the closed-set "after" condition, the response alternatives were presented 1 s after the auditory signal. Results revealed significant effects of lexical competition and talker variability only in the open-set task. These findings suggest that even a delay of 1 s is not adequate to induce the task demands observed in open-set word recognition tests. [Work supported by NIH.]

3aSC6. A test of the orthographic recoding hypothesis. Daniel E. Gaygen (Psych. Dept., Ithaca College, Ithaca, NY 14850, dgaygen@ithaca.edu)

The Orthographic Recoding Hypothesis [D. E. Gaygen and P. A. Luce, *Percept. Psychophys.* **60**, 465–483 (1998)] was tested. According to this hypothesis, listeners recognize spoken words heard for the first time by mapping them onto stored representations of the orthographic forms of the words. Listeners have a stable orthographic representation of words, but no phonological representation, when those words have been read frequently but never heard or spoken. Such may be the case for low frequency words such as jargon. Three experiments using visually and auditorily presented nonword stimuli tested this hypothesis. The first two experiments were explicit tests of memory (old–new tests) for words presented visually. In the first experiment, the recognition of auditorily presented nonwords was facilitated when they previously appeared on a visually presented list. The second experiment was similar, but included a concurrent articulation task during a visual word list presentation, thus preventing covert rehearsal of the nonwords. The results were similar to the first experiment. The third experiment was an indirect test of memory (auditory lexical decision task) for visually presented nonwords. Auditorily presented nonwords were identified as nonwords significantly more slowly if they had previously appeared on the visually presented list accompanied by a concurrent articulation task.

3aSC7. Perception of coda voicing from properties of the onset and nucleus of *led* and *let*. Sarah Hawkins (Dept. of Linguist., Univ. of Cambridge, Sidgwick Ave., Cambridge CB3 9DA, UK) and Noël Nguyen (Université de Provence, France)

Syllable-onset [l] is longer and often has different (usually lower) F_2 frequency before a voiced coda. Five experiments (E1–E5) explore the perceptual power of these properties and f_0 . Listeners identified as *led* or *let* synthetic syllables whose latter part was replaced by noise 80 ms after vowel onset. The duration and F_2 frequency of [l] steady-state were varied in E1–E5, f_0 of [l] in E1, vowel formant frequencies in E2–E5, and stimulus randomization principles in E3–E5. In E1, [l] had one of six

durations, two $F2$ frequencies, and two $f0$ starting frequencies. E2 continued $F2$ differences into the vowel, as in natural speech. E3–E5 varied $F2$ independently in [I] and vowel rather than together; stimuli were presented in one session, or blocked by vowel $F2$. Shorter [I]s, higher $f0$, and higher $F2$ in [I]+vowel produced more *let* responses. $F2$ in [I] (alone) mainly affected responses when vowel quality was constant. However, listeners learned which cues were systematic, and some who initially used $F2$ frequency switched to duration of [I] relatively late in a session. The results suggest coda voicing is a property of the whole syllable, and support word recognition models that are sensitive to systematic variation in subtle phonetic detail.

3aSC8. The effect of syllabification and gemination on $F2$ onsets in Swedish. Augustine Agwuele (Dept. of Linguist., Univ. of Texas, 1 University Station B5100, Austin, TX 78712)

The quest to explain the continuousness of speech on the physical level has been dominated by the co-production theory [Hman, J. Acoust. Soc. Am. **41** (1966)]. According to this view, the production of a VCV sequence involves a diphthongal movement from V1-to-V2, with a superimposed consonantal gesture. However, data from recent studies are at variance with this position—Modaressi [Ph.D. dissertation, UT Austin, 2002], Perkell [Coarticulation Strategies Speech Commun. **5**, 47–68 (1986)]. These studies document a trough-phenomena, which suggests a discontinuity in muscular activity during the production of the consonant. This paper provides additional acoustic evidence in support of sequential programming of consonant-vowel events as advocated by Joos [Acoustic Phonetics Lang. **24** (1948)]. It examines symmetrical VCV sequences in Swedish natural speech with syllable boundaries and duration of consonant gemination altered to produce different types of temporal interval between V1 and V2; i.e., V#CV, VCC#V, VC#CV, VCC#CV. $F2$ of V1mid, V1offset, V2onset, and V2mid were measured. Locus equations were plotted for all VC contacts. Statistical analysis of these data show: (1) de-activation of tongue movement at the CV boundary, (2) a reduction of the influence of V1 on V2 as a function of increasing consonant duration, and (3) a weak degree of CV coarticulation.

3aSC9. Perceptual discontinuities and categorization: Implications for speech perception. Lori L. Holt (Dept. of Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213, lholt@andrew.cmu.edu), Andrew J. Lotto (Washington State Univ., Pullman, WA 99164), and Randy L. Diehl (Univ. of Texas, Austin, TX 78712)

Behavioral experiments with infants, adults and nonhuman animals converge with neurophysiological findings to suggest that there is a discontinuity in auditory processing of stimulus components differing in onset time by about 20 ms. This discontinuity has been implicated as a basis for boundaries between speech categories distinguished by VOT. Here, we investigate how this discontinuity interacts with the learning of novel perceptual categories. Adult listeners were trained to categorize a nonspeech acoustic cue that mimics the temporal distinction of VOT. One group of listeners learned categories with a boundary coincident with the perceptual discontinuity. Another group learned categories defined such that the perceptual discontinuity fell within a category. Listeners in the latter group required significantly more experience to reach criterion categorization performance. The evidence of interactions between the perceptual discontinuity and the learned categories extended to generalization tests as well. It has been hypothesized that languages make use of perceptual discontinuities to promote perceptual distinctiveness among sounds within a language inventory. The present data suggest that these influences interact with category learning. As such, learnability may play a predictive role in selection of language sound inventories. Moreover, it may be possible to observe predictable learning effects in infant speech perception.

3aSC10. Perceptual overshoot with speech and nonspeech sounds. Radhika Aravamudhan and John W. Hawks (School of Speech Pathol. & Audiol., Kent State Univ., P.O. Box 5190, Kent State Univ., Kent, OH 44240, raravam1@kent.edu)

One of the basic quests in speech perception research has been to find the differences or similarities in the mechanisms involved in the perception of speech and nonspeech sounds. The current study will address the differences in perception of speech and nonspeech signals by comparing the perceptual overshoot in synthetic vowels and sinewave acoustic replicas of the synthetic vowels. Lindblom and Studdert-Kennedy (1967) demonstrated that the perceptual boundary for steady state vowels and that for vowels in a CV context with $F2$ transition are different. They called this phenomenon a perceptual compensation or perceptual overshoot. In the current study the perceptual boundaries for synthetic steady state vowels, steady state sinewave acoustic replicas of vowels, a vowel in the CV context with $F2$ transition and sinewave acoustic replicas of vowels in the CV context are compared. The results will be discussed in the poster. For Speech Communication Best Student Paper Award.

3aSC11. The effects of regional dialect on vowel intelligibility from a cross-linguistic perspective. Elaina M. Frieda (Dept of Psych., Auburn Univ., 226 Thach Hall, Auburn, AL 36894-5214, friedem@auburn.edu) and Robert A. Fox (The Ohio State Univ., Columbus, OH 43210)

The present experiment is a continuation of a previously reported study that examined intelligibility of English vowels as a function of dialect spoken by native speakers of English and Japanese. The purpose of this research is to assess how regional dialectal variations affect vowel intelligibility for native and non-native speakers of English. Native English and Japanese subjects were recorded in two divergent dialectal regions of the United States (Ohio and Alabama). These tokens were then employed in a perceptual experiment where native English and Japanese listeners from Ohio and Alabama identified the English vowels. To date, perceptual data from only Ohio native English and Japanese subjects have been reported. A further analysis of the data including Alabama native English and Japanese listeners revealed that native English speakers from Ohio obtained the highest intelligibility scores overall (for example, all four listener groups identified Ohio English more accurately than all other groups). Additionally, native Japanese speakers from Alabama received the lowest overall intelligibility scores. The tentative results of this study imply that non-native speakers of English that are exposed to a nonstandard dialect may have deleterious effects on comprehension.

3aSC12. The recognition of accented and unaccented English words by native speakers of Spanish and English. Satomi Imai, James Flege (Div. of Speech and Hearing Sci., Univ. of Alabama at Birmingham, CH20, 1530 3rd Ave. S., Birmingham, AL 35294, imais@uab.edu), and Amanda Walley (Univ. of Alabama at Birmingham, Birmingham, AL 35294)

This study examined effects of foreign accent and lexical factors (word frequency and neighborhood density) on the recognition of English words in noise. Two groups of native Spanish (NS) adults differing in overall degree of foreign accent (FA) in English (weaker versus stronger FAs) participated as well as a native English (NE) group. Participants identified words that had been spoken by a NE and a NS speaker. It was hypothesized that: (1) the stronger FA group would have less nativelike phonological representations than the weaker FA group, and so would benefit more from hearing Spanish-accented English words; (2) words in dense lexical neighbors would require finer discrimination of English sounds, necessitating more nativelike phonological representations. The results showed that the weaker FA group recognized as many low-neighborhood-density English words as the NE group. For high-neighborhood-density words, the weaker FA group recognized fewer unaccented English words than the NE group, but more accented English words. The stronger FA group recognized as many accented words as the other groups, but fewer

unaccented words. The differences between the weaker and stronger FA groups were interpreted as reflecting a change in their phonological representations for English words. [Work supported by NIH.]

3aSC13. The influence of linguistic experience on pitch perception in speech and nonspeech sounds. Tessa Bent, Ann R. Bradlow (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208, t-bent@northwestern.edu), and Beverly A. Wright (Northwestern Univ., Evanston, IL 60208)

How does native language experience with a tone or nontone language influence pitch perception? To address this question 12 English and 13 Mandarin listeners participated in an experiment involving three tasks: (1) Mandarin tone identification—a clearly linguistic task where a strong effect of language background was expected, (2) pure-tone and pulse-train frequency discrimination—a clearly nonlinguistic auditory discrimination task where no effect of language background was expected, and (3) pitch glide identification—a nonlinguistic auditory categorization task where some effect of language background was expected. As anticipated, Mandarin listeners identified Mandarin tones significantly more accurately than English listeners (Task 1) and the two groups' pure-tone and pulse-train frequency discrimination thresholds did not differ (Task 2). For pitch glide identification (Task 3), Mandarin listeners made more identification errors: in comparison with English listeners, Mandarin listeners more frequently misidentified falling pitch glides as level, and more often misidentified level pitch "glides" with relatively high frequencies as rising and those with relatively low frequencies as falling. Thus, it appears that the effect of long-term linguistic experience can extend beyond lexical tone category identification in syllables to pitch class identification in certain nonspeech sounds. [Work supported by Sigma Xi and NIH.]

3aSC14. Perceptual prothesis in native Spanish speakers. Rachel M. Theodore and Anna M. Schmidt (School of Speech Pathol. and Audiol., Kent State Univ., Kent, OH 44240, rachel_theodore@yahoo.com)

Previous research suggests a perceptual bias exists for native phonotactics [D. Massaro and M. Cohen, *Percept. Psychophys.* **34**, 338–348 (1983)] such that listeners report nonexistent segments when listening to stimuli that violate native phonotactics [E. Dupoux, K. Kakehi, Y. Hirose, C. Pallier, and J. Mehler, *J. Exp. Psychol.: Human Percept. Perform.* **25**, 1568–1578 (1999)]. This study investigated how native-language experience affects second language processing, focusing on how native Spanish speakers perceive the English clusters /st/, /sp/, and /sk/, which represent phonotactically illegal forms in Spanish. To preserve native phonotactics, Spanish speakers often produce prothetic vowels before English words beginning with /s/ clusters. Is the influence of native phonotactics also present in the perception of illegal clusters? A stimuli continuum ranging from no vowel (e.g., "sku") to a full vowel (e.g., "esku") before the cluster was used. Four final vowel contexts were used for each cluster, resulting in 12 sCV and 12 VsCV nonword endpoints. English and Spanish listeners were asked to discriminate between pairs differing in vowel duration and to identify the presence or absence of a vowel before the cluster. Results will be discussed in terms of implications for theories of second language speech perception.

3aSC15. The representation of consonant clusters in the mental lexicon. Lisa J. Incognito and James R. Sawusch (Dept. of Psych., SUNY at Buffalo, Buffalo, NY 14260, incognit@acsu.buffalo.edu)

Previous work has shown that the perception of a phoneme in a syllable is influenced by the number of similar sounding words [lexical neighborhood (Newman, Sawusch, and Luce, 1997)]. This previous work determined neighborhoods for target syllables using a one phoneme change rule. For example, bow, bath and mouth are neighbors of bowth. The present work focused on how consonant clusters are represented in the mental lexicon. Nonsense syllables composed of initial consonant clusters

followed by a vowel and final consonant were used as stimuli. Two rules were used to compute the neighborhood for each target syllable. One was the one phoneme change rule used in previous studies. The second treated clusters of consonants as single units in a one unit change rule. Target syllables with differential neighborhoods based on the two rules were the endpoints of the test series. Results to date agree with the one phoneme change rule. These results are consistent with models of word recognition which treat consonant clusters as a sequence of phonemes. [Work supported by NIDCD Grant R01DC00219 to SUNY at Buffalo.]

3aSC16. Sonority contours in word recognition. Sean McLennan (Indiana Univ., Memorial Hall, Rm. 322, 1021 E. Third St., Bloomington, IN 47405-7005, mmclenna@indiana.edu)

Contrary to the Generativist distinction between competence and performance which asserts that speech or perception errors are due to random, nonlinguistic factors, it seems likely that errors are principled and possibly governed by some of the same constraints as language. A preliminary investigation of errors modeled after the child's "Chain Whisper" game (a degraded stimulus task) suggests that a significant number of recognition errors can be characterized as an improvement in syllable sonority contour towards the linguistically least-marked, voiceless-stop-plus-vowel syllable. An independent study of sonority contours showed that approximately half of the English lexicon can be uniquely identified by their contour alone. Additionally, "sororities" (groups of words that share a single sonority contour), surprisingly, show no correlation to familiarity or frequency in either size or membership. Together these results imply that sonority contours may be an important factor in word recognition and in defining word "neighborhoods." Moreover, they suggest that linguistic markedness constraints may be more prevalent in performance-related phenomena than previously accepted.

3aSC17. Neighborhood effects on speech-perception abilities of adults and children aged 6–9 years. Joan E. Sussman, Devon Dee, and Diana Curcio (Dept. of Communicative Disord. and Sci., Univ. at Buffalo, Buffalo, NY 14214)

According to Luce and Pisoni (1998), the mental lexicon is organized into sparse and dense neighborhoods. In perception studies with adults, typically words from sparse neighborhoods are more quickly identified than words from dense neighborhoods, likely due to a competition effect. In production however, words from dense neighborhoods are produced more quickly than words from sparse neighborhoods, probably due to a facilitation effect. In young children, these so-called "neighborhood effects" have been shown to be absent (Charles-Luce and Luce, 1990), related to the overall smaller size of the lexicon. The current investigation studied single-feature perception by 10 children, aged 6–9 years and 10 adults, looking at differences due to neighborhood density. Each target word was presented in one of two semantically neutral carrier sentences in three levels of babble noise: no noise, +5-dB S/N ratio, and 0-dB S/N ratio (equal signal and noise levels). Results showed that children performed best in quiet and that words from dense neighborhoods were perceived more accurately than those from sparse neighborhoods except for the 0-dB condition. Adults surprisingly performed best in the +5-dB condition and a neighborhood effect shown only for the 0-dB condition.

3aSC18. Children's perception of static noise and static formant cues to stop-consonant place of articulation. Ralph N. Ohde (Dept. of Hearing and Speech Sci., Vanderbilt Bill Wilkerson Ctr., Vanderbilt Univ., Nashville, TN 37212-2197, ralph.n.ohde@vanderbilt.edu)

Children's processing strategies appear to favor dynamic cues such as formant transitions as compared to static cues such as F2 onsets and noise bursts. The purpose of this research was to examine children's perception of place of articulation based only on static cues. Ten children at each of five age levels (3, 4, 5, 6, and 7) and a control group of 10 adults identified

synthesized stop consonants [d g] in two vowel contexts [i a]. The synthesis parameters included variations in *F2* onsets and stop-consonant noise bursts. The *F2* onsets were either “appropriate” or “neutral” for place of articulation. The noise bursts were either short (10 ms) or long (25 ms). Preliminary data show that the *F2* onset is not as salient in children’s perception as in adults’ perception. In addition, children more often than adults categorized neutral *F2* onset stimuli as ambiguous indicating stronger category formation in the latter than former groups. The role of noise bursts was more salient in adult perception than child perception. The findings will provide information on the role of “static” cues, on the perceptual integration of “static” noise and formant cues, and on the influence of sound category formation in perceptual development. [Work supported by NIH and a Vanderbilt University Research Council Grant.]

3aSC19. Cue weighting of static and dynamic vowel properties in children versus adults. Sarah R. Malech and Ralph N. Ohde (Dept. of Hearing and Speech Sci., Vanderbilt Bill Wilkerson Ctr., Vanderbilt Univ., Nashville, TN 37212-2197, ralph.n.ohde@vanderbilt.edu)

The purpose of this study was to determine whether children give more perceptual weight than do adults to dynamic spectral cues versus static cues, when identifying vowel sounds. Three experimental stimulus sets were presented, each with 30-ms stimuli. The first consisted of unchanging formant onset frequencies ranging in value from frequencies for [i] to those for [a], corresponding to a bilabial stop consonant. The second two consisted of either an [i] or [a] onset frequency with a 25-ms portion of a formant transition whose trajectory was toward one of a series of target frequencies ranging from those for [i] to those for [a]. Ten children between the ages of 3;8 and 4;1 and a control group of 10 adults identified each stimulus as [bi] or [ba]. The results showed developmental effects: the children relied more heavily than the adults did on the static formant onset frequency cue to identify the vowels, while the adults appeared to give more equal weight to both static and dynamic cues than the children did. These findings contradict the Developmental Perceptual Weighting Shift theory and are discussed in relation to this theory and other current research on the development of vowel perception.

3aSC20. Perception of place-of-articulation information by monkeys versus humans. Joan M. Sinnott and Casey S. Gilmore (Comparative Hearing Lab, Psych. Dept., Univ. of South Alabama, Mobile, AL 36688)

Four monkeys and six humans representing five different native languages were compared in the ability to categorize natural CV tokens of /b/ vs /d/ produced by four talkers of American-English (two male; two female) in four vowel contexts /i,e,a,u/. A two-choice left/right procedure was used in which percent correct and response time data were compared between species. Both measures indicated striking vowel context effects for monkeys, but none for humans. Specifically, monkeys performed better for back vowels /a,u/ than front vowels /i,e/. Since back vowels have more distinctive *F2* onset transitions differentiating /b/ vs /d/, these results imply that monkey perception is more dependent than human perception on the actual acoustic structure of the syllables. We conclude that humans do not use general mechanisms in place perception, rather they use some sort of special mechanism to eliminate vowel context effects. While monkeys do not provide accurate models of adult humans, they may be able to provide a model of the preverbal human infant before it learns a more speech-specific adult strategy of place information extraction. [Work supported by NIH.]

3aSC21. The role of air pressure and contact force in shaping obstruent consonant onset. Lan Chen (Speech Commun. Group, Res. Lab of Electronics, MIT, 77 Massachusetts Ave., Boston, MA)

Soft tissues (the tongue or lips) are used to form the narrow oral constriction for turbulence noise generation during the production of obstruent consonants. The displacement of the soft tissue subject to oral pressure buildup is comparable to the vertical dimension of the constric-

tion. The contact force during the closure of stop consonants and affricates provides a pressure load over 5 times larger than the air pressure at the surface in contact. It can influence the time variation of the constriction size at onset in the form of elastic energy stored in the compliant structure forming the constriction. A finite element fluid-structure interaction program has been used to simulate the effect of these external forces during the onset of obstruent consonants. Preliminary results from a 2-D tongue tip constriction/closure model show that air pressure and contact force can introduce movement on the order of 0.1–0.2 mm during the first tens of milliseconds after release, which is enough to affect the size of the constriction at onset and the nature of release burst. The results of this kind can be used for speech synthesis, guiding the modification of the trajectories of articulators at the consonant onset. [Work supported by NIH.]

3aSC22. A labial gesture for /l/. Fiona Campbell (Dept. of Linguist., Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada, fionaca@interchange.ubc.ca) and Bryan Gick (Univ. of British Columbia, Vancouver, Canada)

Both in language change and in substitutions during language acquisition and disordered speech, /l/ has often been observed to alternate with labial sounds such as [w] or rounded vowels, particularly in postvocalic position. While there are many possible explanations for this alternation, including acoustic enhancement and articulator coupling, one possibility that has not been tested is whether normal adult speakers of English actually produce lip rounding for /l/. A study was conducted to test for the presence of a labial gesture in normal productions of /l/. Front and side video data of lip positions were collected from three adult English speakers during productions of /l/ and /d/. Significant differences were found for all subjects in lip protrusion (upper and lower) and/or lip aperture (horizontal and vertical) in post-vocalic allophones, as well as between the pre- and post-vocalic allophones of /l/. No significant differences were observed in comparisons of pre-vocalic /l/ and /d/. Results suggest that there is in fact a labial gesture in the post-vocalic allophone of /l/, but not in the pre-vocalic allophone. These findings are consistent with a notion of gestural simplification as a possible explanation for substitutions and in language change. [Research supported by NSERC.]

3aSC23. Language-specific vocal tract configurations during nonspeech. Bryan Gick (Dept. of Ling., Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada) and Clare Cook (Univ. of British Columbia, Vancouver, Canada)

Previous work has been found to be surprisingly low within-speaker variability in baseline articulator positions during inter-utterance nonspeech [Gick, *Phonetica* (2002)], raising the question of whether these baseline positions may in fact be active in speech production. If so, then they should be specified and should vary systematically across languages. A study was conducted to test for cross-language differences in inter-utterance articulator positions. Individual video frames were extracted at the midpoint of interutterance pauses in x-ray films of 5 French and 5 English speakers. Measures were made of articulator positions relative to fixed bone points, and values normalized to jaw size. Frames with potentially confounding surrounding phonetic contexts were omitted. Results for lip measures indicate that French speakers have significantly greater protrusion of the lower lip, but significantly less upper lip protrusion, than English speakers. Additional results will be presented for lingual articulators. Thus these baseline vocal tract configurations do appear to be specified differently for different languages. Additional implications will be discussed, such as possible roles these configurations may play in phonology, potential influence on vowel systems (especially schwa), and cross-language vowel normalization. [Research supported by NSERC and NIH.]

3aSC24. Lip interactions and closure duration in labial consonants. Anders Lofqvist (Haskins Labs., 270 Crown St., New Haven, CT 06511)

This study examines interactions between the upper and lower lips in labial consonants where the duration of the oral closure is varied for linguistic purposes. Earlier work has shown that such interactions occur and that their magnitude is partly related to the duration of the oral closure. Lip movements were recorded in native Japanese speakers using a magnetometer system. Results show a positive correlation between the vertical positions of the upper and lower, when both are measured at the point in time where the lower lip reached its peak position during the oral closure. Since the peak vertical position of the lower lip increased with closure duration, the upper lip also had a higher vertical position at the same point in time for long than for short consonants. During the oral closure, the lower lip continued to move upward with a larger movement for long than for short consonants. Due to the mechanical interaction between the lips, the upper lip reversed its downward movement at the onset of the oral closure and also moved upward during the closure, again with a larger movement for long than for short consonants. [Work supported by NIH.]

3aSC25. When there is way too much stress: A first look at the failure behavior of vocal fold tissue. Roger W. Chan (Vocal Fold Physiol. and Biomechanics Lab., Dept. of Audiol. and Speech Sci., Purdue Univ., West Lafayette, IN 47907) and Thomas Siegmund (Purdue Univ., West Lafayette, IN 47907)

In normal voice production, linear small-amplitude vocal fold oscillation occurs only under restricted conditions, such as during oscillation onset and offset. More often than not, phonation in physiological range involves large-amplitude oscillation that is associated with the development of tissue shear stresses and strains much beyond their linear viscoelastic limit, particularly in the vibrating portion of the vocal fold, i.e., the extracellular matrix (ECM). This study targeted one such large-strain nonlinear viscoelastic behavior of the vocal fold by attempting to quantify the stress failure response of the vocal fold ECM under shear. Sheep vocal fold ECM specimens were subjected to torsional, steady shear in a controlled-strain rheometer *in vitro*, at constant strain rates of 0.01, 0.1, and 1.0 rad/s. Results showed that the vocal fold ECM demonstrated nonlinear stress-strain response, as well as failure response when shear strain reached around 100%. Strain-dependent and rate-dependent onset of partial and complete stress failure was observed. A constitutive approach based on a standard-linear cohesive zone model was formulated to characterize the observed rate-dependent failure behavior of the vocal fold ECM. These findings have important implications for predicting tissue injury and for establishing stress and strain safety limits for large-amplitude vocal fold oscillation.

3aSC26. The effect of physical constraints on articulatory variability in English /r/. Jason Brown (Dept. of Linguist., Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada) and Bryan Gick (Univ. of British Columbia, Vancouver, BC V6T 1Z1, Canada and Haskins Labs., New Haven, CT)

In a recent study of American English /r/, Guenther *et al.* [J. Acoust. Soc. Am. (1999)] hypothesized that articulatory "tradeoff" correlations are the result of the need to maintain stable acoustic targets. Their findings included (1) a positive correlation between tongue back height and tongue front horizontal position for 7/7 subjects, (2) a negative correlation between tongue back height and tongue front height for 5/7 subjects, and (3) a positive correlation between tongue front horizontal position and tongue front height for 2/7 subjects. The present study investigates the possibility that these correlations result from physical constraints on the tongue such as volume preservation and palate angle. Continuous sentences from the Wisconsin x-ray microbeam database were analyzed to determine whether these same correlations were present across whole utterances presumably lacking a stable F3 target. Results to date show some significant correlations despite extremely high noise levels. Results will be presented for additional measures using vowels only to reduce noise. These initial re-

sults suggest that the observed correlations may not result from strict acoustic targets for /r/, but rather from internal and external physical constraints on the tongue. [Work supported by NSERC and NIH.]

3aSC27. An articulatory examination of variable word-final flapping at phrase edges and interiors. Teruhiko Fukaya and Dani Byrd (USC Linguist., 3601 Watt Way, GFS 301, Los Angeles, CA 90089-1693)

Formulations of flapping as a symbolic phonological rule suggest clear articulatory differences between flaps and stops and offer no overt explanation for why phrase boundaries should block the alternation. A current articulatory phonetic study of word-final coronals by de Jong (1998) explored the possibility that gradient changes in articulatory dimensions might give rise to quantal difference in the percept of these coronal stops. In particular, increased overlap between the coronal and the following vowel may result in spatial reduction of the coronal gesture via blending. We present results from an experiment with three speakers examining tongue tip movement and acoustic data for word-final [t], phrase internally and at phrase boundaries, with falling and level stress contours. The results confirm that flaps did not occur at phrase boundaries. Phrase-internally, flaps occurred for one subject invariantly and for another subject variably. While the acoustic durations of flaps were expectedly shorter, this did not correspond directly to their closing movement durations. There was also some tendency for flaps to have a shorter acceleration phase and no change in deceleration duration, arguing against an overlap account. Finally, the phrase boundary was associated with articulatory and acoustic lengthening, both incompatible with patterns seen for flapped tokens. [Work supported by USC Zumburg Grant and by NIH.]

3aSC28. An approach to real-time magnetic resonance imaging for speech production. Shrikanth Narayanan (Dept. of Elec. Eng./Linguist., Univ. of Southern California, Los Angeles, CA 90089, shri@sipi.usc.edu), Krishna Nayak (Stanford Univ.), Dani Byrd, and Sungbok Lee (Univ. of Southern California)

Magnetic resonance imaging has served as a valuable tool for studying primarily static postures in speech production. Now, recent improvements in imaging techniques, particularly in temporal resolution, are making it possible to examine the dynamics of vocal tract shaping during speech. Examples include Mady *et al.* (2001, 2002) (8 images/second, T1 fast gradient echo) and Demolin *et al.* (2000) (4–5 images/second, ultra fast turbo spin echo sequence). The present study uses a non 2D-FFT acquisition strategy (spiral *k*-space trajectory) on a GE Signa 1.5T CV/i scanner with a low-flip angle spiral gradient echo originally developed for cardiac imaging [Kerr *et al.* (1997), Nayak *et al.* (2001)] with reconstruction rates of 8–10 images/second. The experimental stimuli included English sentences varying the syllable position of /n, r, l/ (spoken by 2 subjects) and Tamil sentences varying among five liquids (spoken by one subject). The imaging parameters were the following: 15 deg flip angle, 20-interleaves, 6.7 ms TR, 1.88 mm resolution over a 20 cm FOV, 5 mm slice thickness, and 2.4 ms spiral readouts. Data show clear real-time movements of the lips, tongue and velum. Sample movies and data analysis strategies will be presented. Segmental durations, positions, and inter-articulator timing can all be quantitatively evaluated. [Work supported by NIH.]

3aSC29. Speech task and timing considerations in MRI research. Melissa A. Epstein and Maureen Stone (Univ. of Maryland Dental School, 666 W. Baltimore St., Baltimore, MD 21201)

In order to create dynamic magnetic resonance images, subjects must repeat tokens as much as 30 times in a row, therefore, their precision is critical for image quality. Word repetition rate can be 1 or 1.5 s, and is matched to the rate of image recording. Unlike reciting tokens to the beat of a metronome, subjects must say each token during an extended period of noise that lasts about 75% of the repetition period. Furthermore, some tokens may be more difficult to say repetitively in these conditions. There-

fore, this study simulated the acoustics of a MRI recording session to examine the effects of subject, token, and repetition rate on temporal precision. Subjects repeat up to ten mono- to trisyllabic words 30 times each at two repetition rates. Measurements are made of onsets and offsets of one or more phonemes within each word. Preliminary results (five subjects) indicate that there is an effect of subject and word. Further subjects are being measured to corroborate these results and to determine if one repetition rate is better than another and to test subject reliability.

3aSC30. Hearing smiles and smile suppression in natural speech.

Amy K. Drahota and Vasudevi Reddy (Dept. of Psych., Univ. of Portsmouth, King Henry Bldg., King Henry I St., Portsmouth, Hants PO1 2DY, UK)

That we can hear smiles in speech is an established finding. However smiles in natural speech can be of many different kinds, serving different social functions. Previous research has focused only on one category of smile using either smiles “posed” during speech or degraded samples of smiling speech (to disguise the content of utterances). The present study used naturally occurring speech in three foreign languages (Czech, Spanish, and Finnish) and in English, presented to naive native English speakers. Preliminary analyses extracted two kinds of smiles in speech—“open smiles” and “suppressed smiles” in contrast to speech with “no smiles.” Eighty listeners were presented with 18 audio-clips (six of each type) in randomized order. “No smiles” and “open smiles” were successfully identified across all languages. “Suppressed smiles” were most often coded as “no smiles.” An exploration of the pitch contours showed that “suppressed smiles” had a higher average pitch similar to “open smiles,” but a significantly higher variance in pitch range than others. Natural speech produces smiles serving very different functions. The auditory expression and perceivability of emotion is likely to be influenced by subtle social functions only evident in natural interactions.

3aSC31. Effects of the age of cochlear implantation on the quality of the speech produced by profoundly HOH speakers.

Samantha Lake and Betty Kollia (Dept. of Commun. Disord., William Paterson Univ., Wayne, NJ 07470)

Four hearing-impaired children with prelingual, bilateral, severe-to-profound hearing loss were grouped by age, gender, and age at implantation; the younger group consisted of females approximately 6 years old and implanted between 1–2 years of age and the older group consisted of males approximately 14 years old and implanted at 9 years of age. Each child was diagnosed with prelingual hearing loss, was implanted with the Nucleus 24[®] cochlear implant in 1998, and has approximately 4 years of experience using the implant consistently. All subjects receive 8 h of direct instruction with the implant per week, in a school for the deaf that utilizes total communication. Each subject also receives speech therapy in 30-min sessions four times per week and exhibits intelligible speech. Coarticulation in the children’s speech was studied using five consonant-diphthong-consonant pseudowords, in the carrier sentence “it’s a — again.” The recordings were digitized and analyzed acoustically. The results are discussed with reference to the age of cochlear implantation of the children and its role in the quality of their speech.

3aSC32. The effect of spectrally and temporally altered auditory feedback on speech intonation by hard of hearing listeners.

Dragana Barac-Cikoja, Chizuko Tamaki, and Lannie Thomas (Gallaudet Univ., MTB B09, 800 Florida Ave. NE, Washington, DC 20002)

Eight listeners with severe to profound hearing loss read a six-sentence passage under spectrally altered and/or delayed auditory feedback. Spectral manipulation was implemented by filtering the speech signal into either one or four frequency bands, extracting respective amplitude envelope(s), and amplitude-modulating the corresponding noise band(s). Thus, the resulting auditory feedback did not preserve intonation information,

although the four-band noise signal remained intelligible. The two noise conditions and the unaltered speech were each tested under the simultaneous and three delayed (50 ms, 100 ms, 200 ms) feedback conditions. Auditory feedback was presented via insert earphones at the listener’s most comfortable level. Recorded speech was analyzed for the form and domain of the fundamental frequency (f_0) declination, the magnitude of the sentence initial f_0 peak (P1), and the fall–rise pattern of f_0 at the phrasal boundaries. A significant interaction between the two feedback manipulations was found. Intonation characteristics were affected by speech delay only under the spectrally unaltered feedback: The magnitude of P1 and the slope of the f_0 topline both increased with the delay. The spectral smearing diminished the fall–rise pattern within a sentence. Individual differences in the magnitude of these effects were significant.

3aSC33. The impact of rate reduction and increased vocal intensity on coarticulation in dysarthria.

Kris Tjaden (Dept. of Communicative Disord. & Sci., Univ. at Buffalo, 122 Cary Hall, 3435 Main St., Buffalo, NY 14214-3005, tjaden@acsu.buffalo.edu)

The dysarthrias are a group of speech disorders resulting from impairment to nervous system structures important for the motor execution of speech. Although numerous studies have examined how dysarthria impacts articulatory movements or changes in vocal tract shape, few studies of dysarthria consider that articulatory events and their acoustic consequences overlap or are coarticulated in connected speech. The impact of rate, loudness, and clarity on coarticulatory patterns in dysarthria also are poorly understood, although these prosodic manipulations frequently are employed as therapy strategies to improve intelligibility in dysarthria and also are known to affect coarticulatory patterns for at least some neurologically healthy speakers. The current study examined the effects of slowed rate and increased vocal intensity on anticipatory coarticulation for speakers with dysarthria secondary to Multiple Sclerosis (MS), as inferred from the acoustic signal. Healthy speakers were studied for comparison purposes. Three repetitions of twelve target words embedded in the carrier phrase “It’s a — again” were produced in habitual, loud, and slow speaking conditions. F_2 frequencies and first moment coefficients were used to infer coarticulation. Both group and individual speaker trends will be examined in the data analyses.

3aSC34. Adductor spasmodic dysphonia: Relationships between acoustic indices and perceptual judgments.

Michael P. Cannito (School of Audiol. & Speech Pathol., Univ. of Memphis, 807 Jefferson Ave., Memphis, TN 38105, mcannito@memphis.edu), Christine M. Sapienza, Gayle Woodson (Univ. of Florida, Gainesville, FL), and Thomas Murry (Columbia Univ., New York, NY)

This study investigated relationships between acoustical indices of spasmodic dysphonia and perceptual scaling judgments of voice attributes made by expert listeners. Audio-recordings of The Rainbow Passage were obtained from thirty one speakers with spasmodic dysphonia before and after a BOTOX injection of the vocal folds. Six temporal acoustic measures were obtained across 15 words excerpted from each reading sample, including both frequency of occurrence and percent time for (1) aperiodic phonation, (2) phonation breaks, and (3) fundamental frequency shifts. Visual analog scaling judgments were also obtained from six voice experts using an interactive computer interface to quantify four voice attributes (i.e., overall quality, roughness, brokenness, breathiness) in a carefully psychoacoustically controlled environment, using the same reading passages as stimuli. Number and percent aperiodicity and phonation breaks correlated significantly with perceived overall voice quality, roughness, and brokenness before and after the BOTOX injection. Breathiness was correlated with aperiodicity only prior to injection, while roughness also correlated with frequency shifts following injection. Factor analysis reduced perceived attributes to two principal components: glottal squeezing and breathiness. The acoustic measures demonstrated a strong regression relationship with perceived glottal squeezing, but no regression relationship with breathiness was observed. Implications for an analysis of pathologic voices will be discussed.

3aSC35. Open-source software for speech perception research. Robert T. Gayvert (Gayvert Consulting, 16 Chase View Rd., Fairport, NY 14450) and James M. Hillenbrand (Western Michigan Univ., Kalamazoo, MI 49008)

The purpose of this paper is to describe some relatively simple software that can be used for performing such routine tasks as controlling listening experiments (e.g., simple labeling, discrimination using procedures such as ABX, oddity, same-different, etc., sentence intelligibility, magnitude estimation, and so on), recording responses and response laten-

cies, analyzing and plotting the results of those experiments, displaying instructions, and making scripted audio recordings. The software runs under Windows and is controlled by creating text files that allow the experimenter to specify key features of the experiment such as the stimuli that are to be presented, the randomization scheme, inter-stimulus and inter-trial intervals, the format of the output file, and the layout of response alternatives on the screen. Some simple demonstrations will be provided, along with instructions for downloading the software. [Work supported by NIH.]

WEDNESDAY MORNING, 30 APRIL 2003

ROOM 203, 10:00 A.M. TO 12:00 NOON

Session 3aSP

Signal Processing in Acoustics: General Topics on Signal Processing in Acoustics

Ning Xiang, Cochair

National Center for Physical Acoustics, University of Mississippi, P.O. Box 1848, University, Mississippi 38677-1848

Ronald A. Wagstaff, Cochair

National Center for Physical Acoustics, University of Mississippi, P.O. Box 1848, University, Mississippi 38677-1848

Contributed Papers

10:00

3aSP1. Time-reversal maximum-length sequence pairs for simultaneous acoustical dual source measurements. Ning Xiang, Richard Raspet, and Kevin Dillion (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr. University, MS 38677)

A binary sequence generated by a linear shift-register (maximum-length sequence, M-sequence) and its reversed-order sequence form a time-reversal (reciprocal) M-sequence pair. Their correlation property includes a two-valued pulse-like autocorrelation function and a relatively smaller-valued crosscorrelation function. This unique property, along with other number-theory properties, makes the time-reversal MLS pairs suitable for simultaneous dual source crosscorrelation measurements. In the simultaneous measurement of a dual source system, each of the time-reversal MLS pair simultaneously excites one of two separate sources, one or several receiver signals cross-correlate in turn with each of the MLS pairs resulting in impulse responses associated with two separate sources. The proposed method is particularly valuable for system identification tasks with multiple sound/vibration sources and receivers that have to be accomplished in a limited time period. A fast algorithm: fast MLS transform pair is developed for the crosscorrelation. Its feasibility and potential applications in the acoustical measurements are demonstrated using recent field experimental results.

10:15

3aSP2. Modeling and stability analysis of a dynamic system using control theory and signal processing approaches. Ali T. Herfat (Emerson-Copeland Corp., 1675 W. Campbell Rd., Sidney, OH 45365)

Presented in this paper are the following: Analog and digital modeling of a feedback control system; the system stability analysis, using analog and digital methods; and designing the digital system to meet steady-state error and transient response specifications. The digital control systems can control numerous loops at a reduced cost. System modifications can be implemented with software changes rather than hardware changes. Typically, the digital computer is placed in the forward control path, and modeled as a sample and hold network. Digital-to-analog and analog-to-digital converters are required within the system to ensure compatibility of the analog and digital signals. As we proceed with the analysis along with a

case study, these criteria will be become apparent in this paper. An industrial dynamic system will be presented as a Dynamic Control System (the case study) in this paper.

10:30

3aSP3. Phase alignment for coherent and vector processing. Ronald A. Wagstaff (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

Acoustic phase angle is seldom used to achieve signal-to-noise ratio (SNR) gain for single sensor and beam spectral output. One reason is that phase angles generally progress at a nonuniform rate, with unpredictable changes in the direction of rotation. This causes fluctuations in the phase angles with corresponding reductions in gain, often including severe attenuation and cancellation of signals. By adopting a particular analytically convenient definition for phase fluctuations, the fluctuations, thus defined, constitute a set of aligned-phase angles. The aligned-phase angles can be used instead of phase angles to form phase-aligned coherent and vector averages. Doing so achieves SNR gains that equal or exceed the theoretical value of $10 \log(N)$ for perfectly coherent vector averaging (N is the number of elements averaged). This is accomplished without the signal attenuation and cancellation common to coherent and vector averaging. Furthermore, the aligned-phase angles can also be used to automatically detect signals, based on both phase-aligned coherence and appropriate averaging of the aligned-phase angles. Results are included for wind noise in outdoor measurements. [Work supported by U.S. Army Space and Missile Defense Command.]

10:45

3aSP4. Application of discrete wavelet transform to measurement of traffic using road environmental sound from microphone on vehicle. Hironori Hara and Shinji Ozawa (Dept. of Information and Computer Sci., Keio Univ., Japan, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8522, Japan)

It is important for progress of ITS technology to develop the sensor that receives surrounding information and algorithm that extracts information. Recently the media generally used for ITS research is a picture, and not much research using sound is done. But sound sensor is still expected because sound has global information and less calculation cost than a

picture. And research of a probe car that collects surrounding situations by the sensor is generally done. So improvement in sensing technology is desired. Moreover discrete wavelet transform attracts attention as the technique of signal processing. DWT is suitable for the analysis of nonstationary signals as road environmental sound. So the new system of the probe car using a sound sensor is proposed. Measurement of traffic is assumed as a probe car application example. In this paper, the result of count passing oncoming car with DWT to road environmental sound is reported. Road environmental sound was recorded in two kinds of road, uncovered road and tunnel, the validity of algorithm was verified by experiment.

11:00

3aSP5. Influence of spectral shape on ordinary and higher order correlation detection of broadband transients with prefiltering.

Marcella E. Dean, George E. Ioup, Juliette W. Ioup (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148), and Lisa A. Pflug (Stennis Space Center, MS 39529)

The performances of ordinary and higher order correlation detectors of broadband chirps, amplitude modulated by a Gaussian envelope or a cosine taper and of various center frequencies and widths, are compared. It has been previously reported that using several inputs to the detectors, including functionals such as the autocorrelation, the energy spectrum, and wavelet transforms of the signals, provided no significant improvement over ordinary correlation detection. It was only after prefiltering the signal to some fraction of its passband (a partially-known source scenario, i.e., only the passband is known) that quite significant gains were achieved. These gains, up to 12.5 dB in the signal-to-noise ratio at the minimum detectable level, were obtained for Gaussian-modulated chirps. These chirps have spectra which are fairly sharply peaked in their passbands. For comparison, the detection performance for cosine-tapered chirps, whose spectrum is fairly flat in the passband, is examined. Performance gains are not as large, but still reach almost 10 dB. In all partially-known prefiltering, the tricorrelation detector has the best performance, which is achieved using the received signal input for the Gaussian-modulated chirp and the autocorrelation input for the cosine-tapered chirp. The known source performance is also investigated. [Research supported by ONR.]

11:15

3aSP6. Spatio-temporal gradient signal processing for detecting subsurface cracks.

Kenbu Teramoto (Dept. of Mech. Eng., SAGA Univ., Saga-shi 8408502, Japan) and Kohsuke Tsuruta

This paper presents a crack characterization method by spatio-temporal gradient analysis over the Lamb wave field. The surface ultrasound waves that propagate in the direction of the layer are potential candidates of nondestructive testing (NDT) methodology. The proposed method has an ability to classify the surface acoustic wave field through the rank of the covariance matrix defined over the four-dimensional vector space which is spanned by the following components: a vertical displacement, its vertical velocity, and a pair of shear strains of the surface. The covariance matrix provides the information about cracks. The determinant of the covariance matrix, therefore, is proposed as the index of homogeneity of the object surface. In this study, the computational process in the

wave field near the cracks is discussed and their physical meanings are investigated through FDTD-simulations and acoustic experiments.

11:30

3aSP7. Simultaneous detection and parameter estimation of multiple nonlinear acoustic chirps.

Jasdeep S. Dhanoa, Evan J. Hughes, and Richard F. Ormondroyd (Cranfield Univ., Royal Military College of Sci., Shrivenham, Swindon SN6 8LA, UK)

Historically, short-time spectral analysis using the Fourier transform and its variants has been the primary method for time-frequency analysis. Recently, Wigner-based methods which provide time-frequency distributions have been applied to linear chirp detection. These methods are limited by the effects of windowing and, for the Wigner methods, cross-spectral components. Furthermore, they do not characterize the signal parameters directly but this may be accomplished by applying a Hough or Radon transform to the time-frequency distribution. For the detection of nonlinear chirps, these methods would be computationally intensive and have low accuracy. This paper introduces a new technique based on evolutionary algorithms for direct detection and parametrization of multiple nonlinear chirps within a window of observation. The optimization of the estimated chirp parameters is done in a nonlinear process using the evolutionary algorithm and the performance of the algorithm is not affected by peak broadening or the cross-spectral components. It exploits the windowing of the finite duration signal and provides higher accuracy even under conditions of high noise. The method also detects pure sinusoids and multiple linear chirps implicitly. The output parameters are, the start and stop frequencies, the phase, amplitude and the coefficients of the nonlinear variation of frequency with time.

11:45

3aSP8. Development of an underwater target classifier using target specific features.

M. H. Supriya and P. R. Saseendran Pillai (Dept. of Electron., Cochin Univ. of Sci. and Technol. Kochi, India)

In Sonar, the detection and estimation functions are performed by signal processors, which involve the computation of various statistics, for enhancing the overall performance of the system. This also takes into account all the undesirable propagation effects caused by the underwater channel. Underwater targets can be classified by using certain target specific features such as target strength, target dynamics, and the signatures of the noise generated by the targets. Rough identification of the targets is carried out with target strength values at known aspects while for precise identification, classification clues from target dynamics and target signatures are generated. Databases for the engine noise spectra of various underwater targets, propeller noises, machinery noises and cavitation noises, speed-noise characteristics, etc., have been developed. The signal energy estimated within a finite-time interval is compared with the earlier detection/estimation decisions, which are stored in the target data record and the relevant target data are updated. The algorithm for identification of target from the most matching signature patterns in the database will generate the classification clues, which will help in target identification. Salient highlights of an underwater target classifier using the above-discussed target specific features are presented in this paper.

3a WED. AM

Session 3aUW**Underwater Acoustics, Signal Processing in Acoustics and Engineering Acoustics: Robust Passive Sonar II**

Lisa M. Zurk, Cochair

Lincoln Laboratory, Massachusetts Institute of Technology, 244 Wood Street, Lexington, Massachusetts 02173-6426

Brian H. Tracey, Cochair

*Lincoln Laboratory, Massachusetts Institute of Technology, 244 Wood Street, Lexington, Massachusetts 02173-6426***Chair's Introduction—8:00*****Invited Papers*****8:05****3aUW1. From acoustic observatories to robust passive sonar.** Arthur B. Baggeroer (MIT, Cambridge, MA 02139)

The evolution of the DARPA Robust Passive Sonar (RPS) as well as the ONR Shallow Water Acoustic Testbed (SWAT) programs can be traced from concept of an acoustic observatory posed by Munk in 1980 through several assessment and feasibility studies to their current implementations. During this, the thinking on several key hypotheses matured. (i) Are noise fields directional enough to sustain high array gains? (ii) What are the tradeoffs among nonstationarity caused by ship motion, array configuration (geometry and the number of sensors), and “snapshots” needed for stable adaptive processing? (iii) What is the interaction between gains from vertical and horizontal apertures? (iv) How much signal gain degradation is acceptable? (v) What methods of post-processing can be done for normalization, tracking, and 3-D localization? This presentation will give a brief summary of the history of RPS and SWAT and pose the question of how well we can answer some of hypotheses which motivated them.

8:30**3aUW2. Some approaches to robust, snapshot-deficient, adaptive processing.** Heechun Song (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238)

We typically assume that longer arrays and more sensors are always beneficial. However, one of the issues identified by the ONR Ocean Acoustic Observatory Panel was “snapshot-deficient adaptive processing” arising, for example, from the motion of strong interferers (e.g., ships) through the small resolution cells of large arrays. This paper discusses two different but related concepts to reject the moving interferers: (1) null broadening for individual nulling and (2) model-based projection for collective nulling. The plane wave null broadening approach is combined with the robust white noise constraint and extended to matched field processing. The model-based projection exploits the fact that most interferers are close to the surface while targets of interest are usually in the middle of the water column. The projection matrix is constructed from the most significant eigenvectors of a replica covariance matrix from many independent sources within a surface strip. Simulations and experimental data illustrate the techniques.

8:55**3aUW3. Approaches to robustness.** Henry Cox and Kevin D. Heaney (Orincon Corp., 4350 N. Fairfax Dr., Ste. 470, Arlington, VA 22203)

The term robustness in signal processing applications usually refers to approaches that are not degraded significantly when the assumptions that were invoked in defining the processing algorithm are no longer valid. Highly tuned algorithms that fall apart in real-world conditions are useless. The classic example is super-directive arrays of closely spaced elements. The very narrow beams and high directivity could be predicted under ideal conditions, could not be achieved under realistic conditions of amplitude, phase and position errors. The robust design tries to take into account the real environment as part of the optimization problem. This problem led to the introduction of the white noise gain constraint and diagonal loading in adaptive beam forming. Multiple linear constraints have been introduced in pursuit of robustness. Sonar systems such as towed arrays operate in less than ideal conditions, making robustness a concern. A special problem in sonar systems is failed array elements. This leads to severe degradation in beam patterns and bearing response patterns. Another robustness issue arises in matched field processing that uses an acoustic propagation model in the beamforming. Knowledge of the environmental parameters is usually limited. This paper reviews the various approaches to achieving robustness in sonar systems.

9:20

3aUW4. Motion compensation for adaptive horizontal line array processing: A beam domain approach. T. C. Yang (Naval Res. Lab., Washington, DC 20375)

Large aperture horizontal line arrays have small resolution cells and can be used to separate a target signal from an interference signal by array beamforming. High-resolution adaptive array processing can be used to place a null at the interference signal. But these features are significantly degraded by the source motion, which reduces the time period under which the environment can be considered stationary. For adaptive array processing, a large number of data samples are generally required to minimize the variance of the cross-spectral density between the array elements. The penalty of integrating over a large number of samples, when the source and/or interference moves, is the spread of signal and interference energy to many eigenvalues and consequently, the ability to suppress the interference suffers. We adapt a beam domain approach to compensate for the source motion allowing the beam covariance matrix of the signal to be integrated over a large number of data samples without signal energy loss. We employ an equivalent of a rotating coordinate frame to track the signal bearing-change and use the waveguide invariant theory to compensate the signal range-change by frequency shifting. [Work supported by ONR.]

9:35

3aUW5. Interference suppression by tracking and integrating over source motion. T. C. Yang (Naval Res. Lab., Washington, DC 20375)

In this paper, we will show that the beam-domain motion-compensation algorithm can be used to suppress an interference signal when the interference source is at a different bearing/range, and/or has a different bearing/range rate than the signal source. The motion-compensation algorithm was originally developed to preserve the adaptive processing signal gain for a moving source. We have found that it also suppresses interference (beyond normal adaptive processing) under the above-stated conditions. The initial signal bearing is assumed known based on conventional processing but the range is assumed unknown. (The signal is either detected or its presence is highly suggested.) Signal bearing rate and range rate are assumed unknown and will be searched for by the motion compensation algorithm, which yields in principle a highest beam power at the true signal bearing and range rate. We employ an equivalent of a rotating coordinate frame to track the signal bearing-change and use the waveguide invariant theory to compensate the signal range-change by frequency shifting. The mismatch in the bearing and range rate between the interference and signal suppresses the interference power. [Work supported by ONR.]

9:50

3aUW6. Reduced beamset adaptive matched field processing. Brian Tracey, Srinivas Turaga, and Nigel Lee (MIT Lincoln Lab., 244 Wood St., Lexington, MA 02420, btracey@ll.mit.edu)

Matched field processing (MFP) offers the possibility of improved towed array performance at endfire through range/depth discrimination of contacts. One challenge is that arrays with limited vertical aperture can often resolve only a small number of multipath arrivals. This paper explores ways to capture the array resolution by re-parametrizing the set of MFP replicas. A reduced beamset can be created by performing a singular value decomposition on the MFP replica set. Alternatively, clustering techniques can be used to generate MFP cell families, or regions of similar response. These parametrizations are applied to adaptive MFP algorithms to show speed and performance gains. The use of cell families/regions instead of individual MFP cells also provides a framework for increasing the robustness of MFP by defocusing the MFP beamforming operation.

The techniques are demonstrated for shallow-water towed array scenarios. [Work sponsored by DARPA-ATO under Air Force Contract No. F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the Department of Defense. Approved for Public Release, Distribution Unlimited.]

10:05–10:20 Break

10:20

3aUW7. Adaptive sonar detection performance prediction in an uncertain ocean. Paul Book, Jeffery Krolik (Dept of Elec. & Computer Eng., Duke Univ., Durham, NC 27708-0291), and Shawn Kraut (Queen's Univ., Kingston, ON K7L 3N6, Canada)

This paper addresses the problem of predicting detection performance when the signal wavefront is uncertain and the noise field directionality is unknown. Passive sonar detection in this scenario typically involves robust adaptive beamforming with limited training data. The classical sonar equation, however, assumes the noise field and signal wavefront are known exactly. In this paper, we use the statistics of the generalized likelihood ratio test (GLRT) for the composite hypothesis of a multirank signal in Gaussian noise with unknown covariance matrix to evaluate the detection threshold (DT) as function of ocean uncertainty and number of noise training snapshots. Further, the trade-off between array gain (AG) and detection threshold (DT) is studied as a function of training sample size in a dynamic interference environment. Detection performance of the GLRT is characterized in terms of bounds on the middle 80th percentile of classical passive sonar figure of merit (FOM) and range-of-the-day (RD) over an ensemble of ocean environments. Different classes of environments including downward refracting and upward refracting scenarios are examined with particular attention to the Florida Straits region. Example performance prediction bounds are presented using real horizontal noise field and environmental data. [Work supported by ONR.]

10:35

3aUW8. Mode excision adaptive beamforming for source detection in an uncertain shallow-water waveguide. Vincent E. Premus (MIT Lincoln Lab., 244 Wood St., Lexington, MA 02420, vpremus@ll.mit.edu)

Passive sonar detection is uniquely characterized by the fact that the acoustic clutter distribution is generally confined to the ocean's surface. There is considerable evidence to support the hypothesis that surfaced and submerged sources are well separated in acoustic mode space, and that shallow-water waveguide normal modes are relatively robust to imperfect environmental knowledge. In this work, the use of mode physics is explored for the purpose of identifying an improved adaptive subspace for submerged source detection in the presence of surface interference. The basic premise is to perform adaptive weight computation in a mode subspace that is weakly excited by the submerged source of interest, yet well coupled to the surface interference. The rationale is to excise as much of the target signature as possible from the sample covariance without excessively compromising the measurement of the interference spatial spectrum. This enables more aggressive nulling of the surface clutter spectrum for a given level of signal gain degradation on the submerged source of interest. In this paper, the algorithm for adaptive mode subspace identification will be discussed, and the theoretical performance as a function of imprecise environmental knowledge and array calibration will be examined for a number of different apertures, including vertical line arrays, horizontal line arrays, and volumetric arrays. [Work sponsored in part by

DARPA, under Air Force Contract No. F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the U.S. Air Force.]

10:50

3aUW9. Noise model beamforming. Catherine H. Frazier, Iman W. Schurman, and Bruce K. Newhall (Appl. Phys. Lab., Johns Hopkins Univ., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099, catherine.frazier@jhuapl.edu)

Adaptive beamformers are usually based on the premise that noise is stationary and Gaussian, and hence completely characterized by a covariance matrix. The standard approach then attempts to use measured data to estimate that covariance and form a set of beamformer weights that optimally rejects the estimated noise. This standard method may not be robust to mismatch between real noise and the noise estimate. In particular, low frequency passive sonar noise is usually nonstationary, owing primarily to the motion of the shipping noise sources. We investigate alternative robust noise models that allow for nonstationarity. One such model [W. A. Kuperman and F. Ingenito, *J. Acoust. Soc. Am.* **67**, 1988–1996 (1980)] assumes a uniform probability distribution for surface noise sources. This produces an optimal beamformer that is robust in uniformly rejecting surface noise, no matter where sources occur on the ocean surface or how they move. The noise performance of this beamformer will be compared to standard adaptive approaches for array configurations proposed by ONR for the Acoustic Observatory Program. The high fidelity simulation model used for noise prediction includes nonstationary source motion and realistic range-dependent multipath propagation effects. [Work supported by DARPA.]

11:05

3aUW10. Nonexhaustive array processing. Peter Gerstoft, Katherine Kim, David Battle, W. A. Kuperman (Marine Physical Lab., Univ. of California San Diego, La Jolla, CA 92093-0238), William Hodgkiss, and Heechun Song (Univ. of California San Diego, La Jolla, CA 92093-0238)

Modern arrays may contain of order a thousand elements and the subsequent beamforming requires searching over frequencies, beam angle, source–receiver range, source depth, array depth, array tilt, and array bow. The search space, easily on the order of a trillion cells, suggests seeking alternative approaches to the traditional exhaustive methods. Here we pursue a nonexhaustive global search strategy combined with a local search for significant computational gain. The approach will be illustrated using computationally demanding simulated data and real towed-array data.

11:20

3aUW11. Matched-field processing with an L-shaped array. Gregory J. Orris, Michael Nicholas, and John S. Perkins (Naval Res. Lab., Washington, DC 20375)

Data was collected on a 64 element L-shaped Array deployed in two at-sea experiments: RDS-1 off the coast of Halifax, Nova Scotia during late September of 1997, and KWIX-98 off Key West, Florida in August of 1998. The L-shaped array was deployed in a configuration with 32 elements in the horizontal and 32 elements in the vertical. Environmental variability during both experiments proved to be extreme at times due to

weather and prevailing currents, making processing with so-called “high-resolution” and multi-frequency matched-field processing difficult. It was observed that fully coherent spatial processing using robust linear methods offered only marginal improvement over treating each leg of the array separately and combining the results in an incoherent processor. Possible explanations for these results are explored along with approaches aimed at improving the system performance. [Work supported by ONR.]

11:35

3aUW12. Quantification of deterministic matched-field source localization error in the face of random model inputs. Peter M. Daly and Gerald T. Hebenstreit (SAIC Ocean Systems Operation, 1710 SAIC Dr., M.S. 1-11-15, McLean, VA 22102, peter.m.daly@saic.com)

Deterministic source localization using matched-field processing (MFP) has yielded good results in propagation scenarios where the non-random model parameter input assumption is valid. In many shallow water environments, inputs to acoustic propagation models may be better represented using random distributions rather than fixed quantities. One can estimate the negative effect of random source inputs on deterministic MFP by (1) obtaining a realistic statistical representation of a signal model parameter, then (2) using the mean of the parameter as input to the MFP signal model (the so-called “replica vector”), (3) synthesizing a source signal using multiple realizations of the random parameter, and (4) estimating the source localization error by correlating the synthesized signal vector with the replica vector over a three dimensional space. This approach allows one to quantify deterministic localization error introduced by random model parameters, including sound velocity profile, hydrophone locations, and sediment thickness and speed. [Work supported by DARPA Advanced Technology Office.]

11:50

3aUW13. Blind sound channel deconvolution via artificial time reversal. Karim G. Sabra and David R. Dowling (Dept. of Mech. Eng., Univ. of Michigan, 2019 Lay Auto Lab 2121, Ann Arbor, MI 48109-2121)

Signals that travel through a sound channel are typically distorted when they are received by a remote listener because of interference arising from multiple propagation paths. This presentation introduces a novel passive acoustic technique for blind deconvolution of broadband signals recorded on a vertical array in an unknown multipath sound channel. The technique, called artificial time reversal or ATR, is based on artificially backpropagating the signals measured at the array to their source location using a broadband Green’s function synthesized directly from the measurements. The technique exploits generic features of modal propagation in ocean sound channels to infer a phase relationship between the Green’s function and the signal. The technique allows sound-channel spread signals to be compressed to their original length. Computational examples employing a single source are presented and compared with results obtained from oceanic measurements made in May 1997 off the west coast of Italy [Hodgkiss *et al.*, *J. Acoust. Soc. Am.* **105**, 1597 (1999)]. The temporal correlation between an ATR-compressed signal and the original signal approaches 100% in the computational examples and reaches 90% for the oceanic data. Potential extension of this technique to multiple sources emitting simultaneously in the same frequency band is also presented. [Work supported by ONR; oceanic data provided by Dr. Song of SIO.]

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 subcommittees 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.

Session 3pAO**Acoustical Oceanography: Acoustical Oceanography Prize Lecture**

Peter F. Worcester, Chair

*Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive,
La Jolla, California 92093-0225***Chair's Introduction—12:55*****Invited Paper*****1:00****3pAO1. The sound of rainfall at sea.** Jeffrey A Nystuen (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105, nystuen@apl.washington.edu)

Knowledge of the distribution of oceanic rainfall is needed to more fully describe the marine ecosystem and to provide constraints for climate investigations and for process studies of air-sea interaction. Bubbles generated by raindrop splashes produce a loud and distinctive sound underwater. This sound is a signal that allows the detection and measurement of rainfall at sea. Different raindrop sizes have unique splash characteristics, producing distinctive acoustic signatures for at least four raindrop sizes. These signatures form a mathematical basis for the decomposition of the sound field to measure the drop size distribution in the rain. This information allows the acoustic identification of rainfall type and rainfall rate. The acoustic detection and measurement of rainfall at sea is demonstrated using over 100 buoy-months of data from ocean surface moorings in the tropical Pacific Ocean. This technology is now being transferred to sub-surface instrument platforms. By acoustically monitoring rainfall, and other surface processes including wind speed, from below the surface, many of the fouling, physical damage and vandalism problems that affect surface instruments will be avoided. This technology is passive, introducing no acoustic disturbance into the environment and thus poses no potential harm to marine mammals or other forms of life in the ocean.

Session 3pID**Interdisciplinary: Hot Topics in Acoustics**

Robin O. Cleveland, Chair

*Aerospace and Mechanical Engineering, Boston University, 110 Cummings Street, Boston, Massachusetts 02215***Chair's Introduction—1:55*****Invited Papers*****2:00****3pID1. Hot topics in underwater acoustics.** John S. Perkins (Naval Res. Lab., Washington, DC 20375, perkins@ghost.nrl.navy.mil)

Underwater acoustics is a "big tent" discipline that draws from, and makes contributions to, many diverse fields such as wave propagation, physical oceanography, signal processing and ocean engineering. In recent years, concern for the environmental effects on marine life has spawned research into both the effects of man-made noise, and the use of active and passive acoustics to monitor biological activity. Special Sessions at this meeting reflect this diversity. This paper will draw on these Special Sessions to present some exciting examples of current research: (1) Parabolic equation methods for wave propagation, (2) robust passive sonar techniques, (3) inverse methods for determining geoacoustic parameters, (4) high-frequency acoustic interaction with the sea floor, and (5) bioacoustic resonance. In signal processing, one of the most active areas is the exploitation of the concept of time-reversal. In ocean engineering, there has been a tremendous increase in the applications for autonomous undersea vehicles. A brief overview of developments in these areas will be included. [Work supported by ONR.]

3pID2. Synergistic advances in diagnostic and therapeutic medical ultrasound. Frederic L. Lizzi (Riverside Res. Inst., 156 William St., New York, NY 10038)

Significant advances are more fully exploiting ultrasound's potential for noninvasive diagnosis and treatment. Therapeutic systems employ intense focused beams to thermally kill cancer cells in, e.g., prostate; to stop bleeding; and to treat specific diseases (e.g., glaucoma). Diagnostic ultrasound techniques can quantitatively image an increasingly broad spectrum of physical tissue attributes. An exciting aspect of this progress is the emerging synergy between these modalities. Advanced diagnostic techniques may contribute at several stages in therapy. For example, treatment planning for small ocular tumors uses 50-MHz, 3-D ultrasonic images with 0.05-mm resolution. Thermal simulations employ these images to evaluate desired and undesired effects using exposure strategies with specially designed treatment beams. Therapy beam positioning can use diagnostic elastography to sense tissue motion induced by radiation pressure from high-intensity treatment beams. Therapy monitoring can sense lesion formation using elastography motion sensing (to detect the increased stiffness in lesions); harmonic imaging (to sense altered nonlinear properties); and spectrum analysis images (depicting changes in the sizes, concentration, and configuration of sub-resolution structures.) Experience from these applications will greatly expand the knowledge of acoustic phenomena in living tissues and should lead to further advances in medical ultrasound.

3pID3. ASA education outreach. Uwe J. Hansen (Dept. of Phys., Indiana State Univ., Terre Haute, IN 47809) and E. Carr Everbach (Swarthmore College, Swarthmore, PA 19081-1397)

A number of very successful Hands-on demo sessions for high school students have been a part of regular ASA meetings for some time. In addition, the Education Committee has organized a series of teacher workshops. These workshops are designed to give high school teachers relatively sophisticated tools to enhance their laboratory content. Workshops for teachers in the elementary grades prepare teachers to use music as a vehicle to introduce additional science concepts. Content and methods associated with both workshops will be discussed. Cyberspace outreach by the ASA was accelerated by the establishment of a Home Page Committee, and more recently by the On-Line Education committee, which is creating an educational website. The website provides a fun way for users to access information including acoustics information, history, demos, and links to the Technical Committee's webpages. The ASA has joined other AIP member societies in developing additional mechanisms, including road shows and nightly news spots.

WEDNESDAY AFTERNOON, 30 APRIL 2003

ROOM 102, 1:00 TO 2:30 P.M.

Session 3pNS

Noise, Architectural Acoustics, Speech Communication, Psychological and Physiological Acoustics, Engineering Acoustics and ASA Committee on Standards: Panel Discussion, ASA's Role in Developing a Policy Statement and Outreach for the ANSI Classroom Acoustics Standard

Edward J. Walsh, Cochair

Boys Town National Research Hospital, 555 North 30th Street, Omaha, Nebraska 68131

Louis C. Sutherland, Cochair

27803 Longhill Drive, Rancho Palos Verdes, California 90275-3908

David Lubman, Cochair

14301 Middletown Lane, Westminster, California 92683

Chair's Introduction—1:00

Last year the Acoustical Society of America (ASA) Panel on Public Policy (POPP) was charged with the responsibility for drafting an ASA position statement for the Classroom Acoustics Standard (ANSI S12.60-2002). A draft policy for this Standard was prepared in January 2003 and was submitted to the Executive Council for approval. Both the Panel on Public Policy and the Executive Council expressed the need for a strong statement of affirmation for the Standard, which would allow the ASA to speak with a unified voice and technical accuracy on this matter of societal concern.

One of the principal goals associated with this first POPP initiative is to generate a policy statement that will allow the ASA to extend its outreach to other scientific, engineering and standardization bodies, as well as to the public at large. The content and impact of the policy document will be discussed, along with suggestions for approaches to increase outreach for the Classroom Acoustics Standard.

Session 3pPA

Physical Acoustics: Thermoacoustics

Matthew E. Poese, Chair

Graduate Program in Acoustics, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Contributed Papers

1:00

3pPA1. Cascade thermoacoustic engine. D. L. Gardner and G. W. Swift (Los Alamos Natl. Lab., Los Alamos, NM 87545)

Over the past several years, Los Alamos has successfully fabricated and tested traveling wave thermoacoustic heat engines having efficiencies much greater than those of traditional standing wave thermoacoustic engines. Unfortunately, Gedeon streaming effects present in these traveling wave engines has required using fluid flow "diodes" having greater complexity than desired. The cascade thermoacoustic engine recently fabricated and tested at Los Alamos uses three stages in series, one standing wave and two traveling wave, yielding a configuration in which Gedeon streaming is eliminated and efficiency can still be high. The cascade engine will be described and experimental data will be presented. [Work supported by the U.S. DOE Office of Science and Praxair, Inc.]

1:15

3pPA2. Sensorless control of a harmonically driven linear reciprocating electrodynamic machine for a thermoacoustic refrigerator. Tony L. Shearer (Northrup Grumman Electron. Systems, Sykesville, MD), Robert W. M. Smith (Penn State Appl. Res. Lab., State College, PA), Heath F. Hofmann (Penn State Dept. of Elec. Eng., University Park, PA), and Steven L. Garrett (Penn State Grad. Prog. in Acoust., University Park, PA)

In electrically driven thermoacoustic refrigerators, the acoustic pressure and actuator displacement are typically monitored while a nominally harmonic drive frequency is adjusted to achieve a given operating point (e.g., maximum electroacoustic conversion efficiency, acoustic resonance, maximum power, etc.) Most often, pressure and displacement are measured directly, using sensors. Using a model of the electrical and mechanical dynamics of a linear electrodynamic machine and resonant assembly, the actuator displacement and pressure at the piston can be accurately estimated via direct measurement of the voltage and current at the machine terminals. This provides a means for sensorless control of the machine. The estimation method and system parameter extraction method will be described, as well as results of comparisons to sensors. Results from tests on a thermoacoustic refrigerator operating at 3-kW acoustic input power using a sensorless control system implemented to maintain an acoustic resonance will be shown, wherein the refrigerator cold temperature was varied over a practical operating range. The control system tracked the acoustic resonance frequency and achieved the desired phase relationship between pressure and displacement to within 3 deg based on the estimation, thus maintaining essentially optimal electroacoustic conversion efficiency during the test. [Work sponsored by ONR.]

1:30

3pPA3. Practical considerations for the design of bellows seals in thermoacoustic refrigerators. Robert W. M. Smith (Penn State Appl. Res. Lab., P.O. Box 30, State College, PA 16804)

In many acoustic sources designed for thermoacoustic refrigerators, the means to provide coupling from a reciprocating electrodynamic machine to the thermoacoustic load has been a flexible metallic seal (bellows). Employing bellows provides for improved efficiency and less stringent assembly tolerances in comparison with clearance seal designs, at the

cost of a part that must endure high cycle fatigue. Often the dominant contribution to the stresses in the bellows is associated with deflection. Because the compressional wave speed in the bellows can be small (15 m/s) one must consider the effect of dynamic stress enhancement. For a given bellows geometry and desired end deflection, an optimum bellows length exists which minimizes the deflection stress. Interestingly, the selection of this optimum length also provides some intrinsic reduction in sensitivity of the computed stresses associated with commercial fabrication tolerances in the bellows. It will be shown that with a given bellows geometry, material design stress limit, and a design using the optimum length, results effectively in an end velocity limit on the bellows. In this way, design with conventional bellows can constrain the power available from some commercially available machines. [Work sponsored by ONR.]

1:45

3pPA4. Numerical modeling of inert gas-condensing vapor thermoacoustic engines. W. V. Slaton (Phys. Dept., Eindhoven Univ. of Technol., P.O. Box 513, 5600 MB Eindhoven, The Netherlands), Richard Raspet, and Robert A. Hiller (Univ. of Mississippi, University, MS 38677)

Recent theoretical work by Slaton and Raspet *et al.* describe the acoustic propagation equation [J. Acoust. Soc. Am. **114**, 1414–1422] and the second-order enthalpy and mass transport equations [J. Acoust. Soc. Am. **114**, 1423–1430] for an inert gas-condensing vapor mixture in a porous medium with an imposed temperature gradient. The acoustic propagation and enthalpy transport equations show that the vapor diffusion effects in the mixture are analogous to the heat diffusion effects in the thermoacoustics of inert gases, and that these effects occur in parallel with the heat diffusion effects in the wet system for proper choice of inert gas and vapor. Writing the acoustic propagation equation as two coupled first-order differential equations in terms of the volumetric velocity and acoustic pressure amplitude and utilizing the conservation of enthalpy in the stack allows the system of equations to be solved numerically by interfacing with the well-established thermoacoustic modeling code, DELTAE. Modeling of various thermoacoustic engines utilizing an inert gas-condensing vapor working fluid will be presented. It will be shown how the COP relative to Carnot and the heat pumping power for thermoacoustic refrigerators can be increased significantly by proper choice of gas mixture. [Work supported by ONR.]

2:00

3pPA5. Impedance tube measurements of thermoviscous functions with a condensable working fluid. D. C. Brown, R. A. Hiller, and R. Raspet (Dept. of Phys., Univ. of Mississippi, University, MS 38677)

Theoretical work by Slaton *et al.* [J. Acoust. Soc. Am. **112**, 1414–1422 (2002); **112**, 1423–1430 (2002)] develops the linear theory for thermoacoustics when the working fluid has a component which undergoes condensation and evaporation on the stack during the acoustic cycle. In order to test these ideas an impedance tube is used to measure the acoustic impedance in a closed-end driver-pipe system which contains a sample stack. The impedance is measured on the driver side of the stack from pressure measurements of two closely spaced microphones. This may be compared to the impedance computed from a numerical integration from

the closed end of the pipe through the stack, using the known or assumed values of the thermoviscous dissipation functions in the stack. Measurements and the theory for a variety of stack materials and working fluids will be presented.

2:15

3pPA6. An aeroacoustically driven thermoacoustic heat pump. W. V. Slaton and J. C. H. Zeegers (Dept. of Appl. Phys., Eindhoven Univ. of Technol., P.O. Box 513, 5600 MB Eindhoven, The Netherlands)

The mean flow of gas in a pipe past a side branch, closed at the far end, can excite the resonant acoustic modes of the cavity much like blowing across the top of a bottle. This aeroacoustic whistle can excite very high amplitude acoustic waves within the side branch (easily 10% of the

mean pressure) at optimal gas flow rates and mean pressures within the main pipe. The aeroacoustic whistle uses no moving parts to convert part of the power in the mean flow into acoustic power. Likewise a thermoacoustic heat pump, utilizing this acoustic power, uses no moving parts to pump heat and establish (or maintain) a temperature difference across a porous medium. This new combination of an aeroacoustic sound source and thermoacoustic heat pump (with suitable thermoelectric elements) is part of an electric power generation feasibility study for natural gas wells. Reliable electrical power generation down-hole to provide electricity for sensors, communications devices or energy storage units is an important research and development goal. Experimental results will be presented that demonstrate the performance of a simple thermoacoustic heat pump when powered by an aeroacoustic sound source. [Work supported by Shell International Exploration and Production B.V.]

WEDNESDAY AFTERNOON, 30 APRIL 2003

ROOM 204, 1:00 TO 3:00 P.M.

Session 3pPP

Psychological and Physiological Acoustics: Binaural Processing and Spatial Hearing

Pavel Zahorik, Chair

Waisman Center, Room 571, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, Wisconsin 53705

Contributed Papers

1:00

3pPP1. Investigating the precedence effect for noise bursts of different bandwidths. Jonas Braasch (Institut für Kommunikationsakustik, Ruhr-Universität Bochum, 44780 Bochum, Germany)

The fact that the precedence effect partly fails for signals with a very narrow bandwidth raises the question of if this phenomenon requires an across-frequency-band interaction. To investigate this, a psychoacoustic experiment was conducted, in which the perceived lateralization of a noise burst (ITD: 300 μ s) in the presence of a reflection (ITD: -300 μ s) was measured. The parameters in this experiment were the inter-stimulus interval (ISI: 0.0 ms-4.0 ms) and the bandwidth (100 Hz, 400 Hz and 800 Hz). The data of the listeners clearly show that the localization dominance becomes more stable, especially in regard to the dependence on the ISI, when the bandwidth is increased. At the narrowest tested bandwidth, localization dominance could not be observed for some of the listeners, while others perceived at least the sound coming from the side of the direct sound source, although the position of their auditory events varied strongly with the ISI. A signal analysis on the basis of a modified Lindemann algorithm reveals that, in the first case, the precedence effect does not seem to have any significant influence. The signal analysis also indicates that it is not necessary to assume an interaction across the involved frequency bands and that it is sufficient to simply average over the model outputs of those frequency bands.

1:15

3pPP2. Level effects in monaural and interaural intensity discrimination. Mark A. Stellmack, Neal F. Viemeister, and Andrew J. Byrne (Dept. of Psych., Univ. of Minnesota, 75 E. River Rd., Minneapolis, MN 55455)

Discrimination of interaural level differences (ILDs) is often assessed using a two-interval task in which ILD changes between intervals. In such a situation, overall level is usually roved in order to minimize monaural cues. This roving makes it difficult to assess level-dependent effects. In the present experiment, monaural intensity discrimination in a 2IFC task was compared to ILD discrimination in a single-interval task, thus producing

an analogous situation in terms of number of "observations" and eliminating the need for an overall level rove. Monaural intensity DLs and ILD discrimination thresholds were measured as a function of level for 4-kHz tones and for broadband noise. The Weber functions ($10 \log \Delta I/I$ vs I in dB) in the monaural and binaural conditions were parallel. For the noise the Weber functions had slopes close to zero (Weber's law); for the tones the slopes were -0.082 (near-miss to Weber's law). Overall, the binaural thresholds showed a small, approximately 2 dB, advantage over the monaural thresholds. The important aspect, however, is that the level effects seen monaurally are also seen binaurally. This suggests that the basic processes responsible for Weber's law and the near-miss occur prior to binaural interaction. [Work supported by NIDCD DC05343 and DC00683.]

1:30

3pPP3. Online estimation of room reverberation time. Rama Ratnam (Beckman Inst., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801), Douglas L. Jones, Bruce C. Wheeler, and Albert S. Feng (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

The reverberation time (RT) is an important parameter for characterizing the quality of an auditory space. Sounds in reverberant environments are subject to coloration. This affects speech intelligibility and sound localization. State-of-the-art signal processing algorithms for hearing aids are expected to have the ability to evaluate the characteristics of the listening environment and turn on an appropriate processing strategy accordingly. Thus, a method for the characterization of room RT based on passively received microphone signals represents an important enabling technology. Current RT estimators, such as Schroeder's method or regression, depend on a controlled sound source, and thus cannot produce an online, blind RT estimate. Here, we describe a method for estimating RT without prior knowledge of sound sources or room geometry. The diffusive tail of reverberation was modeled as an exponentially damped Gaussian white noise process. The time constant of the decay, which provided a measure of the RT, was estimated using a maximum-likelihood procedure. The estimates were obtained continuously, and an order-statistics filter was used to extract the most likely RT from the accumulated estimates. The procedure was illustrated for connected speech. Results obtained for simulated and real room data are in good agreement with the real RT values.

3p WED. PM

3pPP4. Reduced order modeling of head related transfer functions for virtual acoustic displays. Joel A. Willhite, Kenneth D. Frampton (Dept. of Mech. Eng., Vanderbilt Univ., Box 1592 Station B, Nashville, TN 37235, joel.a.willhite@vanderbilt.edu), and D. Wesley Grantham (Vanderbilt Univ. Medical Ctr., Nashville, TN 37212)

The purpose of this work is to improve the computational efficiency in acoustic virtual applications by creating and testing reduced order models of the head related transfer functions used in localizing sound sources. State space models of varying order were generated from zero-elevation Head Related Impulse Responses (HRIRs) using Kungs Single Value Decomposition (SVD) technique. The inputs to the models are the desired azimuths of the virtual sound sources (from minus 90 deg to plus 90 deg, in 10 deg increments) and the outputs are the left and right ear impulse responses. Trials were conducted in an anechoic chamber in which subjects were exposed to real sounds that were emitted by individual speakers across a numbered speaker array, phantom sources generated from the original HRIRs, and phantom sound sources generated with the different reduced order state space models. The error in the perceived direction of the phantom sources generated from the reduced order models was compared to errors in localization using the original HRIRs.

2:00

3pPP5. Effects of auralization technique and cross-modal visual stimulus on auditory localization in virtual environments. Paul D. Henderson, Rendell R. Torres, Yasushi Shimizu (Prog. in Architectural Acoust., Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, hendep2@rpi.edu), Richard Radke, and Brian Lonsway (Rensselaer Polytechnic Inst., Troy, NY 12180)

The performance of auditory localization in audiovisual environments is affected by both the complexity and congruence of the auditory and visual sensory inputs. Subjects were asked to identify the perceived location of a virtual acoustic source placed at varying positions in the horizontal plane and rendered by either conventional stereophonic reproduction or through wavefront reconstruction by a loudspeaker array. The virtual source was auralized in simulated reflective and anechoic acoustic environments both with no visual stimulus and with projected imagery of varying congruence. The localization precision of subjects presented with complex aural stimuli in the form of simulated reflective environments is reduced relative to anechoic conditions. Additionally, array-based wavefront reconstruction provides a significant increase in overall localization performance over stereophonic rendering, particularly for subjects positioned off the primary axis of the loudspeaker array. Subjects presented with simultaneous aural and visual inputs are able to accurately locate the virtual sound source when a large incongruence angle exists (typically more than 30 degrees). However, in the presence of a smaller angular separation between stimuli, a bias towards the visual stimulus is detectable. The results help determine the requirements for accuracy in auralization and reproduction in the creation of virtual multimedia environments.

2:15

3pPP6. The effect of stimulus bandwidth and subject position on horizontal-plane localization with virtual source images. D. Wesley Grantham, Daniel H. Ashmead, Robert S. Wall (Dept. of Hearing and Speech Sci., Vanderbilt Univ. Medical Ctr., Nashville, TN 37232), Kenneth D. Frampton, and J. Andrew Willhite (Vanderbilt Univ., Nashville, TN 37235)

In an anechoic chamber normal-hearing subjects performed a localization task in the frontal horizontal plane. The stimulus was a 200-ms burst of filtered noise. Within a block of trials, half of the presentations (randomly determined) were “real”—presented from single loudspeakers—and the other half were “phantoms”—produced by the simultaneous ac-

tivation of two loudspeakers at $\pm 30^\circ$ using a virtual source imaging technique [Takeuchi *et al.*, J. Acoust. Soc. Am. **109**, 958–971 (2001)]. Both phantom and real sources spanned the azimuthal range $\pm 80^\circ$. When the stimulus was a 4 kHz low-pass filtered noise, rms error was only slightly higher for phantom ($\bar{D}=7.1^\circ$) than for real ($\bar{D}=5.5^\circ$) sources. For 8 kHz low-pass filtered noise, performance remained about the same for real sources, but increased for phantom sources ($\bar{D}=11.5^\circ$). Data will also be reported for conditions in which the subject's position is systematically varied outside the “sweet spot.” Results will be discussed in terms of robustness of the virtual imaging technique to stimulus and position factors and its potential usefulness as a tool for the investigation of human auditory spatial perception in static and dynamic environments. [Work supported by NIDCD.]

2:30

3pPP7. The effect of background noise on the perception of auditory distance. Christopher A. Brown, William M. Whitmer, and William A. Yost (Parmlly Hearing Inst., Loyola Univ., 6430 N. Kenmore, Chicago, IL 60626)

Five-ms broadband noise bursts were played and recorded through a KEMAR manikin from 6 different distances in a reverberant room. Distances ranged from 4 to 9 feet, and all recordings were 500 ms in duration. A single loudspeaker sat atop a table, and was moved to each distance prior to recording. During testing, stimuli were presented over headphones to listeners seated in a sound-attenuating chamber. In one condition, the ambient sound of the room was mixed with the stimuli, and served to fill in the silent gaps between stimulus presentations. Previous results have shown that distance judgments improve significantly when the silent gaps are filled with ambient noise. In the other conditions on-line noise, which was filtered to approximate the spectrum of the ambient noise, was added instead. Three levels of attenuation were used, -5 dB, 0 dB, and $+5$ dB, relative to the background noise level. The task was to judge the distance of the sound. A repeated measures design was employed, and listeners ran in only one condition on a given day. Preliminary results suggest that judgments are significantly improved with ambient-room noise, as compared to on-line noise. Additional findings and significance will be discussed. [Work supported by NIH.]

2:45

3pPP8. Auditory and visual distance perception: The proximity-image effect revisited. Pavel Zahorik (Waisman Ctr., Univ. of Wisconsin–Madison, 1500 Highland Ave., Madison, WI 53705, zahorik@waisman.wisc.edu)

The proximity-image effect [M. B. Gardner, J. Acoust. Soc. Am. **43**, 163 (1968)] describes a phenomenon in which the apparent distance of an auditory target is determined by the distance of the nearest plausible visual target rather than by acoustic distance cues. Here this effect is examined using a single visual target (an un-energized loudspeaker) and invisible virtual sound sources. These sources were synthesized from binaural impulse-response measurements at distances ranging from 1 to 5 m (0.25-m steps) in the semi-reverberant room (7.7 m \times 4.2 m \times 2.7 m) in which the experiment was conducted. Listeners ($n=11$) were asked whether or not the auditory target appeared to be at the same distance as the visual target. Within a block of trials, the visual target was placed at a fixed distance of 1.5, 3, or 4.5 m, and the auditory target varied randomly from trial-to-trial over the sample of measurement distances. The resulting psychometric functions are consistent with the proximity-image effect, and can be predicted using a simple model of sensory integration and decision in which perceived auditory space is both compressed in distance and has lower resolution than perceived visual space. [Work supported by NIH-NEI.]

Session 3pSP

Signal Processing in Acoustics: Statistical and Random Analysis

Leon H. Sibul, Chair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, University Park, Pennsylvania 16804

Contributed Papers

1:30

3pSP1. Statistical estimators of a periodically correlated random process for a voiced speech signal. Lesya B. Chorna (School of Audiol. & Speech-Lang. Pathol., Univ. of Memphis, 807 Jefferson Ave., Memphis, TN 38105, lchorna@memphis.edu)

A stochastic model in the form of a periodically correlated random process (PCRP) was applied to a voiced speech signal. In such a signal, the amplitudes and phases of the harmonic components varied randomly; therefore, a stochastic model of the signal was appropriate. A description of the signal in terms of a stationary random process allowed an analysis of the spectral density of the oscillations, but the phase (temporal) structure remained unconsidered. To reveal this structure, a PCRP was used. The speech signal was processed by the co-phase analysis in the time domain and the component analysis in the spectral domain. The algorithms were based on a specific property of the PCRP that the samples selected with a period of correlation formed a series of related stationary sequences. Each sequence was then analyzed using the theory of stationary processes. This technique was used in a study of 70 recordings of the vowel /a/. The recordings were obtained from people with a normal heart rate and provoked arrhythmia. Statistical estimators of the vocal signal calculated on the basis of the PCRP model were statistically different for the normal and arrhythmic heart rates.

1:45

3pSP2. Fisher information for a Gaussian random variable with unequal real and imaginary covariances: Derivation and application to beamforming. Sandra L. Collier (U.S. Army Res. Lab., AMSRL-CI-EE, 2800 Powder Mill Rd., Adelphi, MD 20783-1197, scollier@arl.army.mil) and D. Keith Wilson (Eng. Res. and Development Ctr., U.S. Army Cold Regions Res. and Eng. Lab., Hanover, NH 03755-1290)

There are many physical situations in which a received signal may be modeled as a complex Gaussian random variable. One such situation occurs when an acoustic wave is strongly scattered as a result of propagation through a random medium such as the atmosphere or ocean. However, there are also many conditions under which this model fails to provide an adequate representation. A specific example is the geometric acoustics regime, in which diffraction and scattering by the medium are both weak and the variance of the phase fluctuations is much larger than the variance of the log-amplitude fluctuations. The reduced wavefunction is the wavefunction of a signal propagating in an inhomogeneous medium normalized by the wavefunction in free space. A statistical model is developed for a reduced wavefunction whose real and imaginary components are Gaussian random variables with unequal covariances. A linear transformation is performed and the probability density function of the received signal at a passive array is calculated. The Fisher information is calculated for special conditions of the transformation that are of interest to acoustic beamforming. The coupling of the Cramer-Rao lower bounds on the parameter estimates (e.g., the angles of arrival, source phase, and medium parameters) is addressed.

2:00

3pSP3. Gibbs sampling for multipath arrival time estimation in an uncertain environment. Michele Picarelli and Zoi-Heleni Michalopoulou (Dept. of Mathematical Sci., New Jersey Institute of Technol., Newark, NJ 07102)

Estimation of time delays and amplitudes of multipath arrivals is of great interest in many fields because of the information that can be extracted from these characteristics. In underwater acoustics, multipath features of sound signals can reveal the location of the sound-generating source and properties of the propagating medium; their accurate estimation is, thus, desirable. In this work, a scheme is developed for the estimation of multipath arrival times and amplitudes in a noisy environment considering an unknown signal-to-noise ratio. The number of arrivals at the receiver is considered unknown as well, since, realistically, it depends on properties that would be uncertain *a priori*. The method is based on the efficient calculation of the number of paths, their amplitudes, and arrival times using Gibbs sampling to calculate *a posteriori* probability distributions. Results from the new method are compared to those of conventional approaches and show the potential of the new technique for efficient time delay and amplitude estimation in an environment with several unknowns. [Work supported by ONR.]

2:15

3pSP4. Audio fine classification using the statistical analysis of acoustic images. Edward Chilton and Ioannis Paraskevas (Univ. of Surrey, EE+IT, CVSSP Group, GU2 7XH, UK, E.Chilton@surrey.ac.uk)

The fine classification of audio utterances is an important problem because the features that have to be extracted need to be very accurate in order to contribute to effective classification. In this paper, results are presented for a fine classification problem: namely the classification of two groups of different kinds of gunshots. The problem of accurate classification can be divided into two parts: (i) feature extraction and (ii) classification. The more effective the feature extraction, the more effectively the classifier will be able to categorize the various audio samples. In this paper, a novel method for the automatic recognition of acoustic utterances is presented using acoustic images as the basis for the feature extraction. The feature extraction process is based on the time-frequency distribution of an acoustic unit. A novel feature extraction technique based on the statistical analysis of the spectrogram Hartley transform (distribution) and Choi-Williams distributions of the data is reported as well as a brief discussion of the classifier used. The image is compressed using a statistical analysis of the acoustic image formed from the time-frequency distributions of acoustic data. The kurtosis, L-moments and entropy of the distributions, as well as the energy, contrast, etc., of the corresponding co-occurrence matrices of the distributions are calculated and then combined into a feature matrix. These appropriate features are then presented to the classifier. Initial results obtained indicate that the method is capable of accurate discrimination for fine classification.

3p WED. PM

3pSP5. Localizing a large-dimensional field of sonobuoys. Nicole E. Collison (Defence R&D Canada–Atlantic, 9 Grove St., P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada) and Stan E. Dosso (SEOS, Univ. of Victoria, Victoria, BC V8W 3P6, Canada)

For target localization, multistatic sonar systems require an adequate knowledge of both the source and receiver positions. In this paper, we use a regularized acoustic inversion method on measured direct-arrival times from several impulsive sources to track a freely drifting sonobuoy field. The shallow-water experiment involved 11 sonobuoys within a 6×8 km

field, with 6 sources over approximately 70 min. Regularization allows prior information to be built into the inversion, which in this case consists of estimates (with associated uncertainties) of the source and initial sonobuoy drop positions determined from the GPS position of the aircraft at the instant of drop, as well as a model for smooth sonobuoy tracks. Closely spaced sonobuoys move along similar tracks, although there is considerable movement in different directions over the entire field (260–700 m). Positioning uncertainties are estimated using a Monte Carlo appraisal procedure to be approximately 100 m (absolute) and 65 m (relative). Submitted for the Signal Processing Young Presenter Award.

WEDNESDAY AFTERNOON, 30 APRIL 2003

RENAISSANCE EAST BALLROOM, 3:15 TO 5:15 P.M.

Plenary Session, Business Meeting and Awards Ceremony

Richard Stern, President
Acoustical Society of America

Business Meeting

Presentation of Certificates to New Fellows

Richard H. Campbell	Zoi-Heleni Michalopoulou
Laurel H. Carney	William C. Moss
Bruce D. Cornuelle	Philip A. Nelson
Amy M. Donahue	Marshall H. Orr
Hans P. Gottlieb	Jack E. Randorff
George E. Ioup	Jens H. Rindel
Jules S. Jaffe	Robert S. Schlauch
Leon M. Keer	Scott D. Sommerfeldt
Robert M. Keolian	Mario A. Svirsky
Oswald Leroy	Floyd E. Toole
Raymond Lim	David P. Walsh
Hugh J. McDermott	Beverly A. Wright
Colette M. McKay	Sean F. Wu
Ronald L. McKay	Ning Xiang

Announcement of Prize and Grant

Jeffrey A. Nystuen, 2003 Medwin Prize in Acoustical Oceanography

Lori L. Holt, American Speech, Language, Hearing Foundation's Research Grant in Speech Science

Presentation of Awards

Silver Medal in Psychological and Physiological Acoustics to Brian C. J. Moore

Helmholtz–Rayleigh Interdisciplinary Silver Medal to Arthur B. Baggeroer

R. Bruce Lindsay Award to Dani Byrd

Gold Medal to Richard H. Lyon

Session 4aAA**Architectural Acoustics and Musical Acoustics: Recording Studio Acoustics**

Jeff D. Szymanski, Cochair

Auralex Acoustics, 8851 Hague Road, Indianapolis, Indiana 46256

Alexander U. Case, Cochair

*Fermata Audio and Acoustics, P.O. Box 1161, Portsmouth, New Hampshire 03802***Invited Papers****8:30****4aAA1. A review of the pertinent measurements and equations for small room acoustics.** Douglas Jones (Dept. of Audio Arts & Acoust., Columbia College Chicago, 600 S. Michigan, Chicago, IL 60605)

Confusion about the appropriate measurements and calculations for use in small room acoustics persists. This paper presents an overview of the most important calculations to predict performance during the design phase and the most useful measurements to verify or troubleshoot performance after the facility is constructed. Special attention will be paid to the importance and relevance of reverberation in small rooms.

8:50**4aAA2. A consideration on physical tuning for acoustical coloration in recording studio.** Yasushi Shimizu (Prog. in Architectural Acoust., School of Architecture, Rensselaer Polytechnic Inst., Troy, NY 12180-3590, shimiy@rpi.edu)

Coloration due to particular architectural shapes and dimension or less surface absorption has been mentioned as an acoustical defect in recording studio. Generally interference among early reflected sounds arriving within 10 ms in delay after the direct sound produces coloration by comb filter effect over mid- and high-frequency sounds. In addition, less absorbed room resonance modes also have been well known as a major component for coloration in low-frequency sounds. Small size in dimension with recording studio, however, creates difficulty in characterization associated with wave acoustics behavior, that make acoustical optimization more difficult than that of concert hall acoustics. There still remains difficulty in evaluating amount of coloration as well as predicting its acoustical characteristics in acoustical modeling and in other words acoustical tuning technique during construction is regarded as important to optimize acoustics appropriately to the function of recording studio. This paper presents an example of coloration by comb filtering effect and less damped room modes in typical post-processing recording studio. And acoustical design and measurement technique will be presented for adjusting timbre due to coloration based on psycho-acoustical performance with binaural hearing and room resonance control with line array resonator adjusted to the particular room modes considered.

9:10**4aAA3. If you can't take the room out of your mix, you can't take your mix out of the room!** Peter D'Antonio (RPG Diffusor Systems, Inc., 651-C Commerce Dr., Upper Marlboro, MD 20774, pdantonio@rpginc.com)

The key issue in any recording studio is transferability—the ability of a mix to transfer to other listening environments outside the studio. For a mix to faithfully transfer to a wide range of acoustical environments, it must be created in a room with minimal acoustic distortion. The music industry is very aware of electronic distortion; however, the audible effects of acoustic distortion are only now being fully appreciated. The four forms of acoustic distortion are modal coupling, speaker boundary interference response, comb filtering and poor diffusion or a sparse spatial and temporal reflection density. These phenomena will be explained and methods to minimize them will be suggested.

9:30**4aAA4. NFL Films audio, video, and film production facilities.** Russ Berger, Richard C. Schrag, and Jason J. Ridings (Russ Berger Design Group, 4006 Belt Line Rd., Ste. 160, Addison, TX 75001, russ@rbdg.com)

The new NFL Films 200,000 sq. ft. headquarters is home for the critically acclaimed film production that preserves the NFL's visual legacy week-to-week during the football season, and is also the technical plant that processes and archives football footage from the earliest recorded media to the current network broadcasts. No other company in the country shoots more film than NFL Films, and the inclusion of cutting-edge video and audio formats demands that their technical spaces continually integrate the latest in the ever-changing world of technology. This facility houses a staggering array of acoustically sensitive spaces where music and sound are equal partners with the visual medium. Over 90,000 sq. ft. of sound critical technical space is comprised of an array of sound stages, music scoring stages, audio control rooms, music writing rooms, recording studios, mixing theaters, video production control rooms, editing suites, and a screening theater. Every production control space in the building is designed to monitor and produce multi channel surround sound audio. An overview of the architectural and acoustical design challenges encountered for each sophisticated listening, recording, viewing, editing, and sound critical environment will be discussed.

10:10

4aAA5. NFL Films music scoring stage and control room space. Russ Berger, Richard C. Schrag, and Jason J. Ridings (Russ Berger Design Group, 4006 Belt Line Rd., Ste. 160, Addison, TX 75001, russ@rbdg.com)

NFL Films' new 200,000 sq. ft. corporate headquarters is home to an orchestral scoring stage used to record custom music scores to support and enhance their video productions. Part of the 90,000 sq. ft. of sound critical technical space, the music scoring stage and its associated control room are at the heart of the audio facilities. Driving the design were the owner's mandate for natural light, wood textures, and an acoustical environment that would support small rhythm sections, soloists, and a full orchestra. Being an industry leader in cutting-edge video and audio formats, the NFLF required that the technical spaces allow the latest in technology to be continually integrated into the infrastructure. Never was it more important for a project to hold true to the adage of "designing from the inside out." Each audio and video space within the facility had to stand on its own with regard to user functionality, acoustical accuracy, sound isolation, noise control, and monitor presentation. A detailed look at the architectural and acoustical design challenges encountered and the solutions developed for the performance studio and the associated control room space will be discussed.

10:30

4aAA6. Home studio acoustic treatments on a budget. Gavin A. Haverstick (Auralex Acoust., Inc., 8851 Hague Rd., Indianapolis, IN 46256, gavin@auralex.com)

Digital technology in the recording industry has evolved and expanded, allowing it to be widely available to the public at a significantly lower cost than in previous years. Due to this fact, numerous home studios are either being built inside or converted from bedrooms, dens, and basements. Hobbyists and part-time musicians that typically do not have the advantage of a large recording budget operate the majority of these home studios. Along with digital equipment, acoustic treatment has become more affordable over the years giving many musicians the ability to write, record, and produce an entire album in the comfort of their own home without having to sacrifice acoustical quality along the way. Three separate case studies were conducted on rooms with various sizes, applications, and budgets. Acoustical treatment such as absorption, diffusion, and bass trapping were implemented to reduce the effects of issues such as flutter echo, excessive reverberation, and bass build-up among others. Reactions and subjective comments from each individual studio owner were gathered and assessed to determine how effective home studios can be on a personal and professional level if accurately treated acoustically.

10:50

4aAA7. Studio with a view. Anthony K. Hoover (Cavanaugh Tocci Assoc., 327F Boston Post Rd., Sudbury, MA 01776)

Berklee College of Music (in Boston) needed a new studio in which to teach stereo mixing and critical listening. A small synthesis lab (adjacent to the main lobby, directly over the cafeteria kitchen, penetrated by exhaust ducts, and next to a bathroom) was chosen for renovation. The primary requirements were for maximum visibility to assure hopeful future engineers a full view of all the cool gear, and comfortable seating for fifteen students. The challenges, to be discussed, included isolation with a view, quiet HVAC, and great sound, in a space that was acoustically too small and in the wrong place. The best verification of success is its popularity, which has prevented the author from booking time for listening or testing.

11:10

4aAA8. Design of a critical listening classroom and studio. Bob Alach, Lou Clark (Alacronics, 192 Worcester St., Wellesley, MA 02481, bob@alacronics.com), William Carman (Univ. of Massachusetts—Lowell, Lowell, MA 01854), and Alex Case (Fermata Audio + Acoust., Portsmouth, NH 03802)

The Sound Recording Technology program at the University of Massachusetts—Lowell required design of a space that could serve three functions: a critical listening classroom for 30 students, a live room for the department's premier 24 track recording studio, and a 5.1 surround sound mix room. Accomplishing all of the above required a re-evaluation of room shape and treatments and aggressive use of variable acoustics. In addition, all the usual culprits of HVAC noise and vibration, limited space, a constrained budget, and a difficult to renovate reinforced concrete building added to the challenge. This paper reviews the key design and renovation challenges associated with completing this room, opened for use by students and faculty in the Spring 2003 semester.

Contributed Paper

11:30

4aAA9. Acoustics of a broadcast center. Sergio Beristain (E.S.I.M.E., IPN, IMA, P.O. Box 12-1022, Narvarte, 03001, Mexico D. F., Mexico, sberista@hotmail.com)

A broadcast system in Mexico City had to change facilities in order to concentrate in a single site all related broadcast stations and production studios in order to facilitate its normal operation. This led to a design which included the acoustic noise isolation and the interior acoustics of

every studio and control room, together with the audio interconnection, the electricity layout, the air conditioning system, the office building, etc. This paper presents the acoustics profile of the center, including final results of the construction as they were measured on completion of the installation. The complex has seven AM and FM broadcast stations, plus seven production studios for news, commercials and radio-novels plus an audio master control room, and everything was completed within four months.

11:45–12:00

Panel Discussion

Session 4aAB

Animal Bioacoustics: Nature's Orchestra—Acoustics of Singing and Calling Animals I

Whitlow W. L. Au, Chair

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

Chair's Introduction—8:00

Invited Papers

8:05

4aAB1. The chorus environment and the shape of communication systems in frogs. Vince Marshall (Div. of Biological Sci., Univ. of Missouri—Columbia, 213 Tucker Hall, Columbia, MO 65211, vtml1da@mizzou.edu)

Many species of frogs breed in dense and structurally complex aggregations of calling males termed choruses. Females entering a chorus are faced with the tasks of recognizing and locating mates on the basis of their advertisement calls. The chorus environment poses particular challenges for communication as signalers and receivers will face high levels of background noise and interference between signals. For females, such conditions may decrease the efficiency of communication, with the consequences of increasing the time required to find a mate or errors in mate choice. For males, it will give rise to intense competition for the attention of females. Additionally, the chorus environment for a species is not static, and will vary over both spatial and temporal scales. This complex and dynamic environment has shaped the signals and signaling behaviors of frogs in sometimes surprising ways. In this talk, some of the implications of the chorus environment for both receivers and signalers is discussed. In particular, examples from North American hylid frogs are drawn upon and research on the role of signal timing in influencing the responses of females and plasticity in aggressive behavior between neighbors in choruses are discussed.

8:25

4aAB2. Patterns of fish sound production. David A. Mann (Univ. of South Florida, College of Marine Sci., 140 7th Ave. S., St. Petersburg, FL 33701, dmann@marine.usf.edu)

While vocalization and chorusing behavior has been intensively studied in frogs and birds, little has been done with fishes. This paper presents patterns of sound production in damselfish, toadfish, and spotted seatrout on seasonal, daily, and subdaily time scales. Chorus behavior ranges from highly coordinated behavior between neighboring toadfish to uncoordinated behavior in spotted seatrout. Differences in coordination of sound production (i.e., the degree of overlap in calls) can be related to differences in territoriality and modes of reproduction. Toadfish and damselfish are territorial fishes in which males guard benthic eggs laid in nests. Sciaenids (croakers and drums) spawn planktonic eggs, and form temporary aggregations of calling males.

8:45

4aAB3. Acoustic communication in *Panthera tigris*: A study of tiger vocalization and auditory receptivity. Edward J. Walsh (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131), Lily M. Wang (Univ. of Nebraska, Omaha, NE 68182), Douglas L. Armstrong, Thomas Curro, Lee G. Simmons (Henry Doorly Zoo, Omaha, NE 68108), and JoAnn McGee (Boys Town Natl. Res. Hospital, Omaha, NE 68131)

Acoustic communication represents a primary mode of interaction within the sub-species of *Panthera tigris* and it is commonly known that their vocal repertoire consists of a relatively wide range of utterances that include roars, growls, grunts, hisses and chuffing, vocalizations that are in some cases produced with extraordinary power. *P. tigris* vocalizations are known to contain significant amounts of acoustic energy over a wide spectral range, with peak output occurring in a low frequency bandwidth in the case of roars. von Muggenthaler (2000) has also shown that roars and other vocal productions uttered by *P. tigris* contain energy in the infrasonic range. While it is reasonable to assume that low and infrasonic acoustic cues are used as communication signals among conspecifics in the wild, it is clearly necessary to demonstrate that members of the *P. tigris* sub-species are responsive to low and infrasonic acoustic signals. The auditory brainstem response has proven to be an effective tool in the characterization of auditory performance among tigers and the results of an ongoing study of both the acoustical properties of *P. tigris* vocalizations and their auditory receptivity support the supposition that tigers are not only responsive to low frequency stimulation, but exquisitely so.

9:05

4aAB4. Humpback whale song: A new review. Adam S. Frankel (Marine Acoust., Inc., 706 Giddings Ave., Ste. 2A, Annapolis, MD 20764)

The humpback whale song has been described and investigated since the early 1970s. Much has been learned about the humpback whale social structure, but the understanding of the song and its function remains elusive. The hierarchical nature of the song structure was described early on: Songs can be sung for a long period, apparently by males, and primarily during the mating season. However, singers also become physically competitive, suggesting alternative mating strategies. There are a number of unique structural features of song. Its structure evolves over time and combination. The nature of song evolution strongly implies cultural transmission. Song

structure appears to be shared within an entire population, even though there appears to be little interchange of individuals between sub populations. Despite over thirty years of inquiry there are still numerous unanswered questions: Why is the song structure so complex? Is song a sexual advertisement, an acoustic space mediation mechanism, or both? How do females choose mates, or do they? What drives song evolution, and why is there so much variation in the rate of change? Are there nonreproductive functions of song? What prompts a male to begin or end singing? Our current understanding and the outstanding questions yet to be answered will be reviewed.

9:25

4aAB5. Signals and speciation in plant-feeding insects. Reginald Cocroft, Rafael Rodriguez (Biological Sci., Univ. of Missouri, Columbia, MO 65211), and Randy Hunt (Indiana Univ.-SE, New Albany, IN 47150)

Communication can play a role in speciation when differences in signaling systems reduce gene flow between diverging lineages. However, the observation that closely related species have different signals leaves open the question of cause and effect, because divergence in signals can also occur after speciation is completed. The relationship of signals and divergence is being investigated in the *Enchenopa binotata* species complex of treehoppers, a group of plant-feeding insects that communicate using substrate-borne vibrational signals. Speciation in the *E. binotata* complex is hypothesized to have resulted from repeated shifts to novel host plants. Sources of selection on the signals of these treehoppers can be studied not only in fully differentiated species, but also in experimental populations in the early stages of divergence. Some unique features of vibrational communication through plant stems will be discussed, as well as how shifts to novel host plants may influence signaling systems, and what role(s) communication may have played in the process of evolutionary divergence in this group.

9:45

4aAB6. Listening to Nature's orchestra with peculiar ears. David D. Yager (Dept. of Psych., Univ. of Maryland, College Park, MD 20742)

Insects use hearing for the crucial tasks of communicating with conspecifics and avoiding predators. Although all are based on the same acoustic principles, the diversity of insect ears is staggering and instructive. For instance, a South African grasshopper demonstrates that hearing conspecific calls is possible over distances 1 km with ears that do not have tympana. Actually, these creatures have six pairs of ears that play different roles in behavior. In numerical contrast, praying mantises have just a single ear in the ventral midline. The ear is very effective at detecting ultrasonic bat cries. However, the bioacoustics of sound transduction by two tympana facing each other in a deep, narrow slit is a puzzle. Tachinid flies demonstrate that directional hearing at 5 kHz is possible with a pair of ears fused together to give a total size of 1 mm. The ears are under the fly's chin. Hawk moths have their ears built into their mouthparts and the tympanum is more like a hollow ball than the usual membrane. As an apt last example, cicada ears are actually part of the orchestra: their tympana function both in sound reception and sound production.

10:05–10:20 Break

10:20

4aAB7. Perception of species-specific vocalizations by birds. Robert Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, dooling@psyc.umd.edu)

The complexity of bird vocalizations has fascinated us throughout the ages and there has long been the suspicion that birds are capable of producing, perceiving, and learning features of their songs that are beyond the capabilities of human hearing. Other investigators point to similarities between song production and perception in birds and special perceptual and attentional processes that are involved in the perception of speech by humans. Evidence will be reviewed that birds are born with early perceptual predispositions for learning species-specific vocalizations and demonstrate unusual efficiency in perceiving species-specific vocalizations throughout adulthood. In spite of these similarities with humans, birds have other auditory perceptual biases that are clearly different from humans as in the perception of tonal patterns. Birds, like humans, show a number of phenomena that aid perception under adverse conditions such as the binaural release from masking and the precedence effect and the detection of vocalizations in noise appears to follow from general principles of masking which has implications for studies of the effect of noise on birds in nature. [Work supported by NIH.]

Contributed Papers

10:40

4aAB8. The vocal monotony of monogamy. Jeanette Thomas (Lab. of Sensory Biol., Western Illinois Univ. Quad Cities, 3561 60th St., Moline, IL 61265)

There are four phocids in waters around Antarctica: Weddell, leopard, crabeater, and Ross seals. These four species provide a unique opportunity to examine underwater vocal behavior in species sharing the same ecosystem. Some species live in pack ice, others in factice, but all are restricted to the Antarctic or sub-Antarctic islands. All breed and produce vocalizations under water. Social systems range from polygyny in large breeding colonies, to serial monogamy, to solitary species. The type of mating system influences the number of underwater vocalizations in the repertoire,

with monogamous seals producing only a single call, polygynous species producing up to 35 calls, and solitary species an intermediate number of about 10 calls. Breeding occurs during the austral spring and each species carves-out an acoustic niche for communicating, with species using different frequency ranges, temporal patterns, and amplitude changes to convey their species-specific calls and presumably reduce acoustic competition. Some species exhibit geographic variations in their vocalizations around the continent, which may reflect discrete breeding populations. Some seals become silent during a vulnerable time of predation by killer whales, perhaps to avoid detection. Overall, vocalizations of these seals exhibit adaptive characteristics that reflect the co-evolution among species in the same ecosystem.

4aAB9. Songlike vocalizations from a Sumatran rhinoceros calf (*Dicerorhinus sumatrensis*). Elizabeth von Muggenthaler (Fauna Commun.s Res. Inst., P.O. Box 1126, Hillsborough, NC 27278) and Paul Reinhart (Cincinnati Zoo, Cincinnati, OH)

Within the last ten years the Sumatran rhino population has dropped 50%, and only 200–300 individuals exist, with five in captivity. Their native habitat is dense tropical forest and they are solitary, therefore much of their behavior remains unknown. Sumatrans are the smallest living rhino, standing 0.9–1.5 m tall, and are covered in coarse, reddish-brown hair. The first Sumatran rhinoceros born in captivity in 112 years, and the first calf ever recorded, is 17 months old and weighs 448 kg. At the Cincinnati Zoo this male calf was recorded from 1–3 m, using two Statham radio microphones, and one TCD-D8 Sony DAT recorder (9 Hz–22 kHz). Analysis, including power spectrums, spectrographic functions, and cross correlations were performed using National Instrument's Polynesia. Preliminary analysis indicates that the calf's vocalizations are similar in structure to adult Sumatran vocalizations, although there are some distinctions. "Eeps" and "whales" that are found in adult repertoires are produced by the calf. However, signals from the calf are higher in frequency, and the calf does not vocalize as consistently as the adults. The calf has yet to produce a "whistle blow," which is an adult vocalization that has a strong infrasonic component.

4aAB10. Comparison of St. Lawrence blue whale vocalizations with field observations. 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., St., Lambert, QC J4P 1T3, Canada)

During four field seasons from 1998–2001, vocalizations were recorded in the presence of St. Lawrence blue whales using a single omnidirectional hydrophone. Both long duration infrasonic calls (~18 Hz, 5–20 s) as well as short duration higher frequency calls (85–25 Hz, ~2 s)

were detected and compared with field observations. Two trends were noted. First, the long infrasonic call series were concentrated primarily in the deep (300 m) channel. These call series appear to compare well with blue whale vocalizations recorded by others in the deep open ocean. Second, the shorter audible calls were more evenly distributed over bathymetry and seem to be a form of short distance communication with at least one case occurring during an agonistic interaction. A comparison of these calls with biological parameters such as density of whales in the area, percentages of paired versus single whales, and numbers of males versus females will also be discussed. [Project supported by ARL/PSU, NSF, and the American Museum of Natural History.]

4aAB11. The acoustic field of singing humpback whales in the vertical plane. Whitlow W. L. Au (Marine Mammal Res. Prog., Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), Adam A. Pack (Kewalo Basin Marine Mammal Lab., Honolulu, HI 96814), Marc O. Lammers (Hawaii Inst. of Marine Biol.), Louis Herman (Kewalo Basin Marine Mammal Lab.), Kimberly Andrews (Hawaii Inst. of Marine Biol.), and Mark Deakos (Kewalo Basin Marine Mammal Lab.)

A vertical array of five hydrophones was used to measure the acoustic field of singing humpback whales. Once a singer was located, two swimmers with snorkel gear were deployed to determine the orientation of the whale and to position the boat so that the array could be deployed in front of the whale at a minimum standoff distance of 10 m. The spacing of the hydrophones was 7 m with the deepest hydrophone deployed at depth of 35 m. An 8-channel TASCAM recorder having a bandwidth of 24 kHz was used to record the hydrophone signals. The location of the singer was determined by computing the time of arrival differences between the hydrophone signals. The maximum source level varied between individual units in a song, with values between 180 and 190 dB. The acoustic field determined by considering the relative intensity of higher frequency harmonics in the signals indicate that the sounds are projected in the horizontal direction with the singer's head canted downward 45 to 60°. High-frequency harmonics extended beyond 24 kHz, suggesting that humpback whales may have an upper frequency limit of hearing as high as 24 kHz.

THURSDAY MORNING, 1 MAY 2003

ROOMS 110/111, 9:00 A.M. TO 12:00 NOON

Session 4aAO

Acoustical Oceanography: General Topics in Acoustical Oceanography

Mohsen Badiy, Cochair

College of Marine Studies, University of Delaware, Robinson Hall, Newark, Delaware 19716

Kathleen E. Wage, Cochair

Department of ECE, George Mason University, 4400 University Drive, Fairfax, Virginia 22030

Contributed Papers

4aAO1. Measuring ambient ocean bubble fields using a multibeam sonar. Steve Adelman, David L. Bradley, R. Lee Culver, and Thomas C. Weber (Appl. Res. Lab., Penn State Univ., P.O. Box 30, N. Atherton St., State College, PA 16804-0030, sga121@psu.edu)

For two weeks in July 2002 off the coast of San Diego, a 240 kHz SeaBat 8101 Multibeam Echosounder System was used to measure the backscatter from ambient bubble fields in the near surface layer. The sonar system was mounted approximately 12.5 meters below the surface on the

hydraulic orientation unit of the research platform FLIP and was oriented so that its fan of 101 1.5 degree beams were looking at the surface at an angle of 45 degrees from the vertical. Data collection was part of a joint experiment involving the Applied Research Laboratory (ARL)/The Pennsylvania State University (PSU) and the Marine Physical Laboratory (MPL)/Scripps Institute of Oceanography (SIO). Although sea conditions were benign for most of the test period, wind wave activity near the end provided a number of opportunities to observe near surface bubble entrainment. The sonar system proved to be an effective tool for the observation of the lifetime, spatial structure and dimensions of the ambient bubble fields. [Work supported by ONR under Award No. N00014-02-1-0156.]

4aAO2. Some analytic solutions for the dispersion and attenuation of sound in bubbly seawater. Ralph R. Goodman and Jerald W. Caruthers (The Univ. of Southern Mississippi, Stennis Space Center, MS 39529)

The effect that air bubbles have on sound propagation is well known, and in the past three decades, due to advances in acoustic measurement technologies, has been the subject of several studies that investigate the bubble sizes and distributions due to breaking waves and ship wakes. The distribution of bubbles in size and total void fraction, and in time and space, plays a critical role in the determination of dispersion, refraction and attenuation. Bubble size distributions have been measured under different environments and by several techniques and have been represented usually with a power law or an exponential dependence. The authors have investigated a few distributions that are amenable to exact analytical solutions. These are power law distributions that have been chosen to give insight to realistic cases without significantly modifying the physics. The solutions to the integrals for dispersion and attenuation will be presented. They will be used to show the changes in the magnitudes of sound speed and attenuation under varying conditions.

9:30

4aAO3. On the underwater sound generated by the counterpropagating sea-surface wave. Konstantin A. Naugolnykh (Univ. of Colorado/Zeltech, 325 Broadway, Boulder, CO 80305) and Samuil A. Rybak (N. Andreev Acoust. Inst., Moscow 117036, Russia)

Two counterpropagating gravity-capillary waves of the same frequency, forming a standing wave on the sea surface, give rise to a sound wave traveling away from the surface [M. S. Longuet-Higgins (1950); L. M. Brekhovskikh (1966)]. The effect can be described in terms of three-wave nonlinear interaction [V. V. Goncharov, K. A. Naugolnykh, and S. A. Rybak (1977)]. The frequency and the wave number of the radiated wave are determined by the synchronism conditions. The surface waves propagating in the opposite directions can be generated by the action of the wind on the remaining swell of a cyclone or due to nonlinear interaction of the gravity-capillary waves. The progressive gravity-capillary waves superimposed on the flow produced by the finite-amplitude dominant gravity wave can be blocked at the points where its group velocity balance the convection by the larger-scale flow generating the reflected wave. As a result in the field of progressive waves produced by wind the counter-propagating wave appears generating the sound. [Work partially supported by Collaborative Linkage Grant No. 977 890.]

9:45

4aAO4. Rain parameter estimation using impact generated low-frequency acoustic signals. T. K. Mani, P. R. Saseendran Pillai, James Kurian, and Supriya M. Hariharan (Cochin Univ. of Sci. and Technol., Cochin 682022, India)

The role of rain generated acoustic signals in the frequency range 500 Hz–100 kHz to estimate the rain parameters is well known [Medwin *et al.*, J. Acoust. Soc. Am. **92**, 1613–1623 (1992)]. A study of utilizing the low-frequency component below 500 Hz, present in the rain generated acoustic noise, for estimation of rain parameters is performed in this paper. Raindrops falling from various heights are simulated using a drop generator. Gravity waves are produced in the water surface due to the drop impact and this causes a sinusoidal low-frequency damped pressure wave in water. This low-frequency signal is captured using a sensor assembly and analyzed. It is observed that this component is fairly constant in its frequency while amplitude is found to vary in accordance with the drop size as well as the velocity. Drop-size and its kinetic energy are determined and compared with those obtained from theoretical computations and by direct measurements. The results show that this new technique of

analysis of rain generated acoustic signals yields rain parameters with good degree of accuracy. This method has the advantage of savings in computation time and simplicity of design.

10:00–10:15 Break

10:15

4aAO5. Sound fluctuations due to horizontal refraction in the SWARM-95 experiment. Mohsen Badiy (College of Marine Studies, Univ. of Delaware, Newark, DE 19716), Boris Katsnelson, Serguey Pereselkov (Voronezh State Univ., Voronezh 394006, Russia), James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and William Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

Fluctuations of the sound field measured in the SWARM-95 experiment are interpreted as manifestations of three-dimensional effects in shallow-water sound propagation. It is shown that features of the received sound signal, along an acoustic track placed at 5° to the crests of propagating internal soliton trains, can be explained by significant horizontal refraction of the sound waves. This refraction leads to a remarkable redistribution of the energy in the horizontal plane as has been recently remarked by other investigators. These features are specifically: (1) a high level of fluctuations of the sound energy (up to 6 dB) which are synchronous in time and have same amplitude over depth of waveguide and (2) anomalies in the arrival times and amplitudes of separate waveguide modes. A theoretical analysis based on the theory of vertical modes and horizontal rays is presented, along with computer modeling of the experiment. Analytical estimations of the observed effects as well as results from numerical modeling are in a good agreement with experimental observations and are consistent with simulations. [Work supported by ONR and US CRDF Award No. VZ-010-00.]

10:30

4aAO6. Frequency-dependent anomalies of sound propagation in the SWARM-95 experiment. Boris Katsnelson, Serguey Pereselkov (Voronezh State Univ., Universitetskaya Sq. 1, Voronezh 394006, Russia), Mohsen Badiy (College of Marine Studies, Univ. of Delaware, Newark, DE 19716), and James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Frequency-dependent acoustic propagation is a subject of great interest in shallow-water regions. The phenomenon known as resonance absorption occurs when broadband sound signals propagate across the crests of internal solitary waves (ISWs) if the sound field satisfies the condition of space synchronization. In this paper we show that different anomalies in frequency dependence can take place when sound signals propagate along the wave fronts of ISWs. These anomalies result from the anisotropic properties of the ISWs that can produce frequency-dependent horizontal refraction of acoustic waves in the ocean. Sound propagation of broadband pulses in the SWARM-95 experiment is analyzed from this point of view. The nature of such behavior is connected with the selective (over modes and frequencies) character of thermocline perturbations caused by ISWs. [Work supported by ONR and US CRDF Award No. VZ-010-00.]

10:45

4aAO7. Internal wave observations and transmission losses in the Santa Barbara Channel. Jeff Dunne (Widener Univ., One Univ. Pl., Chester, PA 19013, dunne@pop1.science.widener.edu), Orest Diachok, and Stephen Wales (Naval Res. Lab., Washington, DC 20375)

Measurements from three vertical line arrays of temperature probes, deployed in a triangular configuration in the Santa Barbara Channel in August 2002, permitted determination of speed and direction of tidally

generated internal solitary wave events (ISWEs). Two arrays were deployed 2 km apart along a 60-m isobath, with the third normal to the isobath. Sound propagation measurements were also conducted along the isobath. Highly correlated ISWEs were evident on all three arrays every 12 h. Measurements of time delays suggest that these waves emanated from a single site between two of the Channel Islands. Because the ISWE propagation direction was nearly normal to the acoustic propagation direction, acoustic mode coupling due to ISWEs is expected to be minimal. Time-series temperature measurements reveal that average near-surface sound speeds were systematically different when ISWEs were present or absent, and changed abruptly during transitions. Numerical simulations with a normal mode propagation code (C-SNAP) indicate that frequency-dependent differences in transmission loss on near-bottom hydrophones may be as large as 7 dB in narrow frequency bands (5%–10%) between 0.5 and 2 kHz when an ISWE is present versus when it is absent. [Work supported by ONR.]

11:00

4aAO8. Modeling the acoustic receptions at the NPAL array from the Kauai source. Michael D. Vera, Matthew A. Dzieciuch (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0225), and the NPAL Group^{a)} (SIO-UCSD, APL-UW, WHOI)

Acoustic transmissions from a 75-Hz source near Kauai to a vertical line array near California were recorded as part of the North Pacific Acoustic Laboratory (NPAL) experiment. Extensive environmental measurements were also performed as part of the experiment and were intended to ensure correspondence between numerical simulations and the data. Despite the availability of this information, the process of identifying the recorded arrivals with predictions has not been a simple one. Since the source is near the seafloor at about 800 m depth, and the depth at the receiver is approximately 1800 m, acoustic interaction with the bathymetry has been explored as a possible complication. Ray simulations that allow for specular reflection from the bottom indicate that fully-refracted and bottom-interacting paths can reach the receiver range (about 3900 km) at similar travel times. The simultaneous presence of both kinds of acoustic energy could contribute to the identification difficulties. A series of parabolic-equation simulations have been performed for different geoaoustic parameters in an attempt to correspond more closely to the data. The sensitivity of the predictions to the method used to extract and interpolate the sound speeds has also been investigated. [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.

11:15

4aAO9. Whispering-gallery-mode trapping of sound in shallow water between an up-slope region and internal wave solitons. Allan D. Pierce and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Double-trapping of sound occurs in shallow water when the ocean has a sloping bottom and when a packet of high amplitude internal wave solitons is propagating either up-slope or down-slope. The phase velocity of an adiabatic mode increases in the up-slope direction and also as one approaches the front of the soliton packet, so the mode's horizontal rays can be trapped for propagation that is nominally cross-slope. Because the variation of the horizontal phase velocity is considerably slower in the up-slope direction than it is in the vicinity of the soliton front, an analogy exists with the whispering-gallery effect (found near concave surfaces in

architectural spaces). The equivalent radius of curvature of the "reflecting surface" is found in accord with the earth-flattening approximation to equal the phase velocity divided by its derivative with respect to the up-slope distance. Quantitative substantiation is given for models of sound and solitons in a realistic ocean (two constant sound speed layers separated by a thermocline). Double-trapped propagation has some attenuation because of energy leakage out through the width of the soliton packet.

11:30

4aAO10. Effects of internal solitary waves on the invariance of acoustic intensity striation patterns. Altan Turgut, Marshall Orr (Naval Res. Lab., Acoust. Div., Washington, DC 20375), Daniel Rouseff (Univ. of Washington, Seattle, WA 98105), James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Ching Sang Chiu (Naval Postgrad. School, Monterey, CA 93943)

Effects of Internal Solitary Waves (ISWs) on the acoustic intensity striation patterns are studied using a unique data set collected in the South China Sea during the ONR Asian Seas International Acoustics Experiment (ASIAEx). The data set contains measured broadband (270–330 Hz) acoustic fields on a 32-element, ~400 m aperture horizontal array describing two regimes where the ISW packet being present and not present in a 19 km long propagation track. The geometry of the ISW packet was captured by a Synthetic Aperture Radar (SAR) image when it entered the propagation track at a 50 deg angle. It has been observed that ISWs introduce larger intensity fluctuations as well as frequency-shifts in the striation patterns when they are in the propagation track. However, the slope of the striations seems to be less affected by the presence of ISWs. Results from several broadband simulations are also presented to describe both acoustic intensity fluctuations and variations in the striation patterns in terms of acoustic mode coupling and mode refraction induced by ISWs. [Work supported by ONR.]

11:45

4aAO11. Comparison of broadband mode arrivals at ranges of 3515 km and 5171 km in the North Pacific. Kathleen E. Wage (George Mason Univ., MS1G5, 4400 University Dr., Fairfax, VA 22030) and The ATOC Group^{a)}

The Acoustic Thermometry of Ocean Climate (ATOC) provided an opportunity to observe signals propagating in the low-order modes of the ocean waveguide. Understanding the fluctuations of these mode signals is an important prerequisite to using them for tomography or other applications. In previous work, we characterized the cross-mode coherence and temporal variability of the low-order mode arrivals at 3515 km range [Wage *et al.*, J. Acoust. Soc. Am. (in press)]. This study compares the mode arrivals for two different ranges : 3515 km and 5171 km, using data from the ATOC vertical line arrays at Hawaii and Kiribati. We discuss the mode intensity and coherence statistics for each of the arrays and examine mean arrival time trends over the year-long deployment. Experimental results are compared to PE simulations of propagation through a realistic background environment perturbed by internal waves of varying strengths. The dependence of mode statistics on the path-dependent changes in the background sound speed and the parameters of the internal wave field is explored. [Work supported by an ONR Ocean Acoustics Young Faculty Award.] ^{a)} A. B. Baggeroer, T. G. Birdsall, C. Clark, J. A. Colosi, B. D. Cornuelle, D. Costa, B. D. Dushaw, M. A. Dzieciuch, A. M. G. Forbes, B. M. Howe, D. Menemenlis, J. A. Mercer, K. Metzger, W. H. Munk, R. C. Spindel, P. F. Worcester, and C. Wunsch.

Session 4aBB**Biomedical Ultrasound/Bioresponse to Vibration, Signal Processing in Acoustics and Physical Acoustics:
High-Intensity Focused Ultrasound (HIFU) and Imaged-Guided Therapy I**

Ronald A. Roy, Chair

*Department of Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215***Invited Papers****8:20**

4aBB1. Acoustic hemostasis. L. Crum, M. Andrew, M. Bailey, K. Beach (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Seattle, WA 98105), A. Brayman, F. Curra, P. Kaczkowski, S. Kargl, R. Martin, and S. Vaezy (Univ. of Washington, Seattle, WA 98105)

Over the past several years, the Center for Industrial and Medical Ultrasound (CIMU) at the Applied Physics Laboratory in the University of Washington has undertaken a broad research program in the general area of High Intensity Focused Ultrasound (HIFU). Our principal emphasis has been on the use of HIFU to induce hemostasis; in particular, CIMU has sought to develop a small, lightweight, portable device that would use ultrasound for both imaging and therapy. Such a technology is needed because nearly 50% of combat casualty mortality results from exsanguinations, or uncontrolled bleeding. A similar percentage occurs for civilian death due to trauma. In this general review, a presentation of the general problem will be given, as well as our recent approaches to the development of an image-guided, transcuteaneous, acoustic hemostasis device. [Work supported in part by the USAMRMC, ONR and the NIH.]

8:45

4aBB2. HIFU treatment of liver cancer—Successes and failures. Gail ter Haar, Ian Rivens (Phys. Dept., Inst. of Cancer Res., Royal Marsden Hospital, Sutton, Surrey SM2 5PT, UK), James Kennedy (HAIFU Unit, Churchill Hospital, Oxford OX3, UK), and Feng Wu (Chongqing Univ. of Medical Sci., Chongqing 400016, PROC)

Clinical trials of the HIFU treatment of liver cancer have been underway in the UK at the Royal Marsden Hospital since December 1997, and at the Churchill Hospital since November 2002. Royal Marsden treatments are undertaken using a prototype device known as the Teleson, while those at the Churchill are performed using the machine produced by the Chongqing HAIFU company in China. Both sites have demonstrated the ability to ablate significant volumes of a tumor within the liver. Despite differences in ultrasound exposure delivery, these treatments have highlighted some of the problems associated with the clinical use of extracorporeal HIFU. These problems lie primarily in the areas of targeting, treatment optimization and monitoring of ablation. These problems will be discussed and potential solutions suggested. [Work funded by the UK Department of Health, the Institute of Cancer Research and UTL.]

9:10

4aBB3. Large ultrasound phased arrays for noninvasive focal treatments. Kullervo Hynynen, Gregory Clement, Nathan McDannold, Chris Connor, and Natalia Vykhodtseva (Dept. of Radiol./MRI, Brigham and Women's Hospital, 75 Francis St., Boston, MA 02115)

Ultrasound applicators have become more advanced with the use of large scale phased arrays. These arrays can control the focal spot location and shape and compensate for the ultrasound beam distortions introduced by tissue layers. It has been demonstrated that sharp focus can be formed even through the human skull. However, the exposure control becomes complex with these new arrays, and online monitoring is required to provide clinical safety. For the past few years, magnetic resonance imaging has been tested for imaging the temperature distributions induced by focused ultrasound treatments. Both phantom and animal experiments have shown that the proton resonant frequency is temperature dependent and that maps of the temperature change can be obtained within seconds with phase subtraction methods. Recent clinical treatments have verified the usefulness of large scale phased arrays and temperature mapping during ultrasound surgery.

9:35

4aBB4. Controlled ultrasonic tissue erosion. Charles Cain (Dept. of Biomed. Eng., Univ. of Michigan, Ann Arbor, MI 48109)

Controlled ultrasonic tissue erosion has many clinical applications, including the placement of very precise sharply defined perforations in biological interfaces and membranes with focused ultrasound. With carefully chosen acoustic parameters, tissue can be rapidly eroded away at a constant etching rate. The maximum erosion rate for minimal propagated energy is obtained by using very short high intensity pulses. The etching rate is higher with shorter pulse durations. For short pulses less than 10 cycles of the drive

frequency, an optimum pulse repetition rate exists which maximizes the etching rate. Higher gas saturation in the surrounding medium reduces the etching rate and reduces the spatial sharpness of the holes produced. Most of the erosion appears to be produced in the first several cycles of the therapy pulse. For example, a series of short (about 3 cycles) focused pulses of 2100 W/cm^2 (Isppa) at 788 kHz can erode a very well defined 2 mm diameter hole in a 1 mm thick sample of fresh pork atrial posterior wall in about 1 min at the optimum pulse repetition rate (about 18 kHz). Controlled ultrasonic tissue erosion may provide an effective image guided noninvasive tool in treatment of neonatal patients with hypoplastic left heart syndrome. Without the mixing of oxygenated blood across perforations placed in the atrial septum, these infants do not survive.

10:00–10:15 Break

Contributed Papers

10:15

4aBB5. High intensity focused ultrasound (HIFU) for treatment of T1/T2 prostate cancer. N. Sanghvi (Focus Surgery, Inc., Indianapolis, IN 46226), T. Gardner, and M. Koch (Indiana Univ. School of Medicine, Indianapolis, IN 46226)

This FDA approved phase I/II clinical trial is to evaluate the safety and efficacy of the Sonablate device (Focus Surgery, Inc.) for the treatment of organ confined prostate cancer. 20 patients with biopsy proven prostate cancer, Gleason ≤ 7 and PSA ≤ 10 were treated under general anesthesia. Outcome data included serum PSA collected at day 3, 14, 30, 90, 180, PSA nadir (mean/median), and biopsy results at 6 months. Quality of life was assessed using the International Prostate Symptom Score, International Impotence and Erectile Function score, and the SF-36 health survey. The mean patient age is 62.0, Gleason score of 6.18, PSA of 5.2, and prostate size 26.0 g m. Mean PSA results were 5.62, 44, 20, 1.68, 0.87, and 0.44 ng/ml at screening, 48–72 hours, 14 days, 30 days, 90 days and 180 days, respectively. There was one patient (9%) with a positive TRUS biopsy at 6 months, which resulted in a retreatment. There were no rectal injuries. Average pre-treatment IPSS, IIEF, and SF-36 scores were 9.55, 16.1, and 103.5. At the 30 day follow-up, they were 18.3, 3, and 97.4, respectively. HIFU is a minimally invasive modality that achieves complete prostatic ablation and is efficacious in the treatment of low-stage prostate cancer.

10:30

4aBB6. High intensity focused ultrasound (HIFU) treatment of BPH: results of a multi-center phase III study. N. Sanghvi (Focus Surgery, Inc., Indianapolis, IN 46226), T. Gardner, M. Koch, R. Bihle, R. Foster (Indiana Univ. School of Medicine, Indianapolis, IN 46226), M. Resnick, A. Seftel (Case-Western Reserve Univ., Cleveland, OH), I. Grunberger (Long Island College Hospital, Brooklyn, NY), C. Stiedle (North East Indiana Res. Hospital, Fort Wayne, IN), and J. Corchan (Presbyterian Hospital, Dallas, TX)

The five centers phase III trial was to show that HIFU can treat prostate tissue thermally for symptomatic relief of BPH and improve flow rates. At five sites, 68 BPH patients were treated with the Sonablate device (Focus Surgery, Inc. Indianapolis, IN). A urethral Foley catheter was inserted into the urethra to aid in positioning and was kept *in-situ* during the treatment. A cooling device was used to cool the rectal wall. The patients returned home within a few hours after the procedure. The Foley catheter was kept electively to avoid any incidence of acute urinary retention following the therapy. The catheter was removed after 4–5 days. The average treatment time was 38 minutes. The patients were treated without pain, blood loss or complications. At 90 days post treatment, average Qmax and AUA Symptom Scores improved from 8.7 ml/s to 12.66 ml/s (48%) and 23.06 to 11.62 (52%), respectively. Significant prostate tissue changes took place before and after the treatment. 80% of the patients had cavity formation at the site of treatment at the bladder neck and prostate. Non-surgical HIFU therapy is safe and effective for providing symptomatic relief of BPH symptoms and the treatment can be performed as an outpatient procedure.

THURSDAY MORNING, 1 MAY 2003

ROOMS 108/109, 10:00 A.M. TO 1:00 P.M.

Session 4aED

Education in Acoustics: Hands-on Demonstrations for High-School Students

Uwe J. Hansen, Cochair

Department of Physics, Indiana State University, Terre Haute, Indiana 47809

Corinne M. Darvennes, Cochair

Tennessee Technological University, Box 5014, Cookeville, Tennessee 38505

Approximately 75 female students from rural high schools are invited to participate in this session. They will have the opportunity to sense the excitement of “Doing Acoustics” by participating in about 20 acoustics demonstrations. ASA members are free to play with the toys as well, as long as they do not interfere with student participation.

4a THU. AM

Session 4aPAa

Physical Acoustics: Nonlinear Acoustics and Resonators

Ray S. Wakeland, Chair

*Graduate Program in Acoustics, Pennsylvania State University, Applied Science Building, Room 217,
University Park, Pennsylvania 16804*

Contributed Papers

8:00

4aPAa1. Parametric excitation of an acoustic mode. Bruce Denardo, Derek B. Smith, Larry P. Varnadore (Dept. of Phys., Naval Postgraduate School, Monterey, CA 93943), and Wayne E. Prather (Miltec Res. and Technol., Inc., Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 39677)

One means of driving an oscillatory mode of a system is to modulate a parameter upon which the natural frequency of the mode depends. This "parametric excitation" occurs most readily if the modulation frequency is twice the frequency of the mode. In addition, the modulation amplitude must exceed a threshold that depends upon the dissipation of the mode. The response amplitude then grows exponentially until it is saturated by a nonlinearity of the system, so large amplitudes are possible. Previous efforts to parametrically excite an acoustic mode will be briefly described, and the theory for a length-modulated tube will be discussed. An analysis of this case indicates that it is feasible to parametrically excite the fundamental mode. Results of an experiment that is currently in progress will be presented. The tube has ambient length 2.0 m and diameter 30 cm, and the length is modulated by low-frequency loudspeakers at the ends. Sulfur hexafluoride gas is employed in order to decrease the required drive frequency and thus increase the maximum drive amplitude. A smooth constriction in the middle of the tube serves to detune the second mode, so that this mode is not directly excited.

8:15

4aPAa2. Effect of forcing function on nonlinear acoustic standing waves. Joshua R. Finkbeiner (Illinois Inst. of Technol., Grad. College, 3300 S. Federal St., Chicago, IL 60616, finkjos@iit.edu), Xiaofan Li, Ganesh Raman (Illinois Inst. of Technol., Chicago, IL 60616), Christopher C. Daniels (Ohio Aerosp. Inst., Cleveland, OH 44142), and Bruce M. Steinetz (NASA Glenn Res. Ctr., Cleveland, OH 44135)

Nonlinear acoustic standing waves of high amplitude have been demonstrated by utilizing the effects of resonator shape to prevent the pressure waves from entering saturation. Experimentally, nonlinear acoustic standing waves have been generated by shaking an entire resonating cavity. While this promotes more efficient energy transfer than a piston-driven resonator, it also introduces complicated structural dynamics into the system. Experiments have shown that these dynamics result in resonator forcing functions comprised of a sum of several Fourier modes. However, previous numerical studies of the acoustics generated within the resonator assumed simple sinusoidal waves as the driving force. Using a previously developed numerical code, this paper demonstrates the effects of using a forcing function constructed with a series of harmonic sinusoidal waves on resonating cavities. From these results, a method will be demonstrated which allows the direct numerical analysis of experimentally generated nonlinear acoustic waves in resonators driven by harmonic forcing functions.

8:30

4aPAa3. Determination of dimensionless attenuation coefficient in shaped resonators. Christopher C. Daniels (OAI, 22800 Cedar Point Rd., Cleveland, OH 44142), Joshua R. Finkbeiner (Illinois Inst. of Technol., Grad. College, 10 W. 32nd St., Chicago, IL 60616), Bruce M. Steinetz (NASA Glenn Res. Ctr., Cleveland, OH 44135), Xiaofan Li, and Ganesh Raman (Illinois Inst. of Technol., Chicago, IL 60616)

The value of the dimensionless attenuation coefficient is an important factor when numerically predicting high-amplitude acoustic waves in shaped resonators. Both the magnitude of the pressure waveform and the quality factor rely heavily on this dimensionless parameter. Previous authors have stated the values used, but have not completely explained their methods of determination. This work fully describes the methodology used to establish this important parameter. Over a range of frequencies encompassing the fundamental resonance, the pressure waves were experimentally measured at each end of the shaped resonators. At the corresponding dimensionless acceleration, the numerical code modeled the acoustic waveforms generated in the resonator using various dimensionless attenuation coefficients. The dimensionless attenuation coefficient that most closely matched the pressure amplitudes and quality factors of the experimental and numerical results was determined and will be used in subsequent studies.

8:45

4aPAa4. Numerical simulation of streaming in high amplitude standing wave resonators. Said Boluriaan and Philip Morris (Dept. of Aerosp. Eng., Penn State Univ., 233 Hammond Bldg., University Park, PA 16802)

Classical theories of acoustic streaming in a long standing wave resonator predict a streaming velocity with a parabolic profile across the resonator. Recently, an analysis by Menguy and Gilbert [Acustica **86**, 249–259 (2000)] has shown that the parabolic streaming velocity profile is obtained only for low acoustic pressure amplitudes. At a higher amplitude, the streaming velocity deviates from a parabolic distribution. In a more recent experiment, Thompson and Atchley [J. Acoust. Soc. Am. **111**, 2418 (2002)] have also observed streaming velocities that are different from those predicted by the classical theory when a high amplitude standing wave is used. Furthermore, there are indications that the type of the boundary condition (adiabatic or isothermal, for instance) may affect the mean velocity. In the present research, numerical simulations are performed for a high amplitude standing wave in a resonator. A time-accurate, high-order, finite difference approach on a clustered rectangular grid structure is used and the full Navier–Stokes equations are solved. Results are presented for a wide range of operating parameters and are compared with both the Menguy and Gilbert analysis and the experiments. Moreover, the time evolution of the streaming velocity is investigated.

9:00

4aPAa5. Acoustic streaming at high amplitudes in plane and cylindrical channels. Mark F. Hamilton, Yuri A. Ilinskii, and Evgenia A. Zabolotskaya (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063)

At low amplitudes, acoustic streaming generated by standing waves in a plane channel is similar to streaming in a cylindrical tube. However, numerical simulations based on fully nonlinear equations for the streaming flow reveal significant differences at high amplitudes, at which Reynolds numbers for the streaming exceed unity. In these calculations, the acoustic field driving the streaming is prescribed independently. With increasing acoustic amplitude, the centers of the streaming vortices shift, the streamlines lose their symmetry, and the transverse velocity distribution departs from a parabolic profile. The vortex centers shift in opposite directions in the cylinder and the plane channel. In the cylinder the vortex centers move toward antinodes of the acoustic velocity field as the amplitude is increased, whereas in the plane channel the centers move toward the nodes. Simultaneously, the transverse streaming velocity profiles are observed to become narrower in some regions and broader in others. In general, these nonlinear streaming effects are more pronounced for the plane channel and therefore must be taken into account for Reynolds numbers smaller than for the cylinder. The influence of harmonic generation in the acoustic field will also be discussed. [Work supported by ONR.]

9:15

4aPAa6. Systematic acoustic loading of an aeroacoustic whistle. W. V. Slaton and J. C. H. Zeegers (Dept. of Appl. Phys., Eindhoven Univ. of Technol., P.O. Box 513, 5600 MB Eindhoven, The Netherlands)

Strong self-sustained acoustic oscillations may occur in a gas pipe network under certain gas flow velocities within the network. The pipe network under consideration consists of a main pipe, with a variable mean airflow, with two closed coaxial side branches of variable but equal length

joined to the main pipe at right angles. Coupling between resonant acoustic standing waves and instabilities of the shear layers separating the flow in the main pipe from the stagnant gas in the closed side branches leads to strong acoustic oscillations (easily 10% of the mean pressure) at a frequency corresponding to the half-wavelength acoustic mode defined by the total side-branch length. The acoustic power available in the resonant mode has been measured to compare with theoretical predictions and previous two-microphone measurement techniques as well as to inspire confidence when designing an aeroacoustically driven thermoacoustic heat pump. The acoustic load consists of a variable acoustic resistance and coupled acoustic compliant volume, which is similar to devices used to measure power in thermoacoustic engines. Data will be presented which characterize the performance of the aeroacoustic sound source under variable acoustic loading. [Work supported by Shell International Exploration and Production B.V.]

9:30

4aPAa7. Bifurcations of diffusive soliton solutions to Kuznetsov's equation. Pedro M. Jordan (Code 7181, Naval Res. Lab., Stennis Space Ctr., MS 39529, pjordan@nrlssc.navy.mil)

Exact traveling wave solutions are determined for Kuznetsov's equation, a nonlinear PDE of 3rd order which describes finite amplitude acoustic disturbances in thermoviscous Newtonian fluids. Specifically, it is shown that traveling wave solutions exist, and assume the form of diffusive solitons, if and only if the Mach number is less than or equal to a bifurcation value. It is also shown that the wave speed v is always supersonic, that $\text{Max}[v]$ occurs at the bifurcation value of the Mach number, and that a shock develops as the Reynolds number tends to infinity. Finally, special cases and asymptotic results are listed, a relationship to Burgers' equation is noted, and 3-D bifurcation diagrams are given.

THURSDAY MORNING, 1 MAY 2003

ROOMS 209/210, 10:00 TO 11:45 A.M.

Session 4aPAb

Physical Acoustics: Ultrasonics and Inhomogeneous Media

David L. Gardner, Chair

Material Division, Los Alamos National Laboratory, M.S. K764, Los Alamos, New Mexico 87545

Contributed Papers

10:00

4aPAb1. On the ability of bounded inhomogeneous waves to experimentally verify the behavior of infinite inhomogeneous plane waves. Nico F. Declercq, Joris Degrieck (Dept. of Mech. Constr. & Prod., Soete Lab., Ghent Univ., Sint Pietersnieuwstraat 41, B-9000 Gent, Belgium), and Oswald Leroy (IRC-KULAK Univ., E. Sabbelaan 53, 8500 Kortrijk, Belgium)

Infinite inhomogeneous plane waves have been controversial because of the appearance of a reflection coefficient exceeding unity in some cases. Even though it had already been shown that this is consistent with the principle of energy conservation, and results in a shift in position, full acceptance of the notion of infinite inhomogeneous plane waves appeared only after some experiments have been performed that produce results that are in agreement with theory. However, experiments can only be performed by means of bounded inhomogeneous waves. The current paper answers the ultimate question why bounded inhomogeneous waves are able to behave as if they were infinite inhomogeneous waves and what exactly is the physical and mathematical connection between bounded

inhomogeneous waves and infinite inhomogeneous waves. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

10:15

4aPAb2. Backward displacement of ultrasonic waves reflected from a corrugated interface. A. A. Teklu and M. A. Breazeale (Natl. Ctr. for Physical Acoust., The Univ. of Mississippi, 1 Coliseum Dr., Univ., MS 38677)

A backward displacement can occur when an ultrasonic beam is reflected from an interface having a superimposed periodicity. This was first observed by Breazeale and Torbett [Appl. Phys. Lett. **29**, 456 (1976)]. They observed that at a certain critical angle the reflected beam was shifted at 6 MHz and not at 2 MHz, in agreement with the theory of Tamir and Bertoni [J. Acoust. Soc. Am. **61**, 1397-1413 (1971)]; however, the angle predicted by assuming that the shift was caused by a Rayleigh surface wave did not agree with experiment. Recent experiments using the

schlieren technique showed backward displacement at an incident angle of 25 deg, in good agreement with the value 23.69 deg calculated using the inhomogeneous wave theory [Briers *et al.*, J. Acoust. Soc. Am. **106**, 682–7 (1999) (private communication with N. F. Declercq)]. The inhomogeneous wave theory also predicts a new kind of leaky surface wave at an incident angle of 22.55 deg [N. F. Declercq (private communication)], again in agreement with our measurements. Schlieren photographs of the backward-displaced beams at a water-brass interface are presented and discussed, as is the effect of the presence of an evanescent wave at the interface.

10:30

4aPAb3. Analysis of particular phononic structures using single integral equations. Fadoulourahmane Seydou, Ramani Duraiswami, and Nail A. Gumerov (Perceptual Interfaces and Reality Lab., Inst. for Adv. Computational Studies, Univ. of Maryland, College Park, MD 20742)

A fast method for determining phononic (and photonic) bandgaps in composite materials is developed. It is known that in the propagation of waves in a 3D medium containing N scatterers arranged periodically, there exist refractive indices for which such structures have bandgaps, i.e., frequencies for which no waves can propagate inside. Our task is to find the frequencies that generate these prohibited waves. This requires the solution of an eigenvalue problem for the Helmholtz operator. To solve this problem we choose an alternate route which uses boundary integral equations. We derive a single integral equation on each of the interfaces between the outer region and the scatterers, considering a general transmission boundary condition, by using a hybrid method using layer potentials and Green's formula. This approach reduces the number of unknowns considerably in comparison to other methods, but requires the treatment of large dense matrices and many matrix vector multiplications. To remedy this, we use the Fast Multipole Method. For solving the eigenvalue problem we discuss two methods: the Newton method and a method based on the Cauchy formula. Details of the numerical implementation, and results will be presented for different cases: sound hard, sound soft, impedance and transmission boundary conditions. [Work partially supported by NSF Award 0219681 is gratefully acknowledged.]

10:45

4aPAb4. Band-gap engineering in periodic acoustic stub tuners: Spectral gaps and transmission. Manvir S. Kushwaha (Inst. of Industrial Sci., Univ. of Tokyo, 4-6-1 Komaba, Meguro-Ku, Tokyo 153-8505, Japan), X. F. Wang, and P. Vasilopoulos (Concordia Univ., Montreal, QC H3G 1M8, Canada)

A theoretical investigation on acoustic wave propagation in a periodically stubbed waveguide is reported. In general the waveguide segments and stubs are made of different materials. The acoustic wave in such a system has two independent polarizations: out-of-plane and in-plane modes. The band structure and transmission spectrum are studied for diverse geometries using a simple and efficient version of the transfer-matrix method. For the *same material* between the waveguide and *symmetric* stubs the width of some gaps can change, upon varying the stub length or width, by more than one order of magnitude. A further modulation can be achieved for *different material* between the stubs and the main waveguide or if the stubs are *asymmetric*. The gaps in the band structure of an infinitely long system correspond to those in the transmission spectrum of the same system but with *finite* number n of units. For n finite (i) there exist pseudogaps that gradually turn into complete gaps with increasing n and (ii) the introduction of defects gives rise to states in the gaps and leads to transmission resonances.

11:00

4aPAb5. Group velocity manipulation in simple acoustic band gap filters. W. M. Robertson, C. Baker, and C. Brad Bennett (Dept. of Phys. and Astron., Middle Tennessee State Univ., Murfreesboro, TN 37132)

A simple experimental configuration is presented in which the group velocity of acoustic pulses is adjusted from much slower to much faster than the speed of sound. The experiment is an acoustic analog of two much studied optical phenomena: the superluminal group velocity achieved when optical pulses are tunneled through regions of high absorption or attenuation, and slow (or even stopped) light propagation in the vicinity of strong material dispersion, typically realized using electromagnetically induced transparency in atomic vapors. For the acoustic experiments described here, attenuation and dispersion were created using periodically structured waveguides. The periodicity results in one-dimensional acoustic band gaps, frequency intervals in which propagation of sound is strongly suppressed. The introduction of defects in the perfect periodicity leads to narrow transmission bands within the band gaps. There is strong normal dispersion in the vicinity of these defect modes leading to very slow group velocities. By tuning the carrier frequency of acoustic pulses through the forbidden and defect regions the group velocity can be adjusted from 0.22 to 1.8 times the speed of sound in the straight unstructured waveguide. These results are shown to be consistent with a straightforward theoretical model.

11:15

4aPAb6. Modified Born approximation for solving scattering problems. Igor L. Oboznenko (Kyiv Polytechnic Univ., Kyiv 03057, Ukraine), Vladimir Genis (Drexel Univ., Philadelphia, PA 19104), and Dat H. Tran (Nagoya Univ., Nagoya 464-8603, Japan)

Scattering of ultrasonic waves in inhomogeneous media is described by the inhomogeneous differential equations. Such equations could be solved using a method of consecutive approximations, such as the Born approximation. The Born approximation is applicable when $(n-1) < 1$ and $2ka(n-1) < 1$, where n is the refraction index and ka is the wave dimension of the scatterer. For the Born approximation, it is assumed that the acoustic field inside of a scatterer is substituted by the acoustic field of the incident wave, along with the wave number of the surrounding media. In this work, the modified Born approximation is used, where the acoustic field inside of a scatterer is substituted by the acoustic field of the incident wave, along with the wave number of the scatterer. A similar approach is used for solving the scattering problems of multilayered scatterers, which have weak scattering properties. The computed and experimental scattering characteristics for the elastic scatterers with various acoustical impedance are presented. It is demonstrated that the modified Born approximation more accurately describes the scattering problems for the scatterers with $ka > 1$.

11:30

4aPAb7. Modeling scattering enhancements at isolated resonances using energy conservation, reciprocity, symmetry, and the optical theorem. Philip L. Marston and Curtis F. Osterhoudt (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814, marston@wsu.edu)

Sound scattered by some objects in water exhibits isolated narrow resonances that are sufficiently large in amplitude to dominate the low-frequency scattering. Examples include the quadrupole mode of thin spherical shells and of solid plastic spheres [B. T. Hefner and P. L. Marston, J. Acoust. Soc. Am. **107**, 1930–1936 (2000)] and organ-pipe modes of water-filled pipes [C. F. Osterhoudt and P. L. Marston, J. Acoust. Soc. Am. **110**, 2773 (2001)]. This presentation concerns simple methods for approximating the scattering. In the case of spheres, ray theory for the backscattering reduces to a simple form for high- Q modes: Eq. (58) of Marston [J. Acoust. Soc. Am. **83**, 25–37 (1988)]. This result gives the

backscattering form function at resonance (in the usual normalization) to have the magnitude $2(2n+1)/ka$. Here n is the partial wave index associated with the mode of the sphere and ka is the product of the wave number and the sphere radius. This result may also be derived directly

from energy conservation and the optical theorem. Scattering amplitudes associated with high- Q organ pipe resonances of open cylindrical pipes are also derived here by a related method using the energy conservation, reciprocity, symmetry, and the optical theorem.

THURSDAY MORNING, 1 MAY 2003

ROOM 206, 8:00 A.M. TO 12:00 NOON

Session 4aPPa

Psychological and Physiological Acoustics: Spatial Hearing, Binaural Processing and Masking (Poster Session)

Marc A. Fagelson, Chair

Communicative Disorders, East Tennessee State University, Box 70643, Johnson City, Tennessee 37614-0643

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:30 a.m. and contributors of even-numbered papers will be at their posters from 10:30 a.m. to 12:00 noon. To allow for extended viewing time, posters will remain on display until 4:00 p.m.

4aPPa1. Median plane localization: dependence on spectral content. Gerald Ng and H. Steven Colburn (Hearing Res. Ctr., Boston Univ., 44 Cummington St. #427, Boston, MA 02215, geraldng@bu.edu)

Subjects were asked to report the elevation of noise stimuli differing in their spectral content. Some stimuli contained a single frequency band differing in center frequency and bandwidth; others contained multiple equal-amplitude frequency bands. All of the stimulus frequencies were in the range between 1 kHz and 16 kHz. Speakers were located in the median-sagittal plane at six frontal locations (-30° , -15° , 0° , 15° , 30° , 45°), as well as above (90°) and behind (180°) the listener. The experiments were conducted in a sound-treated room with the speakers concealed from view. Stimuli were 200 ms in duration and presented from only one speaker at a time. On each trial subjects used a pen to mark the perceived stimulus location(s) on a pre-printed coordinate diagram; subjects could indicate the perception of multiple sources and accurately represent auditory image widths and/or ambiguities. Preliminary data confirm historical results showing that localization of single noise bands depends on center frequency and bandwidth. When localizing multiple-band stimuli, responses depended on band density: with a few sparse bands, most subjects heard multiple source locations; for high band densities, subjects heard a single source at or near the true speaker location. [Work supported by NIDCD (R01 00100).]

4aPPa2. Cocktails for Franssen: Asynchronous transient bias and attentional factors in auditory localization. William M. Whitmer, Christopher A. Brown, Raymond H. Dye, Jr., and Noah F. Jurcin (Parmlly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, wwhitme@luc.edu)

The cocktail-party paradigm was applied to nonspeech signals. Pure-tone stimuli were used in an extension of the Franssen effect, an auditory illusion wherein the location of a slow-onset signal is perceived to be the same as a simultaneous, contralateral sudden-onset signal. Listeners heard simultaneous sudden-onset (transient) and contralateral slow-onset (steady-state) tones in a reverberant environment with a second delayed transient from a third azimuthal location. Results showed that the Franssen effect was either maintained or "reset," but not reduced. The ongoing steady-state tone was perceived either at the initial-transient and then

delayed-transient location, or at the initial transient location only. None of the listeners showed a location bias to the delayed transient tone when its frequency differed from the initial signal frequency or was replaced with noise. In additional conditions based on the "false Haas effect," consonant-vowel pairs representing transient and steady-state signals were segregated contralaterally in a reverberant space. Results showed no resemblance to the Franssen effect. In general, results indicated that the role of attention is fundamental to the localization of an ongoing stimulus. [Work supported by NIH.]

4aPPa3. Effects of reverberation and experience on distance perception in simulated environments. Matthew Schoolmaster, Norbert Kopco, and Barbara G. Shinn-Cunningham (Boston Univ. Hearing Res. Ctr., Boston, MA 02215, shinn@cns.bu.edu)

Individually measured head-related-transfer functions were used to simulate different acoustic environments in order to see how listener experience influences auditory distance perception. Three environments were simulated: an anechoic space, a classroom with the listener in the room center, and the same classroom with the listener in the room corner. Each subject completed two series of trials consisting of 360 trials in each environment. Each series consisted of six experimental sessions, with only one session performed per day. In the fixed-room series, one environment was tested in each session (ordered randomly). In the mixed-room series, trials from the three environments were intermingled within each session. Listeners were divided into two groups: Group A performed the fixed-room series followed by the mixed-room series, while Group B did the reverse. Sources were simulated from ahead and to the right of the listener at distances ranging from 0.15 to 1.7 m. Preliminary analysis indicates that performance improves with experience in the fixed-room but not the mixed-room trials. These results suggest that listeners learn to calibrate auditory distance percepts based on recent experience with the reverberation and echoes in a particular environment. [Work supported by AFOSR and the Sloan Foundation.]

4a THU. AM

4aPPa4. Cues for “front-back confusion.” Mitsuo Matsumoto, Kiyooki Terada, Mikio Tohyama (Kogakuin Univ., 2665-1 Nakano-machi, Hochioji, Tokyo 192-0015, Japan, matsu@tohyama.cc.kogakuin.ac.jp), and Hirofumi Nakajima (Nittobo Acoust. Eng. Co., Ltd., 1-13-12 Midori, Sumida-ku, Tokyo 130-0021, Japan)

Front-back confusion in the median plane disturbs sound image localization, especially for a sound image simulated using a transaural system and a dummy head. In the median plane, the spectral cue is a major cue for front-back discrimination. The front-back difference in HRTF characteristics between real heads and a dummy head was measured. For real heads, the frequency amplitude at 1.5 kHz with a front-source HRTF was significantly smaller than with a rear-source HRTF. In contrast, for the dummy head, no significant difference was found. Two signals band-limited up to 8 kHz in a subjective test were used: one had a dip amplitude frequency of -15 dB at 1.5 kHz on the frequency axis (characteristic of gearh) and the other had a peak of 15 dB (characteristic of gearh) at the same frequency. A sound image reproduced by an actual frontal sound source and the signal with the peak was perceived ambiguously. A sound image reproduced by an actual rear sound source and the signal with the dip was perceived to be ambiguous.

4aPPa5. Auditory performance in an open sound field. Kim F. Fluit, Tomasz Letowski, and Timothy Mermagen (US Army Res. Lab., Auditory Res. Team, AMSRL-HR-SD, Bldg. 520, kfluit@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 type of sound, distance to a sound source, terrain configuration, meteorological conditions, hearing capabilities of the listener, level of background noise, and the listener's 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. The purpose of the present study was to determine the listener's abilities to detect, recognize, localize, and estimate distances to sound sources from 25 to 800 m 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 to 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.

4aPPa6. Insensitivity to large differences in interaural correlation. John F. Culling (School of Psych., Cardiff Univ., Cardiff CF11 9BQ, UK)

Sensitivity to differences in interaural correlation was measured using an adaptive threshold technique. The 2-down/1-up adaptive track began at the opposite correlation value from the standard (e.g., starting at 1 and converging on -1). Each 3I-FC trial was followed by correct/incorrect feedback. Steps in the adaptive track were initially linear and switched to logarithmic as the track passed zero interaural correlation. First, four listeners were trained to asymptote using a band of noise (462–540 Hz). The listeners were then trained under identical circumstances, but with remote flanking bands (0–330 and 719–3000 Hz) of interaurally correlated, synchronously gated noise, equal in spectrum level to the target band. The listeners were also occasionally given an adaptive track beginning at zero interaural correlation. With no flanking bands, three listeners experienced difficulty learning to adapt from 1 to -1 , frequently recording threshold differences in correlation close to 2. Little difficulty was experienced when adapting from -1 to 1 and none when adapting from zero in either direction. When remote flanking bands were added, all listeners experienced

greater difficulty. Only one listener learned to converge on -1 . The salience of cues used by the listeners seemed to decrease when comparing high positive and negative correlations.

4aPPa7. Binaural versus better-ear listening. Jacob W. Scarpaci, N. I. Durlach, and H. Steven Colburn (Hearing Res. Ctr. and Dept. of Biomed Eng., Boston Univ., Boston, MA 02215, scarpaci@bu.edu)

Advantages of binaural over monaural hearing in noisy environments are reduced when the monaural stimulation is derived from the monaural signal with the better signal-to-noise ratio (better-ear listening). In the reported experiments, conducted in a soundproof room with two speakers and a custom-designed, noise-cancellation headset, speech intelligibility in the presence of interference was measured for both binaural and better-ear configurations. The headset, which incorporated two microphones (located at the two ears) and two insert earphones, was used to present binaural stimulation or better-ear (better-microphone) monaural stimulation. Although the results varied significantly with the locations of the target and interference sources, the advantage of binaural listening over better-ear listening was no more than a few dB. In addition to reporting the data obtained in these experiments, relations to previous work on better-ear listening and CROS hearing aids, as well as to current work on cochlear implants, are discussed. [Work supported by NIDCD (00100).]

4aPPa8. Modulation masking of lateralization based on envelope interaural phase differences. Stanley Sheft and William A. Yost (Parmlly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, ssheft@luc.edu)

Modulation-filterbank models discard envelope phase above very low modulation rates. To evaluate the role of relative phase in binaural envelope processing, modulation masking of envelope-based lateralization was measured with a left/right task. A diotic 400-ms wideband noise was modulated by a two-tone function. An interaural phase difference was applied to the probe component of the modulator with the masker component always diotic. The probe modulation rate was 80, 160, or 320 Hz with masker rate varying from 40 to 1280 Hz. Masker modulation interfered with probe lateralization at all probe rates. With 80-Hz probe modulation, the masking function was asymmetric with maskers above the probe more effective than those below. At the higher probe modulation rates, interference was generally greatest when the probe and masker rates were close. However, function asymmetry persisted at these higher rates. In some conditions, a diotic phase shift of the masker had a significant effect. Unlike monaural results, binaural envelope processing showed high-rate phase sensitivity and asymmetric tuning. The asymmetric results and masker-phase effects are consistent with viewing the envelope manipulations as interaural gating asynchronies described by group delay in the modulation domain. [Work supported by NIH.]

4aPPa9. Speech perception in tight acoustic spaces. Nandini Iyer, Rama Ratnam, Sandeep Phatak (Beckman Inst., 405 N. Mathews Ave., Urbana, IL 61801), Charissa Lansing, and Albert Feng (Beckman Inst. & Univ. of Illinois at Urbana—Champaign, Urbana, IL 61801)

Acoustic cues for sound localization and speech perception in reverberant environments are more complex than those in anechoic conditions. The ability to utilize these cues deteriorates as a function of age. This study examined the speech perception ability of elderly listeners in well-defined, enclosed spaces and identified signal-processing strategies that might enhance this ability. For this, we performed tests in three environments: a car, a plywood cube, and a conference room. Pre-recorded mono-

syllabic words (Modified Rhyme Test) were played in quiet or in the presence of an 8-talker babble from a loudspeaker and recorded using two microphones placed at KEMARs ears. These signals were played to 10 young and 10 elderly listeners with normal hearing. We evaluated speech perception abilities of the listeners as a function of the following variables: (1) the location and distance of the loudspeaker from the listener, and (2) the number of reflecting surfaces. We tested the hypothesis that diotic presentations of the signal (better ear signal to both ears, or active steering followed by a presentation of the summed signal) improved speech perception abilities in these situations.

4aPPa10. Discrimination between harmonic tone complexes and bandpass noise signals presented within lowpass noise. Steven van de Par, Armin Kohlrausch, and Jeroen Breebaart (Philips Res., Prof. Holstlaan 4, 5656 AA, Eindhoven, The Netherlands)

Threshold levels were measured at which harmonic tone complex targets could be *discriminated* from bandpass noise targets both presented within a diotic 2-kHz lowpass filtered noise masker with 65 dB SPL. In the 3IFC paradigm, one interval contained a sine-phase tone complex (20-Hz spacing) while the remaining intervals contained a bandpass noise with the same bandwidth, center frequency (1 kHz) and overall level as the tone complex. In addition, *detection* thresholds were measured for both targets using the same masker. When interaurally in-phase targets with a 100-Hz bandwidth were used, discrimination thresholds were about 8 dB higher than the detection thresholds. When out-of-phase targets were used, the difference between discrimination and detection thresholds was 18 dB, which was mainly due to a decrease of about 15 dB of the detection thresholds (BMLD). For 1500-Hz wide targets, the discrimination and detection thresholds were more similar. For the in-phase condition, the discrimination and detection threshold overlapped while for the out-of-phase conditions, discrimination thresholds exceeded the detection thresholds by only 5 dB. Assuming that discrimination is based on the processing of the temporal structure of the targets, this processing seems to be more efficient for wide- than for narrow-band targets.

4aPPa11. Basilar-membrane nonlinearity effects on tones and speech. Judy R. Dubno, Amy R. Horwitz, and Jayne B. Ahlstrom (Dept. of Otolaryngol.–Head & Neck Surgery, Medical Univ. of South Carolina, 135 Rutledge Ave., P.O. Box 250550, Charleston, SC 29425, dubnojr@muscc.edu)

The contribution of nonlinearities in the basilar-membrane response to the understanding of speech in noise was estimated by measuring growth of simultaneous masking for tones and speech. Speech signals were bandpass-filtered nonsense syllables; consonant recognition and pure-tone thresholds were measured as a function of masker level in two conditions: (1) masker within the passband of the speech and (2) masker below the passband of the speech. With this low-frequency masker, the effect of the masker at the “signal” place is assumed to grow linearly whereas the growth of response to the speech at the “signal” place is compressed. Growth-of-masking functions were also measured for short-duration tones in narrowband maskers centered at or below the signal frequency. Subjects were younger and older adults with normal hearing; these older adults were included given that age-related changes in nonlinearities may occur even with minimal threshold elevation. For both subject groups, large individual differences were observed in upward spread of masking and in the extent to which the subsequent variance in audibility accounted for the variance in speech-recognition scores. These results will be discussed along with estimates of basilar-membrane nonlinearities obtained from subject’s growth-of-masking functions and otoacoustic emissions. [Work supported by NIH/NIDCD.]

4aPPa12. Upward spread of informational masking in normal-hearing and hearing-impaired listeners. Joshua M. Alexander and Robert A. Lutfi (Univ. of Wisconsin–Madison, Waisman Ctr., 1500 Highland Ave., Madison, WI 53705)

Thresholds for pure-tone signals of 0.8, 2.0, and 5.0 kHz were measured in the presence of a simultaneous multitone masker in 15 normal-hearing and 8 hearing-impaired listeners. The masker consisted of fixed-frequency tones ranging from 522–8346 Hz at 1/3-octave intervals, excluding the 2/3-octave interval on either side of the signal. Masker uncertainty was manipulated by independently and randomly playing individual masker tones with probability $p=0.5$ or $p=1.0$ on each trial. Informational masking (IM) was estimated by the threshold difference ($p=0.5$ minus $p=1.0$). Decision weights were estimated from correlations of the listener’s response with the occurrence of the signal and individual masker components on each trial. IM was greater for normal-hearing listeners than for hearing-impaired listeners, and most listeners had at least 10 dB of IM for one of the signal frequencies. For both groups, IM increased as the number of masker components below the signal frequency increased. Decision weights were also similar for both groups—masker frequencies below the signal were weighted more than those above. Implications are that normal-hearing and hearing-impaired individuals do not weight information differently in these masking conditions and that factors associated with listening may be partially responsible for the greater effectiveness of low-frequency maskers. [Work supported by NIDCD.]

4aPPa13. Criteria placement in staircase procedures. Virginia M. Richards and Rong Huang (Dept. of Psych., Univ. of Pennsylvania, 3815 Walnut St., Philadelphia, PA 19104, rongh@psych.upenn.edu)

Data from several masking studies are evaluated to determine whether observers alter their decision criteria as signal levels increase and decrease according to an adaptive staircase algorithm. In a Yes/No procedure, a 1000-Hz signal tone is added to one of twelve maskers. Each masker is composed of six equal-amplitude tones with frequencies drawn at random from a 200–5000 Hz range. In two conditions, the maskers are either fixed across several sessions or randomly drawn prior to each trial. Data from at least 15 staircases are combined, and for each signal level the overall hit and false alarm rates are determined. Regardless of condition, fixed or random, observers tend to adopt a single criterion across all signal levels. If it is assumed that in the random condition a single criterion is used across all trials, the location of the mean of the no-signal distributions relative to the criterion can be estimated. For the current data, the pattern of no-signal distributions is not obviously related to masker properties. [Work supported by NIH.]

4aPPa14. Development of a fast method for determining psychophysical tuning curves. Aleksander Sek, Ewa Skrodzka (Inst. of Acoust., A. Mickiewicz Univ., 61-614 Poznan, Umultowska 85, Poland), and Brian C. J. Moore (Dept. of Exp. Psych., Univ. of Cambridge, Downing St., Cambridge CB2 3EB, UK)

A psychophysical tuning curve (PTC) is usually measured by determining the level of a narrowband noise required to mask a fixed low level tone, for several masker center frequencies. PTCs could be useful clinically for assessing frequency selectivity and for the diagnosis of dead regions in the cochlea. When the signal frequency falls in a dead region, the tip of the PTC is shifted away from the signal frequency. However, PTCs determined in the traditional way are too time consuming for use in clinical practice. A fast method for determining PTCs is being developed and evaluated. This uses a band of noise that sweeps in center frequency. A Bekey method is used to track the masker level required for threshold. For normally hearing subjects, the new method gives stable results, and PTCs similar in shape to those determined in the traditional way, when the masker sweeps over a 2-octave range in about 4 minutes, and the level

changes by 1–2 dB/s. Preliminary results using hearing-impaired subjects also show a good agreement with the traditional method. However, further work is required to determine optimum values for the noise bandwidth and rate of change of frequency and level.

4aPPa15. Phase rotation thresholds and its use in watermarking technique. Aleksander Sek, Yoiti Suzuki, Ryouichi Nishimura, and Kotaro Sonoda (Res. Inst. of Elec. Commun., Tohoku Univ., Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan)

Strong demands for the protection of copyrights for high fidelity audio recordings as well as broadcasting signals via a network brought about a development of so-called watermarking techniques. One of them is based on a phase modulation referred to in here as a phase rotation (PR). This method assumes a poor sensitivity of the auditory system to the phase changes in a signal and embeds a watermark in the recording by means of a filter with linear phase characteristics. However, the sensitivity of the auditory system to the phase rotation has not been studied in detail, especially for complex signals. In this paper we present a preliminary result of threshold measurements of the phase rotation applied to different types of signals. It appears that the thresholds for detecting the phase rotation, when expressed by means of a time group delay, depend on the type of the signal and the frequency region that the phase rotation is applied for. The detection of the phase rotation is approximately independent on the rotation rate up to 120 Hz. When the phase rotation was applied to several components of a harmonics complex, the highest values of the thresholds were observed in the low frequency region.

4aPPa16. Psychometric functions for informational masking. Robert A. Lutfi, Doris J. Kistler, Michael R. Callahan, and Frederic L. Wightman (Waisman Ctr., Univ. of Wisconsin, Madison, WI 53706)

The method of constant stimuli was used to obtain complete psychometric functions (PFs) from 44 normal-hearing listeners in conditions known to produce varying amounts of informational masking. The task was to detect a pure-tone signal in the presence of a broadband noise and in the presence of multitone maskers with frequencies and amplitudes that varied at random from one presentation to the next. Relative to the broadband noise condition, significant reductions were observed in both the slope and the upper asymptote of the PF for multitone maskers producing large amounts of informational masking. Slope was affected more for some listeners while asymptote was affected more for others. Mean slopes and asymptotes varied nonmonotonically with the number of masker components in much the same manner as mean thresholds. The results are consistent with a model that assumes trial-by-trial judgments are based on a weighted sum of dB levels at the output of independent auditory filters. For many listeners, however, the weights appear to reflect how often a nonsignal auditory filter is mistaken for the signal filter. For these listeners adaptive procedures may produce a significant bias in the estimates of threshold for conditions of informational masking. [Work supported by NIDCD.]

4aPPa17. Effect of signal-masker similarity and signal uncertainty on informational masking. Eunmi L. Oh (Samsung AIT, P.O. Box 111, Suwon 440-600, Korea, oh@sait.samsung.co.kr) and Robert A. Lutfi (Univ. of Wisconsin, Madison, WI 53706)

An experiment was conducted to determine the relative importance of signal-masker similarity and signal uncertainty as factors affecting informational masking. The masker was a complex of fixed-frequency tones whose levels varied independently and at random on each presentation. The signal was a complex of fixed-frequency tones whose levels were sampled from one of two Gaussian distributions differing in mean ($SD = 5$ dB). The levels of the signal tones were sampled independently (IS condition) or were the same as determined by a single sample (SS condi-

tion). The listener's task was to select the signal from the distribution with the higher mean. It was expected that if uncertainty were the dominant factor better performance would be obtained in the IS condition—this due to the reduction in uncertainty associated with multiple independent “looks” at the signal. If, however, similarity were the dominant factor better performance would be obtained in the SS condition—this due to the dissimilarity of signal and masker spectra (flat versus irregular). Four of six listeners showed the first result, the remaining two the second result. The remaining two listeners, however, showed results inconsistent with a similarity effect with maskers having flat spectra. [Work supported by NIDCD.]

4aPPa18. Acceptance of background noise as a function of speech presentation level. Clifford A. Franklin, Jr., Anna K. Nabelek, and Samuel B. Burchfield (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, South Stadium Hall, Knoxville, TN 37996-0740)

The acceptance of background noise while listening to speech (ANL) at different speech presentation levels was assessed. Twenty listeners (10 male) between 18–30 years with normal hearing listened to a narrative at speech presentation levels of 20, 34, 48, 62, and 76 dB HL, then adjusted the background noise to the highest intensity level that they would be willing to accept for an extended listening period. The ANL is the intensity of the speech presentation level minus the intensity level of the background noise. The group mean ANLs for presentation levels of 20, 34, 48, 62, and 76 dB HL were 10.60, 14.25, 17.10, 21.80, and 24.55 dB, respectively. The group mean ANLs differ by approximately three and one half decibels between each presentation level. This difference between adjacent speech presentation levels is representative of a linear function. The average MCL was 43 dB HL with a standard deviation of 6.7 dB. The group mean ANL for speech presented at MCL was 15.5 dB with a standard deviation of 7.27 dB. A MANOVA for repeated measures indicated a statistically significant main effect for speech presentation level ($F = 18.624$, $p = 0.001$). [Work partially supported by NIDCD (NIH) R01 DC 05018.]

4aPPa19. Comparison of acceptance of background noise and speech reception threshold in quantifying the hearing aid directivity benefit. Melinda C. Freyaldenhoven, Anna K. Nabelek, and Samuel B. Burchfield (Dept. of Audiol. and Speech Pathol., The Univ. of Tennessee, South Stadium Hall, Knoxville, TN 37996-0740)

Hearing aid directivity benefit was compared as improvement in acceptance of background noise and speech reception threshold (SRT). Forty adult subjects were tested wearing binaural hearing aids in omnidirectional and directional listening conditions. Acceptance of background noise was determined by having subjects select their most comfortable listening level (MCL) for a story delivered from a loudspeaker (0). Next, speech babble was added (180) and the subjects selected the maximum background noise level (BNL) which was acceptable while listening to and following the story. The MCL minus the BNL yielded the acceptable noise level (ANL), all in dB. The difference between the ANL for the omni-directional and directional conditions is the directivity benefit. The SRT was determined by delivering spondaic words (0) at the subjects MCL. Next, speech babble was delivered (180) and adjusted until the subject could repeat 50% of the spondees. The difference between the SRT for the omnidirectional and directional conditions is the directivity benefit. Mean directional benefit, $ANL = 3.50$ dB and $SRT = 3.60$ dB, were not significantly different. The individual ANLs and SRTs were significantly correlated ($r = -0.36$, $p = 0.002$). The ANL procedure appears to be a viable tool for quantifying hearing aid directivity benefit. [Work supported by NIDCD (NIH) 3 R01 DC 05018-01S1.]

4aPPa20. Relationship between acceptance of background noise and hearing aid use. Anna K. Nabelek, Samuel B. Burchfield, and Joanna D. Webster (Dept. of Audiol. and Speech Pathol., The Univ. of Tennessee, South Stadium Hall, Knoxville, TN 37996-0740, samburch@utk.edu)

Background noise produces complaints among hearing-aid users, however speech-perception-in-noise does not predict hearing-aid use. It is possible that hearing-aid users are complaining about the presence of background noise and not about speech perception. To test this possibility, acceptance of background noise is being investigated as a predictor of hearing-aid use. Acceptance of background noise is determined by having subjects select their most comfortable listening level (MCL) for a story. Next, speech-babble is added and the subjects select the maximum background noise level (BNL) which is acceptable while listening to and following the story. The difference between the MCL and the BNL is the acceptable noise level (ANL), all in dB. ANLs are being compared with hearing-aid use, subjective impressions of benefit (APHAB), speech perception in background noise (SPIN) scores, and audiometric data. Individuals who accept higher levels of background noise are more successful users than individuals who accept less background noise. Mean ANLs are 7.3 dB for full-time users ($N=21$), 12.6 dB for part-time users ($N=44$), and 13.8 dB for rejecters ($N=17$). ANLs are not related to APHAB, SPIN, or audiometric data. Results for about 120 subjects will be reported. [Work supported by NIDCD (NIH) RO1 DC 05018.]

4aPPa21. Relation between measures of speech-in-noise performance and measures of efferent activity. Brad Smith, Ashley Harkrider, Samuel Burchfield, and Anna Nabelek (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, 457 South Stadium Hall, Knoxville, TN 37996)

Individual differences in auditory perceptual abilities in noise are well documented but the factors causing such variability are unclear. The purpose of this study was to determine if individual differences in responses measured from the auditory efferent system were correlated to individual variations in speech-in-noise performance. The relation between behavioral performance on three speech-in-noise tasks and two objective measures of the efferent auditory system were examined in thirty normal-hearing, young adults. Two of the speech-in-noise tasks measured an acceptable noise level, the maximum level of speech-babble noise that a subject is willing to accept while listening to a story. For these, the acceptable noise level was evaluated using both an ipsilateral (story and noise in same ear) and a contralateral (story and noise in opposite ears) paradigm. The third speech-in-noise task evaluated speech recognition using monosyllabic words presented in competing speech babble. Auditory efferent activity was assessed by examining the resulting suppression of click-evoked otoacoustic emissions following the introduction of a contralateral, broad-band stimulus and the activity of the ipsilateral and contralateral acoustic reflex arc was evaluated using tones and broad-band noise. Results will be discussed relative to current theories of speech in noise performance and auditory inhibitory processes.

4aPPa22. Effects of multiple background talkers on word recognition and response awareness. Edward L. Goshorn and Elizabeth K. Robertson (Speech Dept., P.O. Box 3165 Tech Station, Louisiana Tech Univ., Ruston, LA 71272, egoshorn@ltparts.latech.edu)

Effects of background talkers (0, 1, 2, 3, 4, 5, 7, 10, and 14) on word recognition and awareness of errant/accurate responses were examined. Diagnostic Rhyme Test (DRT) words and background talkers were presented at 70 dB SPL (sound field) to ten normal-hearing subjects. DRT words and background talkers were digitally processed to produce equal VU meter levels. Three replicates were obtained for each condition. Performance measures were: (1) percent correct, corrected for guessing, transformed to rational arcsine units (PCGRAU); (2) subject's awareness of accurate responses (AA); (3) subject's awareness of errant responses

(AE); and (4) a geometric-based symmetric awareness (SA). Awareness measures were derived from subject's confidence ratings to DRT stimuli. Both informational and direct masking effects were present. PCGRAU varied nonlinearly as number of talkers increased. One talker provided significantly more masking than two talkers and was equally effective as three, four, five, seven, and ten talkers as well as speech-spectrum noise. Fourteen talkers provided the most masking and was equally effective as white noise. In general, AA tended to diminish while AE tended to improve as additional background talkers were added. SA was best for 14-talker background noise but poorest for one-talker and speech-spectrum noise.

4aPPa23. Speech perception in gated noise: The effects of spectral and temporal resolution and auditory streaming. Su-Hyun Jin and Peggy B. Nelson (Dept. of Commun. Disord., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Hearing-impaired (HI) listeners often report difficulty understanding speech in the presence of background noise, even in the presence of mild degrees of hearing loss. In addition, HI listeners show significantly less benefit from fluctuations in noise than do normal hearing (NH) listeners. Furthermore, HI listeners often show less accuracy of processing frequency and temporal information in acoustic signals and grouping them into a whole meaningful speech. The purpose of this study is to examine differences in performances of NH and HI listeners for speech perception in various types of noise. Three hypotheses will be examined in this study. First, the amount of masking release in HI listeners with mild hearing loss will be smaller than in NH listeners even when speech and noise are presented at intensities sufficient to overcome the hearing loss. Masking release will be measured for consonant recognition and sentence recognition by subtracting the percent correct in steady noise from that in gated noise. Second, the performance of HI listeners on spectral resolution, temporal resolution, and auditory streaming tasks will be significantly poorer than that of NH listeners. Third, the performance differences in these tasks may account for differences in masking release.

4aPPa24. Noise suppression algorithm based on the auditory masked threshold in listeners with cochlear hearing loss. Kathryn Arehart, Jessica Rossi-Katz (Speech-Lang.-Hearing Sci., Univ. of Colorado, 409 UCB, Boulder, CO 80309), Ajay Natarajan, and John Hansen (The Ctr. for Spoken Lang. Res., 594 UCB, Boulder, CO 80309-0594)

This study describes the formulation and evaluation of a new noise suppression scheme, with the goal of improving speech-in-noise perception by hearing-impaired listeners. Arehart *et al.* [Speech Commun. (2003)] implemented and evaluated a noise suppression algorithm based on an approach that used the auditory masked threshold in conjunction with a version of spectral subtraction to adjust the parameters used in the subtraction process based on the masked threshold of the noise across the frequency spectrum. That original formulation was based on masking properties of the normal auditory system, with its theoretical underpinnings based on MPEG-4 audio coding [Johnston (1988)]. This paper describes details of a new enhancement formulation based on masking characteristic of cochlear hearing loss. The new algorithm improves on previous formulations in two ways. First, the algorithm is implemented with generalized minimum mean-square error estimators, which provide improvements over spectral subtraction estimators. Second, estimation of the auditory masked thresholds and masking spreading functions are adjusted to address elevated thresholds and broader auditory filters characteristic of cochlear hearing loss. In addition to algorithm details, results of the algorithm evaluation (using objective quality measures; intelligibility and quality measures in hearing-impaired listeners) will be presented. [Work supported by the Whitaker Foundation.]

Session 4aPPb

Psychological and Physiological Acoustics: Pitch, Speech Perception and Loudness

Craig A. Champlin, Chair

Department of Communication Sciences and Disorders, University of Texas, 1 University Station A1100, Austin, Texas 78712-0114

Contributed Papers

8:15

4aPPb1. Pitch discrimination as a function of the inter-stimulus interval: Evidence against a simple model of perceptual memory.

Laurent Demany, Gaspard Montandon, and Catherine Semal (CNRS and Univ. Victor Segalen, 146 rue Leo Saignat, F-33076 Bordeaux, France, Laurent.Demany@psyac.u-bordeaux2.fr)

A listener's ability to compare two sounds separated by a silent time interval T is limited by a sum of "sensory noise" and "memory noise." The present work was intended to test a model according to which these two components of internal noise are independent and, for a given sensory continuum, the memory noise depends only on T . In three experiments using brief sounds (<80 ms), pitch discrimination performances were measured in terms of d' as a function of T (0.1–4 s) and a physical parameter affecting the amount of sensory noise (pitch salience). As T increased, d' first increased rapidly and then declined more slowly. According to the tested model, the relative decline of d' beyond the optimal value of T should have been slower when pitch salience was low (large amount of sensory noise) than when pitch salience was high (small amount of sensory noise). However, this prediction was disproved in each of the three experiments. It was also found, when a "roving" procedure was used, that the optimal value of T was markedly shorter for very brief tone bursts (6 sine cycles) than for longer tone bursts (30 sine cycles).

8:30

4aPPb2. F_0 discrimination interference: Effects of resolved tone complexes and noise on fundamental frequency discrimination of unresolved complex tones.

Hedwig Gockel, Robert P. Carlyon (MRC Cognition and Brain Sci. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK, hedwig.gockel@mrc-cbu.cam.ac.uk), and Christopher J. Plack (Univ. of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK)

F_0 discrimination of a 400-ms complex tone with only unresolved components ("target") was investigated in the absence and presence of a synchronously gated resolved complex tone ("interferer"). The target and the interferer were bandpass filtered from 1375–15 000 Hz and 125–625 Hz, respectively. In a 2I-2AFC task, listeners indicated the interval containing the target with the higher pitch. The nominal F_0 of the target was 88 Hz; that of the interferer was constant across the two intervals and was either 88 Hz or increased by various amounts. Although the target and interferer were in well-separated frequency regions, performance (percent correct) dropped by about 16% when the interferer's F_0 was 88 Hz. The impairment was halved when the interferer's F_0 was 10% higher than that of the target, and almost eliminated when it was 30% higher. Increasing the level of a 1375-Hz low-pass-filtered noise, gated synchronously with the target and the interferer (F_0 equaled 88 Hz), improved performance, further demonstrating that the deterioration produced by the resolved complex was not due to peripheral masking. The results are consistent with a form of across-frequency interference at the level of pitch perception. [Work supported by EPSRC Grant GR/R65794/01.]

8:45

4aPPb3. Effects of relative frequency, absolute frequency, and phase on fundamental frequency discrimination: Data and an autocorrelation model.

Joshua Bernstein and Andrew Oxenham (MIT Res. Lab. of Electronics and Harvard-MIT Speech & Hearing Bioscience & Technol. Prog., 77 Massachusetts Ave., Cambridge, MA 02139, jgbern@mit.edu)

Fundamental frequency (F_0) difference limens (DLs) were measured versus F_0 for sine- and random-phase harmonic complexes bandpass-filtered into low- or high-frequency regions, with 3-dB passbands of 2.5–3.5 and 5–7 kHz, respectively. In all cases, F_0 DLs decreased dramatically with increasing F_0 as approximately the tenth harmonic appeared in the passband. Generally, F_0 DLs were similar in both frequency regions for complexes with similar harmonic numbers and phase relationships. However, F_0 DLs were larger in the high-frequency than the low-frequency region for random-phase complexes containing only harmonics above the tenth, suggesting a possible role for additional fine-structure information in the low-frequency region. The dependence of F_0 discrimination on relative frequency presents a significant challenge to autocorrelation (AC) models of pitch perception, in which predictions generally depend more on absolute frequency and phase locking. To represent this relative frequency effect, a "lag window" modification to the Meddis and O'Mard [J. Acoust. Soc. Am. **102**, 1811–1820 (1997)] AC model was introduced, restricting each channel's AC representation to a limited range of lags relative to the center frequency (CF). Thus in each channel, the AC model responds best to F_0 's whose dominant harmonics fall near the CF. [Work supported by NIH Grant Nos. 5T32-DC-00038 and R01-DC-05216.]

9:00

4aPPb4. Evidence for distinct types of "perfect pitch."

David A. Ross (Dept. of Diagnostic Radiol., Yale School of Medicine, Box 208043, New Haven, CT 06520), John C. Gore (Vanderbilt Univ. Medical Ctr., Nashville, TN 37232), and Lawrence E. Marks (John B. Pierce Lab., New Haven, CT 06519)

The ability to identify and reproduce sounds of specific frequencies, typically called "perfect pitch," is remarkable and uncommon. Whether this skill is learned early in life or inherited has been a matter of great controversy. Further, a substantial literature suggests that "perfect pitch" may be heterogeneous. Previously, we proposed a model to account for heterogeneity. The model subdivides individuals capable of naming notes accurately into two groups: possessors of true absolute pitch (AP), who automatically encode the frequency of all tonal stimuli, precategorically and independent of their source; and possessors of heightened tonal memory (HTM), who identify tones by comparing them to a memorized tonal template. The ability of individuals with HTM to identify tonal stimuli should depend strongly on the tones' acoustical properties, such as timbre or chroma. Three experiments sought to test this hypothesis directly. Individuals claiming "perfect pitch" were recruited and initially classified as having AP or HTM. Consistent with the model, the two groups differed significantly in their sensitivity to the targets' timbre, chroma, and tonal context, suggesting that they may use different mecha-

nisms to identify tonal stimuli. The model may help reconcile the long-standing controversy between early learning and genetic theories of "perfect pitch."

9:15

4aPPb5. Effects of real and illusory glides on pure-tone frequency discrimination. Johannes Lyzenga, Robert P. Carlyon (MRC Cognition and Brain Sci. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK), and Brian C. J. Moore (Univ. of Cambridge, Cambridge CB2 3EB, UK)

Pure-tone frequency difference limens (DLs) were measured for 500-ms, 4-, and 6-kHz pure tones, where the interval between the initial and final constant-frequency portions of the tones in each 2IFC trial consisted of either a silent gap, a frequency glide, or a noise burst. The noise was inserted to create the illusion of the tone continuing through the gap. The interval between the two constant-frequency portions was 0, 10, 50, or 200 ms; a condition with a 500-ms frequency glide without constant-frequency portions was also used. To prevent subjects from using information from the end point of the glides, a frequency rove of approximately four ERBs was used in all conditions. It was found that DLs were lower for the glide than for the gap condition, supporting the conclusion that the auditory system contains a mechanism specific for the detection of dynamic frequency changes [A. Sek and B. C. J. Moore, *J. Acoust. Soc. Am.* **106**, 351–359 (1999)]. For a number of subjects, the DLs were smaller for the noise condition than for the gap condition, suggesting that this mechanism may be able to operate on a stage of processing at which an illusory glide has been introduced.

9:30

4aPPb6. Comparison of behavioral discrimination, MMN, and P300 to speech and nonspeech stimuli. Joanna Webster, Ashley Harkrider, and Mark Hedrick (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, 457 South Stadium Hall, Knoxville, TN 37996)

Our objective is to examine the relation between central auditory processes and discrimination of speech (consonant–vowel) and nonspeech (frequency glide) stimuli. Behavioral responses and auditory evoked potentials (MMN and P300) of ten adults were evaluated to synthetically generated consonant–vowel (CV) speech and nonspeech contrasts. The CVs were two within-category stimuli and the nonspeech stimuli were two frequency glides whose frequencies matched the formant transitions of the CV stimuli. It was found that listeners exhibited significantly better behavioral discrimination to the nonspeech versus speech stimuli in same/different and oddball behavioral paradigms. MMN responses were present in all subjects to both stimulus contrasts, and were not significantly different with regard to stimulus type. P300's were present in nine of ten subjects to both stimulus contrasts. However, the CV speech contrasts produced P300's with significantly smaller amplitudes and longer latencies than those to the nonspeech stimuli. These results suggest that the stimuli were processed differently when measured behaviorally and with the P300, but not when measuring the MMN. The enhanced discrimination of the frequency glide stimuli versus the CV stimuli of analogous acoustical content supports the idea that different levels of processing mediate the auditory perception of speech versus nonspeech stimuli.

9:45

4aPPb7. Phoneme contrasts with two, same-talker speakers. Mark Ericson (2610 Seventh St., Wright–Patterson AFB, OH) and Pamela Mishler (Dept. of Veteran Affairs, Dayton, OH)

People often have difficulty hearing speech in the presence of concurrent conversations. This well-known cocktail-party effect can be parsed into energetic and informational masking effects. The purpose of this study was to measure and model the effects of energetic and informational

masking that occur when two words, spoken by the same talker are heard at the same time. The word identification test used in the experiments was the Modified Rhyme Test (MRT). The MRT was used as both the stimulus and the masker, which afforded a multitude of consonant contrasts. The phrases 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 spoken by three male and three female talkers. The dependent variable was the percent correct identifications of the two consonants. Listeners performed at 94% correct for the first word choice and 72% correct for the second word choice. The distribution of errors was analyzed by place of articulation, manner of articulation, and speech-to-speech-ratio of the phoneme pairs and compared to articulation index predictions for speech intelligibility.

10:00–10:15 Break

10:15

4aPPb8. Tactual displays of consonant voicing to supplement speechreading. Hanfeng Yuan, Charlotte M. Reed, and Nat Durlach (Res. Lab. of Electron., MIT, Rm. 36-757, 77 Massachusetts Ave., Cambridge, MA 02139)

This research is concerned with the development of tactual displays of voicing to supplement speechreading in persons with profound hearing impairment. The voicing cue was based on the onset-time difference between amplitude envelopes derived from two different filtered bands of speech (a low-pass band at 350 Hz and a high-pass band at 3000 Hz). This envelope-onset-asynchrony cue was presented through a multifinger tactual stimulator. The amplitude envelopes of the low- and high-frequency bands were used to modulate the amplitude of a 50-Hz sinusoid presented to the thumb and a 250-Hz sinusoid presented to the index finger (respectively). Acoustic measurements of envelope-onset asynchrony in CVC syllables indicate that this timing difference provides a reliable and robust cue for voicing. Perceptual measurements were made of the ability to discriminate (1) temporal-onset order for sinusoidal signals (fixed in frequency but roving in amplitude and duration); and (2) eight pairs of consonant contrasts in CVC syllables under conditions of speechreading alone, tactual cue alone, and the combined condition. Relations among the results for the two perceptual tasks, as well as for the associated acoustic measurements, will be discussed. [Work supported by NIH/NIDCD.]

10:30

4aPPb9. Effects of pitch, level, and tactile cues on speech segregation. Rob Drullman and Adelbert W. Bronkhorst (TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands, drullman@tm.tno.nl)

Sentence intelligibility for interfering speech was investigated as a function of level difference, pitch difference, and presence of tactile support. A previous study by the present authors [*J. Acoust. Soc. Am.* **111**, 2432–2433 (2002)] had shown a small benefit of tactile support in the speech-reception threshold measured against a background of one to eight competing talkers. The present experiment focused on the effects of informational and energetic masking for one competing talker. Competing speech was obtained by manipulating the speech of the male target talker (different sentences). The PSOLA technique was used to increase the average pitch of competing speech by 2, 4, 8, or 12 semitones. Level differences between target and competing speech ranged from –16 to +4 dB. Tactile support (B&K 4810 shaker) was given to the index finger by presenting the temporal envelope of the low-pass-filtered speech (0–200 Hz). Sentences were presented diotically and the percentage of correctly perceived words was measured. Results show a significant overall increase in

intelligibility score from 71% to 77% due to tactile support. Performance improves monotonically with increasing pitch difference. Louder target speech generally helps perception, but results for level differences are considerably dependent on pitch differences.

10:45

4aPPb10. Phase effects in forward masking for normally hearing and hearing-impaired subjects. Brian C. J. Moore, Thomas Stainsby, and Esme Terasewicz (Dept. of Exp. Psych., Univ. of Cambridge, Downing St., Cambridge CB2 3EB, UK)

The forward masking produced by harmonic complex tones depends on the phases of the masker components; phases giving “peaky” waveforms on the basilar membrane result in less forward masking than phases giving less peaky waveforms. This difference has been attributed to the effects of peripheral compression and suppression, which depend on the active mechanism in the cochlea. Hence, no phase effect would be expected for subjects with moderate cochlear hearing loss. Growth-of-masking functions were measured for forward maskers containing the first 40 harmonics of a 100-Hz fundamental, with components added in cosine (C) or random (R) phase. The signal frequency was 1 or 2 kHz. For three normally hearing subjects, the R masker produced considerably more masking than the C masker, at high masker levels. Of four subjects with moderate cochlear hearing loss, one showed no effect of masker phase, but for the other three there was a significant effect for at least one signal frequency. It is suggested that peripheral compression and suppression play a role in the phase effect, but other factors, possibly depending on the operation of the efferent system, are involved. [Work supported by MRC (UK).]

11:00

4aPPb11. Role of short-term memory in loudness comparisons. Wolfgang Ellermeier (Dept. of Acoust., Fredrik Bajers Vej 7 B5, DK-9220 Aalborg Ost, Denmark) and Birgit Werner (Boston Univ., Boston, MA 02215)

In an earlier study of the auditory discrimination of time-varying noise bursts [Ellermeier and Schrödl, J. Acoust. Soc. Am. **108**, 2596 (2000)], listeners were found to place greater weight on the beginning and end of the sounds than on the middle portion. To investigate whether this outcome is due to primacy and recency effects in short-term memory which tend to be sensitive to manipulations of the inter-stimulus interval (ISI), the ISI separating the two noise bursts in a 2IFC task was varied systematically. Six participants performed loudness comparisons on 1-s samples of white noise randomly changing in level every 100 ms. In different blocks of trials, the two noise bursts to be compared were either separated by a 500-ms or a 2-s ISI. COSS analysis [Berg, J. Acoust. Soc. Am. **86**, 1743–1746 (1989)] of the overall loudness judgments revealed elevated weights for the beginning and end of the noises, as in the earlier study. These weighting patterns were largely unaffected by the manipulation of ISI, suggesting that the temporal weights found characterize loudness integration in general, and are not just due to idiosyncrasies of the timing used in the 2IFC procedure.

11:15

4aPPb12. Objective limits for four varieties of putative loudness adaptation. Ernest M. Weiler (Psychoacoustics Lab., CSD, CAHS, ML #394, Univ. of Cincinnati, Cincinnati, OH 45267-0394, ernest.weiler@uc.edu), Hongwei Dou, Joel S. Warm, and David E. Sandman (Univ. of Cincinnati, Cincinnati, OH 45267-0394)

The four varieties included three monaural techniques: (1) tone decay (TD); (2) simple adaptation (SA); (3) ipsilateral comparison paradigm (ICP); and the binaural (4) simultaneous dichotic loudness balances

(SDLB). “Loudness adaptation” indicates that over time, under some conditions, there is a perceived decrease in loudness, when the initial baseline stimulus is progressively assessed. The authors have found the following limits: (1) the classic TD occurred within about 30 dB of threshold for all values tested (250 to 8000 Hz). (2) Except near threshold, SA for the loudness of a continuous unmodulated tone was observed at or above 40 dB when the stimulus reached 6000 Hz, or more. (3) ICP adaptation, which depends on at least 5-s intensity modulation, was found at all values tested from 40 to 80 dB, and from 250 to 8000 Hz. It correlates significantly with TD, with suppression of transient evoked otoacoustic emissions, and at 8000 Hz with SA adaptation. (4) Binaural SDLB adaptation has been repeatedly found at 20 to 100 dB, at all frequencies tested. Factor analysis indicated that binaural SDLB has at least four component factors, including binaural interaction. Unless factors are separated, SDLB adaptation does not correlate with monaural adaptation.

11:30

4aPPb13. Cancelled harmonics—How high does the effect go? Matthew J. Goupell, Peter Xinya Zhang, and William M. Hartmann (Phys. and Astron., 4230 BPS Bldg., Michigan State Univ., East Lansing, MI 48824)

Demonstration Number 1 in the IPO-NIU-ASA collection of auditory demonstrations (compact disk) periodically cancels and reinserts a harmonic of a complex tone having a 200-Hz fundamental and 20 equal-amplitude harmonics. This procedure causes a listener to hear out the manipulated harmonic as a separate tone. In this way, the demonstration exposes harmonics 1 through 10. The following question arises: What is the highest harmonic that can be made audible, and what is responsible for the limitation? Listening experiments, using random harmonic phases, fundamental frequencies (f_0) from 50 to 2000 Hz, and a maximum harmonic frequency of 20 kHz, show that for high fundamental frequencies ($f_0 > 200$ Hz) the highest audible harmonic frequency is insensitive to f_0 and is only about 10 percent less than the highest audible sine frequency in quiet. For lower fundamental frequencies, the highest audible harmonic number tends to be insensitive to f_0 and is 50–70 for normal hearing listeners. In this region the highest audible harmonic number can be predicted from noise-masked threshold data, but with a large uncertainty. [Work supported by NIDCD.]

11:45

4aPPb14. Underwater loudness for pure tones: Duration effects. Edward A. Cudahy, Derek Schwaller, David Fothergill, and Keith Wolgemuth (Naval Submarine Medical Res. Lab., Box 900, Groton, CT 06349-5900)

The loudness of underwater pure tones was measured by loudness matching for pure tones from 100 to 16,000 Hz. The standard was a one second tone at 1000 Hz. The signal duration was varied from 20 milliseconds to 5 seconds. Subjects were instructed to match the loudness of the comparison tone at one of the test frequencies to the loudness of the standard tone. Loudness was measured at the threshold, the most comfortable loudness, and the maximum tolerable loudness. The intensity of the standard was varied randomly across the test series. The subjects were bareheaded U.S. Navy divers tested at a depth of 3 meters. All subjects had normal in-air hearing. Tones were presented to the right side of the subject from an array of underwater sound projectors. The sound pressure level was calibrated at the location of the subject's head with the subject absent. Loudness increased and threshold decreased as duration increased. The effect was greatest at the lowest and highest frequencies. The shape of the loudness contours across frequency and duration derived from these measurements are different from in-air measurements. [Research supported by ONR.]

Session 4aSA

Structural Acoustics and Vibration: General Topics in Structural Acoustics and Vibration

Kenneth D. Frampton, Chair

Department of Mechanical Engineering, Vanderbilt University, Box 1592, Station B, Nashville, Tennessee 37235-1592

Contributed Papers

9:00

4aSA1. Radiation efficiency of a fluid-loaded rib-stiffened finite cylindrical shell. Y. F. Hwang and W. K. Bonness (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804-0030)

Due to the difficulty of computing the free modes of heavily fluid-loaded structures, radiation efficiency of such structures has been frequently calculated from the forced vibratory responses, which are usually contaminated with the direct radiation of the force field. This paper discusses two methods of calculating the radiation efficiency of a fluid-loaded rib-stiffened finite cylindrical shell. Solutions in both methods are obtained from Fourier series expansion of fluid loading and rib dynamic stiffness in terms of the shell *in vacuo* modes. In the first method, the forced radiated power and the spatial mean-square vibratory velocity are calculated excluding the supersonic elements of the forcing function. In the second method, free modes of the fluid-loaded and rib-stiffened shell are calculated from the coupled equations of motion in terms of *in vacuo* modes as generalized coordinates. From this, the modal radiated power and the modal spatial mean-square velocity are calculated. Radiation efficiency is determined from the ratio between the radiated power and the product of ρc , surface area, and the spatial mean-square vibratory velocity. Calculated results from the two methods are discussed. Membrane waves' contributions to the radiation efficiency are also discussed. [Work supported by ONR, Code 333.]

9:15

4aSA2. Vibration of fluid-loaded hemi-prolate spheroidal shells. 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 hemi-prolate spheroidal shells of constant thickness were derived using Hamilton's principle. The shell is clamped at the equator and is excited by mechanical surface force fields. The shell theory used in this derivation includes shear deformations and rotatory inertias. The displacements and rotations were expanded in an infinite series of comparison functions. The shell is fluid-filled and is submerged in an infinite fluid medium. The external and internal fluid loading impedances were computed using expansions of prolate spheroidal wavefunctions in each domain. The dynamic response of the fluid-loaded shell was determined using an axisymmetric normal surface force as the excitation input. Numerical results were obtained for the driving and transfer mobilities for several shell thickness-to-length ratios ranging from 0.005 to 0.1, and for various shape parameters, " a ," ranging from an elongated hemi-spheroidal shell ($a=1.01$) to a hemispherical 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.]

9:30

4aSA3. Chaotic properties and eigenfrequency distribution of Semi-Stadium 2-D Fields. Mikio Tohyama, Yoh-ichi Fujisaka, and Kazuaki Yoshida (2665-1, Nakano-machi Hachioji-shi, Tokyo 192-0015, Japan, ctdeneuv@sin.cc.kogakuin.ac.jp)

This paper describes the eigenfrequency distribution of semi-stadium 2-D fields whose geometrical figure is formed by rectangular and hyper-circular segments. The sound ray propagation in the fields exhibits chaotic properties when order n of the hyper-circular parts defined by $x^n + y^n = r^n$ decreases to 2. The chaotic behavior can be estimated according to the degree of freedom of the gamma distribution that represents the eigenfrequency spacing statistics and the correlation dimensions of the chaotic structure. The experimental and simulation results of the 2-D-membrane and plate vibration analysis show that the degree of freedom of the gamma distribution changes from 1 to 2 including nonintegers as the boundary of the 2-D space changes from rectangular (n : infinity, "regular field") to stadium (n : 2, "irregular field"), which corresponds to the change in the correlation dimensions from 1 to 2 for the sound ray propagation in the field. The family of gamma distributions includes the Wigner distribution, which Lyon [J. Acoust. Soc. Am. **45**, 545–565 (1969)] assumed for irregularly shaped boundaries as the case where the degree of freedom is 2.

9:45

4aSA4. Development of a new spectral energy formulation based on structural driving point mobilities or impedances. Seungbo Kim and Raj Singh (Acoust. and Dynam. Lab., The Ohio State Univ., 206 W. 18th Ave., Columbus, OH 43210-1107, kim.873@osu.edu)

A new approximate method is proposed here that characterizes harmonic kinetic and potential energies of a system based on the driving point impedance or mobility. The proposed complex-valued formulation is compared with an existing, real-valued dynamic energy approximation method [J. Sound Vib. **217**, 351–386; **247**, 683–702]. The scope of this study is limited to the frequency domain analysis and harmonic excitation is applied to a linear time-invariant (LTI) system. Longitudinal and flexural motions of a finite beam along with some discrete system examples are employed to illustrate the proposed scheme. Our method is based on an alternate interpretation of the associated driving point transfer functions and it approximates total time-averaged harmonic kinetic and potential energies. Numerical results show that our method yields a more accurate estimate than the existing method. Further, the proposed scheme is found to be insensitive to the driving point measures as consistent results using either impedance or mobility formulations, unlike the existing method, are predicted. Our characterization method is suitable for a moderately damped system at low and mid-frequencies, like the existing method.

10:00

4aSA5. Assessment of a spectral energy formulation based on driving point structural mobilities or impedances. Seungbo Kim and Raj Singh (Acoust. and Dynam. Lab., The Ohio State Univ., 206 W. 18th Ave., Columbus, OH 43210-1107, kim.873@osu.edu)

To determine the contributions of parallel paths or interactions between structural components at any frequency, dynamic kinetic and potential energies are needed. However, an exact method to characterize an arbitrarily damped system in kinetic and potential energy forms is currently not available. One approximate characterization method examines time-averaged dynamic energies of subsystem(s) via the driving point transfer functions [J. Sound Vib. **217**, 351–386 and **247**, 683–702]. This concept, based on a real-valued energy formulation, is critically examined here. Linear time-invariant (LTI) system is assumed and harmonic excitation is applied to the system. Finite continuous (bar and beam) structures along with discrete system examples are utilized to illustrate concepts. Predictions show that the energy estimates deviate from the exact values, especially at higher frequencies and with higher damping. Further, the results from the impedance formulation differ from the ones from the mobility representation. The cause of discrepancies is explained by interpreting the transfer functions involved.

10:15

4aSA6. How can one tell if noisy frequency response data contain a modal contribution? Christopher W. Moloney and Jerry H. Ginsberg (Woodruff School of Mech. Eng., Georgia Inst. of Technol., 801 Ferst Dr., Atlanta, GA 30332-0405, gtg824d@prism.gatech.edu)

A key question that arises in some techniques for experimental modal analysis is whether coherent modal information is buried within a frequency band of complex frequency response data. This question is especially problematic if the data are contaminated by a substantial level of noise. Furthermore, some identification algorithms introduce small anomalies to the data in the course of processing. To explore techniques for resolving this issue frequency response data for a one-degree-of-freedom system is generated analytically and contaminated with various levels of white noise. Five statistical measures are computed for a sequence of bands that are obtained by forming overlapping halves of the previous frequency bands. The five measures are the expected value, mean-square, and variance of the frequency response in each band, as well as eigenanalysis of the autocorrelation matrix and wavelet decomposition. Graphical results are presented and discussed, and an analytical assessment method is proposed.

10:30

4aSA7. A comparison of the rational fraction polynomial method and the algorithm of mode isolation for fitting noisy data. Matthew S. Allen and Jerry H. Ginsberg (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

A multitude of methods exist for fitting linear, modal models to mechanical system response data. Many of these methods involve rearranging the modal equations of motion of the system so that a linear-least-squares solution is possible. Some methods require simplifying assumptions, such as light damping, while most popular methods are exact for clean data. Because the overall process is nonlinear, two different exact, linear-least-squares methods, can give different results for noisy data. This work compares two methods which are exact for linear, viscous, state-space (or nonproportionally damped) systems, the well known Rational Fraction Polynomial Method (RFP) and the Algorithm of Mode Isolation (AMI). It is shown that while RFP performs exceptionally for clean data, it is much less robust than AMI for noisy data. The performance of both algorithms is compared when applied to noise contaminated analytical data for a multi-

degree of freedom frame structure. The frame can be tuned so that high damping ratios and heavy modal coupling are present. [Work supported under a National Science Foundation Graduate Research Fellowship.]

10:45

4aSA8. Experiments and interpretation of surface waves acoustically generated on spherically endcapped cylindrical shells. A. Claude Ahyi, Hui Cao, P. K. Raju (Dept. of Mech. Eng., Auburn Univ., Auburn, AL 36849), and H. Überall (Catholic Univ. of America, Washington, DC 20064)

The impact of an acoustic pulse on a submerged elastic shell (which we shall assume evacuated) generates three types of circumferentially propagating surface waves: those that are analogous to plate waves of type $A_0(A_1, A_2, \dots)$ and $S_0(S_1, S_2, \dots)$, and a Scholte–Stoneley wave of type A that propagates in the surrounding fluid. A computer program devised by Hui Cao *et al.* renders visualizations of the generation and propagation of circumferential pulses on spherically endcapped shells, visualized by the re-radiation into the surrounding fluid of A , A_0 , and S_0 waves in sequential pictures. These surface waves are generated experimentally by us in the laboratory from an ultrashort-pulse source, at axial incidence on a hemispherically endcapped glass tube. Sequential visualizations of the re-radiated pulses are obtained by using the Schardin–Cranz Schlieren method. These observations lead to an experimental measure of the group velocity dispersion curves of the surface waves.

11:00

4aSA9. Phonon localization methods through time and space: Experimental SAW phonon aspects on a cylindrical shell. Loic Martinez (ECIME IUP GE, 5 mail Gay-Lussac, F 95 031 Neuville sur Oise Cedex, France), Bruno Morvan, and Jean Louis Izbicki (Universit du Havre, 76610 Le Havre, France)

A transient experimental study of Surface Acoustic Wave (SAW) propagation along a 1D medium as a function of time leads to a 2D space–time signal collection. Previous studies have shown that 3D space–wave number–frequency representation $S(x, k, f)$ allows the characterization of the space transient aspects of SAW generation [L. Martinez, J. Acoust. Soc. Am. **105**, 952 (1999)]. In order to analyze the time transient aspect of phonon propagation, the 3D time–wave Number-Frequency representation $Z(t, k, f)$ is proposed. The $Z(t, k, f)$ matrix is obtained by short time Fourier transforming each time signal of the 2D time–wave number representation. The $Z(t, k, f)$ representation is used to experimentally investigate SAW generation and propagation around a cylindrical shell (the relative thickness is equal to 0.03) surrounded by water and excited by a pulse (0.1 s duration). For the air filled shell, the $Z(t, k, f = \text{const})$ slices show the time sequence of the continuous flow of incident phonons striking the insonified side of the shell and their reflection or conversion in SAW phonons. The complex frequencies, direction of propagation and the time origin of all the phonons are identified. For the first time, the elastic Franz waves due to diffraction around the shell are also observed. For the water-filled shell the $Z(t, k, f = \text{const})$ slices point out two new major features. On the one hand, phonons trapped inside the fluid column of the shell are clearly identified. On the other hand, when the initial incidence angle enables the generation of a SAW, the SAW is generated at each internal striking point. A precise experimental ray model is deduced from these results, allowing us to link this transient experimental study with the theoretical results based on a steady state approach.

Session 4aSC

Speech Communication: Acoustical and Perceptual Characteristics of Special Speech Registers

Sarah H. Ferguson, Chair

*Department of Speech, Language, Hearing Sciences and Disorders, University of Kansas, 3001 Dole Center,
1000 Sunnyside Avenue, Lawrence, Kansas 66045*

Chair's Introduction—8:30

Invited Papers

8:35

4aSC1. Production and perception of clear speech. Ann R. Bradlow (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208, abradlow@northwestern.edu)

When a talker believes that the listener is likely to have speech perception difficulties due to a hearing loss, background noise, or a different native language, she or he will typically adopt a clear speaking style. Previous research has established that, with a simple set of instructions to the talker, "clear speech" can be produced by most talkers under laboratory recording conditions. Furthermore, there is reliable evidence that adult listeners with either impaired or normal hearing typically find clear speech more intelligible than conversational speech. Since clear speech production involves listener-oriented articulatory adjustments, a careful examination of the acoustic-phonetic and perceptual consequences of the conversational-to-clear speech transformation can serve as an effective window into talker- and listener-related forces in speech communication. Furthermore, clear speech research has considerable potential for the development of speech enhancement techniques. After reviewing previous and current work on the acoustic properties of clear versus conversational speech, this talk will present recent data from a cross-linguistic study of vowel production in clear speech and a cross-population study of clear speech perception. Findings from these studies contribute to an evolving view of clear speech production and perception as reflecting both universal, auditory and language-specific, phonological contrast enhancement features.

9:00

4aSC2. Elderspeak. Susan Kemper (3090 Dole, Gerontology Ctr., 1000 Sunnyside, Univ. of Kansas, Lawrence, KS 66045)

Elderspeak has been assumed to be an accommodation to the perceived communication needs of older adults as it involves a slow rate of speaking, simplified syntax, vocabulary restrictions, and exaggerated prosody. It also is judged to be patronizing and disrespectful in that its use presumes that the older adult is cognitively impaired. Using a controlled referential communication task, my colleagues and I have investigated elderspeak addressed by young adults to older adults. Our studies suggest that the speech register is composed of two sets of parameters. One set of parameters is linked to the perception that the older listener is cognitively impaired; these parameters affect how much information is conveyed and include semantic elaborations such as expansions and repetitions of previous map directions. The other set of parameters includes fluency, prosody, and grammar. The modifications to fluency and prosody do not appear to benefit older listeners. In these studies, the young adults' use of elderspeak did improve the performance of the older listeners on the referential communication task. The use of elderspeak by the young partners appeared to trigger older adults' perceptions of themselves as cognitively impaired, consistent with the "communicative predicament of aging" model of Ryan *et al.* (1986).

9:25

4aSC3. An acoustic comparison of two women's infant- and adult-directed speech. Jean Andruski and Shiri Katz-Gershon (Audiol. & Speech-Lang. Pathol., Wayne State Univ., 581 Manogian Hall, Detroit, MI 48201)

In addition to having prosodic characteristics that are attractive to infant listeners, infant-directed (ID) speech shares certain characteristics of adult-directed (AD) clear speech, such as increased acoustic distance between vowels, that might be expected to make ID speech easier for adults to perceive in noise than AD conversational speech. However, perceptual tests of two women's ID productions by Andruski and Bessega [J. Acoust. Soc. Am. **112**, 2355] showed that is not always the case. In a word identification task that compared ID speech with AD clear and conversational speech, one speaker's ID productions were less well-identified than AD clear speech, but better identified than AD conversational speech. For the second woman, ID speech was the least accurately identified of the three speech registers. For both speakers, hard words (infrequent words with many lexical neighbors) were also at an increased disadvantage relative to easy words (frequent words with few lexical neighbors) in speech registers that were less accurately perceived. This study will compare several acoustic properties of these women's productions, including pitch and formant-frequency characteristics. Results of the acoustic analyses will be examined with the original perceptual results to suggest reasons for differences in listener's accuracy in identifying these two women's ID speech in noise.

4aSC4. Predicting children's hyperarticulate speech during human-computer error resolution. Sharon Oviatt, Rachel Coulston (Dept. of Computer Sci. & Eng., Oregon Health & Sci. Univ., Beaverton, OR 97006, oviatt@cse.ogi.edu), and Courtney Darves (Psych. Dept., Univ. of Oregon)

When speaking to interactive systems, people sometimes *hyperarticulate*—or adopt a clarified form of speech that has been associated with increased recognition errors. The goal of the present study was to provide a comprehensive assessment of the type and magnitude of linguistic adaptations in children's speech during human-computer error resolution, and to compare these adaptations with those typical of adult hyperarticulation. A study was conducted in which twenty-four 7- to 10-year-old children interacted with a simulated conversational system, which permitted a comparison of their verbatim repetitions immediately before and after system recognition errors. Matched original-repeat utterance pairs then were analyzed for acoustic, prosodic, and phonological adaptations. Like adult speech, the primary hyperarticulate changes in children's speech included durational phenomena such as lengthening of pauses and the speech segment, and a more deliberate, hyper-clear articulatory style. However, children's speech also displayed large increases in amplitude that are not typical of adult hyperarticulation, as well as substantially larger magnitude adaptations than those observed in adult speech. These results corroborate and generalize the Computer-elicited Hyperarticulate Adaptation Model, and have implications for improved error handling in next-generation spoken language and multimodal systems. [Work supported by NSF Grant No. IIS-0117868.]

10:15–10:30 Break

Contributed Papers

10:30

4aSC5. Does an infant-directed speaking style aid in the separation of different streams of speech? Rochelle S. Newman (Dept. of Hearing & Speech Sci. & Prog. in Neurosci. & Cognit. Sci., Univ. of Maryland, College Park, MD 20742, rnewman@hesp.umd.edu), Tammy Weppelman, and Isma Hussain (Dept. of Psych., Univ. of Iowa)

This work explores whether an infant-directed speaking style (IDS) may be easier to separate from background speech noise than is adult-directed speech (ADS). Many of the acoustic cues found in infant-directed speech are similar to ones shown to be important in adult stream segregation (such as differences in voice pitch and pitch variability). In addition, using an infant-directed speech style may serve to make the talker's voice more dissimilar from any background speech (which is likely to be adult-directed). We explored this issue in three different ways: by examining adults' ability to separate two different streams of speech varying in register, by examining whether adult speakers use IDS to a greater extent when in the context of noise, and by examining whether infants' preference for IDS over ADS might be greater in the context of background noise. Implications of these three lines of work will be discussed. [Work supported by NSF and NICHD.]

10:45

4aSC6. Acoustic characteristics of listener-constrained speech. Simone Ashby and Fred Cummins (Dept. of Computer Sci., Univ. College Dublin, Belfield, Dublin 4, Ireland, simone.ashby@ucd.ie)

Relatively little is known about the acoustical modifications speakers employ to meet the various constraints—auditory, linguistic and otherwise—of their listeners. Similarly, the manner by which perceived listener constraints interact with speakers' adoption of specialized speech registers is poorly understood. Lindblom's Hyper & Hypo (H&H) theory offers a framework for examining the relationship between speech production and output-oriented goals for communication, suggesting that under certain circumstances speakers may attempt to minimize phonetic ambiguity by employing a "hyperarticulated" speaking style (Lindblom, 1990). It remains unclear, however, what the acoustic correlates of hyperarticulated speech are, and how, if at all, we might expect phonetic properties to change respective to different listener-constrained conditions. This paper is part of a preliminary investigation concerned with comparing the prosodic characteristics of speech produced across a range of listener constraints. Analyses are drawn from a corpus of read hyperarticulated speech data comprising eight adult, female speakers of English. Specialized registers include speech to foreigners, infant-directed speech, speech produced under noisy conditions, and human-machine interaction. The authors gratefully acknowledge financial support of the Irish Higher Education Authority, allocated to Fred Cummins for collaborative work with Media Lab Europe.

11:00

4aSC7. Very loud speech over simulated environmental noise tends to have a spectral peak in the F1 region. Sten Ternstrom, Mikael Bohman (Speech Music & Hearing, Kungliga Tekniska Hogskolan, Drottning Kristinas Vaeg 31, SE-100 44 Stockholm, Sweden, sten@speech.kth.se), and Maria Sodersten (Huddinge Univ. Hospital, SE-141 86 Stockholm, Sweden)

In some professions, workplace noise appears to be a hazard to the voice, if not to hearing. Several studies have shown that teachers and sports instructors, for example, are more prone to voice problems than average, prompting research on loud voice. Since on-location recordings are in many ways impractical, the running speech of 23 untrained speaker subjects (12 female, 11 male) was instead recorded in several types of loud noise that was presented over high-quality loudspeakers. Using adaptive cancellation techniques, the noise was then removed from the recordings, thus exposing the strained voices for analysis. The experiment produced a large body of data, only one aspect of which is reported here. In most subjects, the vowel spectrum as a function of voice SPL showed the expected behavior for low to moderate efforts, but developed a very pronounced peak in the F1 region at the highest efforts. This peak can be ascribed to the concerted action of several acoustic mechanisms, including source waveform asymmetry, F1 approximating one of the lower partials, and increased formant Q values due to a longer closed phase. [Work supported by the Swedish Council for Working Life and Social Research, Contract No. 2001-0341.]

11:15

4aSC8. Effect of speaking styles on the relevance of the Speech Transmission Index in the presence of reverberation. Sander J. van Wijngaarden and Tammo Houtgast (TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands)

The Speech Transmission Index (STI) predicts speech intelligibility under various speech degrading conditions, including noise and reverberation. Measurements of sentence intelligibility in noise and reverberation, for three different talkers adopting different speaking styles, indicate that the STI is sometimes inaccurate in the presence of reverberation. For a trained talker speaking clearly, the STI predictions are accurate, but for conversational speech by untrained talkers, the effect of reverberation is underestimated. By using a wider range of modulation frequencies than prescribed for the standardized STI calculation (up to 31.5 Hz instead of 12.5 Hz), more accurate predictions are obtained for conversational speech. This can be understood by comparing the envelope spectrum between talkers and speaking styles: conversational speech tends to spread

more of the energy in its envelope spectrum to higher modulation frequencies. Since higher modulation frequencies are more susceptible to reverberation, conversational speech tends to be affected more by reverberation than clear speech. To investigate the range of between-talker variations,

envelope spectra were calculated for a population of 134 talkers. Upper boundaries of the STI modulation frequency range of 12.5 and 31.5 Hz seem approximately appropriate for the 5th and 95th percentile of this population.

THURSDAY MORNING, 1 MAY 2003

ROOM 203, 10:30 A.M. TO 12:00 NOON

Session 4aSP

Signal Processing in Acoustics: Arrays and Beamforming

Alan W. Meyer, Chair

Lawrence Livermore National Laboratory, P.O. Box 808, L-154, Livermore, California 94551

Contributed Papers

10:30

4aSP1. High resolution beamforming for small aperture arrays. Chris Clark, Tom Null (Miltec Res. and Technol., Inc., NCPA, 1 Coliseum Dr., Univ. of MS 38677), and Ronald A. Wagstaff (Univ. of Mississippi, MS 38677)

Achieving fine resolution bearing estimates for multiple sources using acoustic arrays with small apertures, in number of wavelengths, is a difficult challenge. It requires both large signal-to-noise ratio (SNR) gains and very narrow beam responses. High resolution beamforming for small aperture arrays is accomplished by exploiting acoustical fluctuations. Acoustical fluctuations in the atmosphere are caused by wind turbulence along the propagation path, air turbulence at the sensor, source/receiver motion, unsteady source level, and fine scale temperature variations. Similar environmental and source dependent phenomena cause fluctuations in other propagation media, e.g., undersea, optics, infrared. Amplitude fluctuations are exploited to deconvolve the beam response functions from the beamformed data of small arrays to achieve high spatial resolution, i.e., fine bearing resolution, and substantial SNR gain. Results are presented for a six microphone low-frequency array with an aperture of less than three wavelengths. [Work supported by U.S. Army Armament Research Development and Engineering Center.]

10:45

4aSP2. Characterization of the left/right bearing ambiguities for passive line arrays. Frank A. Boyle and David E. Grant (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

The symmetry of a straight line underwater acoustic array produces a left/right bearing ambiguity that makes it impossible to determine which side of the array a source is located on. A simple technique for resolving left/right ambiguities is to break the symmetry by curving the array. Under certain common conditions, however, significantly curved arrays still possess left/right ambiguities that depend on the source's range and the beamformer's focal length. These ambiguities can be modeled, given a source and array configuration, to produce a prediction of source localization capability. A map can be generated that indicates the spatial regions where a source is likely to be localized on the incorrect side of the array. The presentation will include a description of how the ambiguities are modeled, and simulated results. Considerations on how curved arrays can be configured to minimize ambiguities will be discussed.

11:00

4aSP3. Multiple speech signal enhancement using a microphone array. Heather E. Ewalt and Michael T. Johnson (Dept. of Elec. and Computer Eng., Marquette Univ., Milwaukee, WI 53223, heather.ewalt@mu.edu)

The ability to extract and enhance a primary speech signal from an environment with multiple speakers is an important issue. While methods exist for a variety of beamforming techniques [M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications* (Springer, New York, 2001)] as well as for multi-source filtering in stationary noise [H. Saruwatari *et al.*, "Speech Enhancement Using Nonlinear Microphone Array With Noise Adaptive Complementary Beamforming," Proc. of IEEE ICASSP, pp. 1049–1052 (2000)], the theory has yet to be developed for integrating spatial filtering with additional enhancement methods to deal with non-stationary interference from interfering talkers. This paper presents a novel method for incorporating multiple parallel beamformers with traditional speech enhancement algorithms, particularly the Wiener filter and spectral subtraction. By iteratively improving the spectral magnitude estimates of each speech source, substantial improvement in overall signal separation can be obtained. The performance of the algorithm is illustrated using a simulated multiple speaker environment with resulting SNR and sSNR plots. [Work supported by the GAANN Fellowship.]

11:15

4aSP4. Beamforming for a microphone array embedded in asymmetrically shaped objects. Philippe Moquin, Stéphane Dedieu (Mitel Networks, 350 Legget Dr., Kanata, ON K2K 2W7, Canada), and Rafik Goubran (Carleton Univ., Ottawa, ON K1S 5B6, Canada)

Broadband frequency invariant beamforming for circular arrays or linear arrays are quite common but not when they are embedded in a diffracting structure. Meyer [J. Acoust. Soc. Am. **109**, 185–193 (2001)] describes arrays embedded in a diffracting sphere, and provides an analytical solution for the wave equation in acoustics. For arrays of simple shape like circular rings embedded in a more complex shape one must make use of numerical methods (e.g., boundary element methods). Microphone arrays in shapes that are not symmetric or axisymmetric can also be solved this way but result in very asymmetrical beams. One example of such an obstacle is a telephone incorporating a microphone array. This presentation will show results from simulations and measurements of a six-microphone array. A design approach to obtain reasonably well behaved beams relies on constrained optimization, with a constraint build using a set of vectors containing the sensor signal for acoustic waves with specific directions of arrival. [Work supported in part by Carleton University.]

4a THU. AM

4aSP5. Inferring array geometry from multiple sources of opportunity. Frank A. Boyle and David E. Grant (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

A beamformer's performance is sensitive to the accuracy of the element location estimates. Array element localization is a reasonably simple task when a source is present with a known location. The problem is more difficult with a source whose position is unknown. Techniques commonly require assumptions to be made regarding the source's range. A method for determining the array's geometry from multiple sources of opportunity has been formulated. Array geometry is determined from multiple observations of parallax from each source, for a collection of hydrophone pairs in the array. The array's geometry is built up from the combined orientations of hydrophone pairs. In its current form, the technique requires accurate knowledge of hydrophone pair separation. The presentation will include an application of the method to simulated array data and a discussion of current limitations.

4aSP6. Beamforming using indefinite term for arbitrarily located microphones. Hirofumi Nakajima, Yamanaka Takaaki, and Hiroshi Nakagawa (Nittobo Acoust. Eng. Co., Ltd., 1-13-12, Midori, Sumida-ku, Tokyo 130-0021, Japan, nakajima@noe.co.jp)

The design of filter coefficients for a beamforming algorithm constructed of finite impulse response filters and adders was investigated. Designing filter coefficients can be represented as a linear inverse problem, and using the indefinite term of the linear equation, a filter having the smallest sidelobe beam pattern based on the least-square estimation criterion can be designed. This design method is represented by a closed form equation using singular value decomposition and a pseudo-inverse matrix, so it does not need parameter adjustment, as do adaptive methods. It can be used for arbitrarily located microphones; it was applied to a sound visualization system using a spherical microphone array (SMA). At the meeting, this design method and its directivity performance when applied to a SMA will be described.

THURSDAY MORNING, 1 MAY 2003

ROOM 102, 8:00 A.M. TO 12:00 NOON

Session 4aUW

Underwater Acoustics and Acoustical Oceanography: High Frequency Sediment Acoustics

Eric I. Thorsos, Chair

Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105-5598

Chair's Introduction—8:00

Invited Papers

8:05

4aUW1. Underwater sand acoustics: A perspective derived from the sediment acoustics experiment (SAX99). Kevin L. Williams, Eric I. Thorsos, Darrell R. Jackson, Dajun Tang, and Steve G. Kargl (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St, Seattle, WA 98105-6698, williams@apl.washington.edu)

The sediment acoustics experiment (SAX99) included investigations of the following three questions. What are the dominant mechanisms responsible for backscattering from sand sediment? What are the dominant mechanisms responsible for subcritical penetration into sand? What are the appropriate constitutive equations for sand? In this paper a summary is presented of APL-UW SAX99 experiments and data/model comparisons relevant to each question. Perspectives are also given on some of the issues that remain or arose during SAX99 and the associated analyses. In general, these issues are tied to the frequency dependencies seen in the data but not fully captured by present models. For backscattering the issue is that as the frequency increases the measured backscattering strength does not follow predictions based on surface roughness scattering models. In the case of penetration it is a frequency cutoff effect seen in SAX99 buried array data but seemingly violated in the detection of buried objects near the SAX99 site. Regarding the constitutive equations, it is the frequency dependence of the attenuation above 50 kHz. Recent experiments will be described that have been motivated by these issues. Finally, research proposed as part of a follow-on sediment acoustics experiment (SAX04) will be outlined. [Work supported by ONR.]

8:25

4aUW2. Time domain modeling of seafloor scatter at low grazing angle using the small slope approximation. Eric Pouliquen, Lucie Pautet, and Gaetano Canepa (Saclant Undersea Res. Ctr., Viale San Bartolomeo, 400, 19138 La Spezia, Italy, pouliq@saclantc.nato.int)

Sea surface and bottom scattering components are usually quantified only in terms of scattering strength (SS). The SS corresponds to an ensemble averaged plane wave intensity scattered from a unit surface at a unit distance. However for non-steady-state cases when the acoustic footprint approaches the size of the acoustic wavelength, higher moment statistics and probability distribution functions (PDF) provide additional and essential information for detection and classification purposes. As a step toward higher moment scattering prediction the time domain model BORIS [Pouliquen *et al.*, J. Acoust. Soc. Am. **105** (1999)] has been extended to low grazing angles. It uses the fourth-order small slope approximation (SSA-4) and the small perturbation theory for interface and volume scattering, respectively. This bistatic 3-D model accounts for the full sensing geometry and sonar properties. It uses stochastic realizations of the boundaries and volumes with controlled statistics. BORIS-SSA offers numerous possibilities for practical applications. In particular it provides an objective assessment of detection and classification performances for most current sonar systems (conventional and synthetic aperture sonars) operating in realistic and complicated environments (cluttered or patchy). The model principles and simulation results will be presented and illustrated with real sonar data.

4aUW3. Tempo and scale of biogenic effects on high-frequency acoustic propagation near the marine sediment–water interface in shallow water. Peter Jumars (Darling Marine Ctr., Univ. of Maine, 193 Clark's Cove Rd., Walpole, ME 04573)

Organisms have natural scales, such as lifetimes, body sizes, frequencies of movement to new locations, and residence times of material in digestive systems, and each scale has potential implications for acoustic effects. The effects of groups of organisms, like organisms themselves, aggregate in space and time. This review, including an assortment of unpublished information, examines examples of such aggregations, many of them documented acoustically. Light synchronizes many activities. Macroscopic animals forage primarily under cover of darkness. This phasing applies both to animals that extend appendages above the sediment–water interface and to animals that leave the seabed at night. Whereas their bottom-modifying activities are concentrated in nocturnal or crepuscular fashion, the bottom-modifying activities of the visual feeders follow a different phasing and often dominate the rate of change in acoustic backscatter from the interface. Light also acts through its effects on primary production, often concentrated in a very thin surficial layer atop the seabed. The supersaturation of oxygen does, and microbubble nucleation may, result. Where tidal velocities are large, light-set patterns are often tidally modulated. Activities of animals living below the seabed, however, remain a mystery, whose primary hope for solution is acoustic. [Work supported by ONR and DEPSCoR.]

9:05

4aUW4. The evolution of rippled seafloor topography with acoustic implications. Michael D. Richardson (Marine Geosciences Div., Naval Res. Lab., Stennis Space Center, MS 39529-5004) and Peter Traykovski (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Rippled seafloors are often responsible for anisotropic patterns of acoustic backscattering and allow penetration of high-frequency energy into the seafloor below the critical angle. Both natural and manipulative experiments conducted during the Sediment Acoustic Experiment (SAX99) demonstrate the importance of understanding the temporal evolution and characterizing the spatial statistics of naturally occurring ripple fields for prediction of sonar performance and the detection of buried targets. Current and wave-induced sand ripples evolve in a more or less predictable pattern. Numerous empirical and semi-empirical predictive models, based on well-established principles of sediment transport, allow prediction of ripple wavelength, height, and shape. Degradation of sand ripple fields, especially by biological processes such as feeding, burrowing, and emergence is less known and has not been modeled. The temporal evolution of rippled topography measured with sector scanning sonar in high-energy environments is presented. These high-fidelity and nearly continuous observations coupled with measurements of bottom currents and near-bottom wave-induced orbital motion provide improved insights and new models of the evolution of rippled seafloor topography. In low-energy environments (SAX99 and the proposed SAX04) the longer times between storms allow characterization of rates of biological processes which destroy ripple structure and create isotropic roughness. [Work supported by ONR.]

9:25

4aUW5. Time and space scales of bedform evolution in sandy nearshore and inner shelf environments. Alex E. Hay (Dept. of Oceanogr., Dalhousie Univ., Halifax, NS B3H 4J1, Canada)

Recent experiments in wave-forced nearshore and inner shelf environments have demonstrated through acoustic and optical imaging techniques a remarkable range of bedform patterns in mobile sandy sediments at sub-10 m horizontal scales. The observed bedstates differ both in the level of complexity of their spatial patterns, and in the time scales of their response to changes in hydrodynamic forcing conditions. Yet, the (re)occurrence of these different bed states at a given location can be quite repeatable and, once formed, the migration of the bedforms characterizing a particular state can in some instances be predicted reasonably well from the local hydrodynamics. These results provide a basis against which the present capability for predicting bedstate evolution in wave-forced mobile bed environments can be assessed. It is argued that, with the possible exception of a few simple cases, current models are able to predict neither the level of complexity in the observed characteristic patterns, nor the time evolution of the bed given the prior history of the forcing. Nevertheless, the repeatability of bedstate occurrence, and the reasonably good predictions of migration velocity and shear stress at the bed, together suggest that such a predictive capability may be achievable in the not-too-distant future.

10:00–10:15 Break

Contributed Papers

9:45

4aUW6. A comparison of models for the reflection loss of acoustic waves from a smooth water/sediment interface. Marcia J. Isakson and Andrew Worley (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

Reflection-loss measurements have been inverted using a composite poro-elastic model to produce realistic sediment properties. This type of inversion could provide a noninvasive method of characterizing sediments. However, it is not well understood how measurement variability affects the outcome of the inversion. When measurement variability is taken into account, other models may be able to predict reasonable sediment properties. In this study, several models will be considered for their ability to produce realistic sediment parameters and their fit to the data. Models considered will be the visco-elastic model, the Biot/Stoll model, the effective density fluid model, the Buckingham model, and the composite Biot squirt flow/shear model. Each model will be inverted using a

simulated annealing algorithm with OASR as the forward model. Special emphasis will be placed on inverting within the statistical nature of the data. [Work supported by ONR, Ocean Acoustics.]

10:15

4aUW7. Buried target detection with a synthetic aperture sonar. John Piper (Coastal Systems Station, Code R21, Panama City, FL 32407-7001, piperje@ncsc.navy.mil)

Synthetic aperture sonar (SAS) systems can provide the high resolution and high signal-to-noise ratios that are important in detecting buried targets. Previous experiments with the Coastal Systems Station SAS have shown good results for above critical grazing angle targets and mixed results for below critical angle targets. Interest in characterizing the environmental conditions, which can allow detection at sub-critical angles, has prompted additional SAS testing. This paper describes a buried target experiment that included two 1.5-meter by 0.5-meter cylindrical targets,

one 35-cm silicon-fluid-filled sphere, and two 58-cm air-filled spheres. These targets were buried under sandy sediment in 17 meters of water approximately 1 mile offshore from Panama City Beach, Florida. Above critical angle and below critical angle runs were made during November 2001 and June 2002. Environmental conditions and the seafloor ripple structure were reported by divers and, during the 2002 experiment, measured by the Applied Physics Laboratory/University of Washington with their second-generation *in situ* measurement of porosity (IMP2) instrumentation.

10:30

4aUW8. Measurements of subcritical grazing angle detection of targets buried under a rippled sand interface. J. L. Lopes, C. L. Nesbitt, R. Lim (Coastal Systems Station, Code R21, 6703 W. Hwy. 98, Panama City, FL 32407-7001), D. Tang, K. L. Williams, and E. I. Thorsos (Univ. of Washington, Seattle, WA 98105-6698)

A series of controlled measurements were conducted to investigate shallow grazing angle acoustic detection of targets buried in sand having a rippled sediment-water interface. The measurements were performed in a 13.7-m deep, 110-m long, 80-m-wide test-pool with a 1.5-m layer of sand on the bottom. A silicone oil filled target sphere was buried under a rippled interface with contours formed by scraping the sand with a machined rake moved along a guide frame. Broadband, broad beam transducers were placed onto the shaft of a tilting motor. The transducers and tilting motor were attached to an elevated rail that enabled this assembly to be translated horizontally, permitting acquired data to be processed using synthetic aperture sonar techniques. Acoustic backscatter data were acquired at subcritical grazing angles in the frequency range of 10 to 50 kHz for various ripple wavelengths and heights. For each bottom configuration, the ripple profile over the buried target was measured using the In-situ Measurement of Porosity 2 (IMP2) system. Measurement results are presented that illustrate target detection via ripple scattering. The characteristics of the target return are found to depend sensitively on the ripple height and wavelength. [Work supported by ONR.]

10:45

4aUW9. Model/data comparisons of subcritical grazing angle detection of targets buried under a rippled sand interface. Kevin Williams, Eric Thorsos, Dajun Tang (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698, williams@apl.washington.edu), Raymond Lim, and Joe Lopes (Coastal Systems Station, Code R21, Panama City, FL 32407-7001)

Subcritical grazing angle target detection results for a sphere buried in Coastal System Station's Facility 383 pond are compared to predictions from two different models. Both models use perturbation theory to calculate the transmission across the rough interface due to ripple scattering. One is a simple sonar equation model that uses a first-order expression [Eq. (16), Jackson *et al.*, IEEE J. Ocean. Eng. **27**, 346-361 (2002)] to calculate the penetration, taking into account the measured ripple profile and the target strength estimate for the buried sphere. The other model uses the penetrating field to calculate target backscattering via a T-matrix formalism adapted to account for the specified sinusoidal roughness. These models have been combined with a reverberation model to predict detection SNR. The reverberation model uses a power law fit to the sand/water interface roughness spectrum determined from measured roughness profiles. [Work supported by ONR.]

11:00

4aUW10. Exploiting sediment acoustics properties for subcritical buried target detection using a synthetic aperture sonar aboard the AUV Reliant. Kerry W. Commander, Jose E. Fernandez, John E. Piper, and John S. Stroud (Coastal Systems Station, Code R21, Panama City, FL 32407-7001, commanderkw@nsc.navy.mil)

The Coastal Systems Station High Frequency/ Low Frequency Synthetic Aperture Sonar (HF/LFSAS) was recently integrated into a Bluefin Robotics, Inc. Odyssey III Family autonomous underwater vehicle (AUV). This vehicle, designated Reliant, has been used in several at-sea experiments to determine the feasibility of using bottom roughness to enhance detection of buried targets at subcritical grazing angles. Previously, the

HF/LFSAS was deployed on a towed 21-in. vehicle and made some remarkable buried target detections at subcritical angles. However, due to the complexities associated with using a towed platform, only a limited amount of data was obtained during these exercises. This sparse data set made it difficult to quantify the enhancing effects of bottom roughness on buried target detection. An extensive run matrix for buried target detection was completed this year using the Reliant AUV. The primary variables in the experiments were ripple orientation, standoff range, and grazing angle. Results from these sea tests are presented and compared to theoretical models.

11:15

4aUW11. Variation of sea-bed backscattering strength due to Bragg scattering. K. W. Commander, J. L. Lopes, and R. Lim (Coastal Systems Station, Code R21, 6703 W. Hwy. 98, Panama City, FL 32407-7001)

Bragg scattering from a rippled sea-bed may lead to reduced SNRs from proud targets, which adversely affects sonar performance. To quantify this effect, acoustic backscatter from a rippled bottom is investigated in the frequency range of 2-10 kHz by conducting a laboratory-type experiment to measure the reverberation levels from a bottom with a rippled interface. The experiment was performed in a 13.7-m deep, 110-m long, 80-m wide test pool that has 1.5 m of sand covering the bottom. In this experiment a calibrated parametric sonar was attached to a tower that was fitted with horizontal pan and vertical tilt motors. The rippled bottom was artificially formed with the aid of a sand scraper that consists of a frame and a rake that glides along the frame [Lopes *et al.*, "Shallow Grazing Angle Sonar Detection of Targets Buried Under a Rippled Sand Interface," Oceans 2002 MTS/IEEE, pp. 461-467]. The bottom backscatter measurements are compared to predictions of a model in which the interface roughness is represented by a Gaussian spectrum centered on the imposed ripple frequency band.

11:30

4aUW12. Mechanical loading of a spherical hydrophone embedded in a sediment. Steven G. Kargl (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

The free-field calibration of the response of a hydrophone is typically performed in water and it is often assumed that the hydrophone is insensitive to loading when immersed in other materials. Hydrophones are routinely deployed in sandy sediments to measure the penetration of an acoustic field from a sonar. Models of the mechanical response of a spherical hydrophone will be discussed. Space is split into three concentric regions and an infinite external region. The innermost region is a vacuum, the inner spherical shell is an active element (e.g., a piezoelectric ceramic), and the outer shell is an impedance matching material. The active element is modeled as a homogenous elastic material such that its piezoelectric properties are neglected. The external medium can be either a fluid, elastic, or a poroelastic medium. The mechanical impedance shows that the loading causes a shift in the resonance frequency of the hydrophone as well as a change in the Q of the resonance. The sensitivity of the hydrophone will also be discussed. [Work supported by ONR.]

11:45

4aUW13. Measurements of bottom reflection loss in the East China Sea. Jee Woong Choi and Peter H. Dahl (Appl. Phys. Lab., Univ. of Washington, Seattle, WA)

Measurements of bottom reflection loss were conducted in 105 m of water in the East China Sea (ECS) at location 29 39N, 126 49E as part of the Asian Seas International Acoustic Experiment (ASIAEX), in the spring of 2001. Sources in the frequency range 2-20 kHz were located at depths 25 m and 50 m, and receivers were deployed at a range of 500 m and at depths 26 m and 52 m. Bottom reflection loss measurements are obtained as a function of frequency and, to a limited extent, grazing angle. Experimental values are compared with model calculations based on independently measured geoaoustic and bottom roughness parameters. The frequency dependence of sediment attenuation, influence of bottom roughness, and role of the critical angle will be discussed in the context of these measurements. [Work supported by ONR, ASIAEX program.]

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. Nedzelnsky, 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
S1	1:45 p.m. to 3:15 p.m.
S3	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, 30 April 2003, 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>U.S. TAG Chair/Vice Chair</u>	<u>TC or SC</u>	<u>U.S. TAG</u>
ISO		
P. D. Schomer, Chair	ISO/TC 43 Acoustics	S1 and S3
H. E. von Gierke, Vice Chair		
P. D. Schomer, Chair	ISO/TC 43/SC 1 Noise	S12
H. E. von Gierke, Vice Chair		
D. J. Evans, Chair	ISO/TC 108 Mechanical Vibration and Shock	S2
R. Eshleman, Acting Chair	ISO/TC 108/SC 1 Balancing, including Balancing Machines	S2
A. F. Kilcullen, Chair	ISO/TC 108/SC2 Measurement and Evaluation of Mechanical Vibration and Shock as Applied to Machines, Vehicles and Structures	S2
R. Eshleman, Vice Chair		
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	ISO/TC 108/SC5 Condition Monitoring and Diagnostics of Machines	S2
R. F. Taddeo, Vice Chair		
G. Booth, Chair	ISO/TC 108/SC6 Vibration and Shock Generating Systems	S2
IEC		
V. Nedzelnsky, 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

R. D. Hellweg, Chair S12

Compaq Computer Corporation, Acoustics Lab, Mechanical Engineering Group, MR01-3/03, 200 Forest Street, Marlborough, Massachusetts 01752

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

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, 1 MAY 2003

ROOMS 103/104, 1:15 TO 5:00 P.M.

Session 4pAA**Architectural Acoustics, Noise and Education in Acoustics: Teaching Architectural Acoustics to Non-Acousticians**

Rendell R. Torres, Chair

Architectural Acoustics Program, Rensselaer Polytechnic Institute, School of Architecture, Greene Building, 110 8th Street, Troy, New York 12180-3590

Chair's Introduction—1:15

Invited Papers

1:20

4pAA1. Teaching room acoustics as a product sound quality issue. Mendel Kleiner and Daniel Vastfjall (Dept. of Appl. Acoust., Chalmers Univ. of Technol., S-41296 Gothenburg, Sweden)

The department of Applied Acoustics teaches engineering and architect students at Chalmers University of Technology. The teaching of room acoustics to architectural students has been under constant development under several years and is now based on the study of room acoustics as a product sound quality issue. Various listening sessions using binaural sound recording and reproduction is used to focus students' learning on simple, easy to remember concepts. Computer modeling using ray tracing software and auralization is also used extensively as a tool to demonstrate concepts in addition to other software for simple sound generation and manipulation. Sound in general is the focus of an interdisciplinary course for students from Chalmers as well as from a school of art, a school of design, and a school of music which offers particular challenges and which is almost all listening based.

1:35

4pAA2. Acoustics class at Berklee College of Music. Anthony K. Hoover (Cavanaugh Tocci Assoc. and Berklee College, 327F Boston Post Rd., Sudbury, MA 01776)

Berklee College of Music (in Boston) was developing its outstanding Music Technologies Division, and understood the need for a comprehensive class on acoustics. The result was a three-credit-hour class, offered twice per year, covering the fundamentals, architectural acoustics (outdoors, indoors, and transmission), vibration isolation, hearing and psychoacoustics, and more. One outgrowth was the Acoustical Society at Berklee, with presentations by local and visiting ASA members, yearly visits to an anechoic chamber, special studio sessions, tours, and joint meetings with professional societies. Over 2000 students have completed and performed well in the class. The author's favorite measure of success is the growing number of students who have chosen a career in acoustics. This paper will summarize and discuss this class.

1:50

4pAA3. Studio-based courses in architectural acoustics. Gary W. Siebein (School of Architecture, Univ. of Florida, P.O. Box 115702, Gainesville, FL 32611-5702, gsiebein@siebeinacoustic.com)

Examples of 4 nontraditional, elective courses offered to students in the Master of Architecture, Master of Science and Doctor of Philosophy degree programs in architecture are presented. Listening to Buildings involves experiential exercises in architectural acoustics for undergraduate students in a required large lecture course to form links between what people hear in a room and the architectural shapes and materials of the room that affects what is heard. The Applied Acoustics Design Lab takes students through each of the major components of an actual consulting project. Student work includes room acoustic design, sound isolation design, sound system design and HVAC noise control design for an actual building project. An Acoustical History of Theaters and Concert Halls teaches fundamentals of computer-based acoustic modeling and has students compare the physical design and aural simulations of a historic and a contemporary performance space. Graduate design studio in acoustics examines sound and silence as potential form-givers in architecture and as generative ideas for the basis of architectural interventions in complex sites and programs. Interior and exterior soundscapes and phenomenological philosophy form the background research for the studio.

2:05

4pAA4. Using acoustic measurement and analysis computer programs to teach architectural acoustics to architectural students. Robert C. Coffeen (School of Architecture and Urban Design, Marvin Hall, The Univ. of Kansas, Lawrence, KS 66045)

Explaining the concepts of sound amplitude and frequency to architectural students is difficult unless suitable visual and audible presentations are used. Computer programs designed for professional sound measurement and evaluation, and computer programs designed for acoustic analysis and auralization provide tools that can be employed to bring to life the basic sound concepts that must be understood by those students who will later be responsible for designing building spaces where hearing conditions are important.

2:20

4pAA5. Teaching noise control to architectural engineers. Ralph T. Muehleisen (Civil Environ. and Architectural Eng., Univ. of Colorado, 428 UCB, Boulder, CO 80309, ralph.muehleisen@colorado.edu)

Architectural engineers have to deal with acoustics in a variety of situations. HVAC engineers, structural engineers, safety engineers, lighting engineers, and construction engineers all need to be aware of noise, vibration, sound transmission, and general room acoustics. In this talk, the teaching of acoustics to architectural engineers is presented. In particular, the talk will focus on how to teach useful building noise control and architectural acoustics to architectural engineers with no previous background in acoustics and who may be less mathematically adept than other engineers or physicists.

2:35

4pAA6. A real challenge: Teaching acoustics to architecture students. Daniel R. Raichel (CUNY Graduate Ctr. and Douglas Eilar & Assoc., Fort Collins, CO 80526, raichel@juno.com)

The key to instilling the fundamentals of acoustics in architecture students is to arouse their interest. Because so many of the students are interested in music and high-fidelity equipment, it does not take much to ignite their interest in acoustics, particularly when they come to realize that perfectly good equipment can be undermined by poor room acoustics. Because they generally are not comfortable with mathematics, having had received perhaps no more than one or two semesters of introductory calculus; they need to be spoon-fed mathematics, even to the point of reviewing logarithmic manipulations, which are normally taught in secondary schools. The purpose of teaching acoustics to architects is not to make them acoustic experts, per se, but to make them appreciative of the effect of room acoustics and to understand that they must work hand-in-hand with acousticians when they design listening spaces that range in size from small classrooms to lecture halls to large concert halls. A regular acoustics text, such as that by Beranek, or Kinsler and Frey, or Raichel would be beyond the scope of an architectural course, but a text written especially for nonscience majors (such as that by Apfel) should and did serve admirably.

4p THU. PM

4pAA7. Teaching architectural acoustics to students of engineering versus students of architecture. Lily M. Wang (Architectural Eng. Prog., Univ. of Nebraska–Lincoln, 200B PKI, Omaha, NE 68182-0681, lwang4@unl.edu)

The author currently teaches architectural acoustics at the University of Nebraska to students in the College of Engineering and to students in the College of Architecture. These two groups have different backgrounds and focuses: for example, engineers are more mathematically prepared, while architects are more visually oriented. However, the instructor's goals for the members of both groups are similar, namely that they (1) master the fundamentals of good architectural acoustic design and (2) learn to respect and communicate well with each other during the design process, leading to more successful projects. This presentation will demonstrate the similarities and differences in the methods that are used by the author to address these two groups.

3:05–3:15 Break

3:15

4pAA8. One approach to architectural acoustics in education. J. Christopher Jaffe (Jaffe Holden Acoust., Inc., 114A Washington St., Norwalk, CT 06854)

In the fall of 1997, Dean Alan Balfour of the School of Architecture at the Rennselaer Polytechnic Institute asked me to introduce an undergraduate 14 credit certificate course entitled "Sonics in Architecture." Subsequently, the program was expanded to include a Master's Degree in Building Science. This paper discusses the trials and tribulations of building a scientific program in a liberal arts school. In addition, the problem of acquiring the research funds needed to provide tuition assistance for graduate students in Architectural Acoustics is reviewed. Information on the curriculum developed for both the lecture and laboratory courses is provided. I will also share my concerns regarding the teaching methods currently prevalent in many schools of architecture today, and how building science professionals might assist in addressing these issues.

3:30

4pAA9. The acoustics teaching drawdown at architectural schools in America: How to prepare learning materials for the new curricula. M. David Egan (Clemson Univ. and Egan Acoust., Box 365, Anderson, SC 29622)

In the early 1990s, the Association of Collegiate Schools of Architecture (ACSA) directory listed more than forty faculty members as teaching a course on acoustics. Today at the more than 120 ACSA full-member schools, there are only eight faculty members listed as teaching acoustics. Clearly the focus at most architecture schools in North America does not include the thorough study of science and technology. Courses in acoustics, illumination, HVAC systems, and fire protection are no longer included in curricula of most schools. Unlike studies in law and medicine, each school controls its own curriculum and what it calls its degree [A. O. Dean, *Arch. Record* (August 2002), pp. 84–92]. Learning materials have been developed to aid architecture professors who now must cover acoustics in the design studio. These materials, based on more than three decades of experience teaching and writing seven books on building technologies, also are intended to support self-study by students and design professionals. Examples of self-study exercises and hands-on exercises will be described [M. D. Egan, *Arch. Acoustics Workbook* (2000)]. A checklist of "writing principles" for the preparation of educational materials in architectural acoustics also will be presented.

Contributed Papers

3:45

4pAA10. Architectural acoustics to non acousticians: Prediction auralization in design and consulting. Quinsan Cio (Dept. of Architecture, Ball State Univ., Muncie, IN 47306)

Acoustic theories and principals are not necessarily familiar to most of the parties closely related to architectural acoustic issues. For example, architects involved in building designs, occupant communicates involved in space improvements are often the clients of architectural acoustic consultants. On one hand they are responsible to make knowledgeable decisions, on the other hand they do not have time or interest to be systematically trained in acoustics principals and theories. Properly educating such clients is one of the challenges constantly faced by architectural acoustic consultants. Traditionally, consultants are tormented between two different approaches. One is to deeply educate clients in architectural acoustics. The other is to convince the clients to trust and leave the matter to experts. Neither has been satisfactory. The former is insufficient and inefficient to establish appreciation of practical results, while the latter is resisted for lack of transparency. Thanks to the latest development of computing technologies, a new approach is available, which is to demonstrate the acoustic effects in audible media through auralization and prediction in conjunction with explanation of simple principles. This paper discusses this approach of client education in practices with case study examples.

4:00

4pAA11. The use of lattice gas method in explaining architectural acoustics principles. Sung Y. Yoon and Dana M. Smith (Dept. of Architecture, Architectural Acoust., Greene, RPI, Troy, NY 12180)

Many concepts in architectural acoustics are based on physical phenomena, such as sound propagation, reflection, etc. In explaining those phenomena, it is necessary to employ the principles of physical acoustics. Unfortunately, the methods of physical acoustics are not easily approachable unless one is familiar with mathematical physics techniques, partial differential equations, for example. The lattice gas method, based on the cellular automata concept, can be an alternative to avoid the use of complicated mathematical equations. A relatively simple code can be used to generate computer simulations. The existing models for sound propagation by Sudo and Sparrow, are examined. Models for other phenomena involving architectural acoustics are suggested.

4:15

4pAA12. Teaching acoustics courses online. Courtney B. Burroughs (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

Three graduate-level courses in noise control engineering have been taught by the Graduate Program in Acoustics through the Penn State World Campus for more than four years. In each course, students receive

(1) a study guide which gives an overview of the course, the software, and assignments, and (2) a CD which contains the content of the course with animations embedded in the text. Because all communications are online, students have been able to take these courses asynchronously on their own schedule with minimum disruptions in their home and work schedules. Students work on group projects, known as collaborative learning activities, and conduct simulated measurements using virtual instruments. Individual learning activities include study questions, written problems, computer coding, and interactive demonstrations. The first course has been offered four times and several students have completed all three courses. Overviews of the elements of these online courses that have been successful and the elements that have presented problems, along with a discussion of adjustments that we have been able to make and those that we would like to make to improve the courses, will be presented.

4:30–5:00

Panel Discussion

THURSDAY AFTERNOON, 1 MAY 2003

ROOM 201, 1:10 TO 2:45 P.M.

Session 4pABa

Animal Bioacoustics: Nature's Orchestra—Acoustics of Singing and Calling Animals II

Whitlow W. L. Au, Chair

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

Chair's Introduction—1:10

Contributed Papers

1:15

4pABa1. Nonlinear phenomena in northern mockingbird (*Mimus polyglottos*) vocalizations: Acoustics and physiology. Sue Anne Zollinger, Tobias Riede, and Roderick A. Suthers (Indiana Univ., 1001 E. Third St., Bloomington, IN 47405)

The bird sound source (syrinx), like the mammalian larynx, has been discovered to behave as a nonlinear system of coupled oscillators. Modeling approaches as well as excised syrinx experiments show that the syrinx can be considered a nonlinear sound source. Until now, only few acoustic data were available showing to what extent a birdsong species' vocal repertoire actually contains nonlinear phenomena, like subharmonics, frequency jumps, biphonation and deterministic chaos. The passerine syrinx, in contrast to the mammalian larynx, consists of two sound sources, one situated in each of the bronchi that are independently controlled. Bronchial airflow measurements showed that individual mockingbirds use each side of the syrinx independently. However, there are interactions of both sound sources, which have not yet been investigated acoustically/spectrographically. The present study shows how many of the spectrographically observed nonlinear phenomena are attributed to the interaction of the two sound sources, i.e., nonlinear phenomena, while both sound sources show positive airflow values, compared to single sound source utterances, i.e., positive flow through one sound source only. Nonlinear phenomena could be attributed to the interaction of the two sound sources (earlier referred to as "two-voice-phenomenon") but were also found to be produced by a single sound source.

1:30

4pABa2. Nonlinear acoustics in the pant-hoot vocalization of common chimpanzees (*Pan troglodytes*). Tobias Riede (Indiana Univ., 315 Jordan Hall, Bloomington, IN 47405), Adam Clark Arcadi, and Michael J. Owen (Cornell Univ., Ithaca, NY)

Pant-hoots produced by chimpanzees are multi-call vocalizations. While predominantly harmonically structured, pant-hoots can exhibit acoustic complexity that has recently been found to result from inherent nonlinearity in the vocal-fold dynamics. This complexity reflects abrupt shifts between qualitatively distinct vibration patterns (known as modes), which include but are not limited to simple, synchronous movements by the two vocal folds. Studies with humans in particular have shown that as the amplitude and vibration rate increase, vocal-fold action becomes in-

creasingly susceptible to higher-order synchronizations, desynchronized movements, and irregular behavior. We examined the occurrence of these sorts of nonlinear phenomena in pant-hoots, contrasting quieter and lower-pitched introduction components with loud and high-pitched climax calls in the same sounds. Spectrographic evidence revealed four classic kinds of nonlinear phenomena, including discrete frequency jumps, subharmonics, biphonation, and deterministic chaos. While these events were virtually never found in the introduction, they occurred in more than half of the climax calls. Biphonation was by far the most common. Individual callers varied in the degree to which their climax calls exhibited nonlinear phenomena, but we are consistent in showing more biphonation than any of the other forms. These outcomes demonstrate that understanding these calls requisitely requires an understanding of such events.

1:45

4pABa3. Intensity discrimination as a function of level and frequency in three species of birds. Amanda M. Lauer, Kirsten Poling, and Robert J. Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, alauer@psyc.umd.edu)

Many studies have examined frequency discrimination in birds, but there has not been as complete a description of avian intensity discrimination abilities. Birds appear to be slightly less sensitive to changes in intensity than humans and other mammals; however, few studies have systematically looked at the effects of both frequency and presentation level on intensity discrimination in birds. Here we describe intensity discrimination as a function of frequency and sensation level in two small songbird species, the canary (*Serinus canarius*), the zebra finch (*Taeniopygia guttata*), and a nonsongbird species, the budgerigar (*Melopsittacus undulatus*). Intensity difference limens (DLs) for pure tones were obtained from birds using standard operant conditioning procedures and the Method of Constant Stimuli. DLs ranged from approximately 2–6 dB, which are slightly larger than the DLs reported in mammals. For all three species, DLs become smaller with increasing presentation level, but show little effect across frequency for a given level. These results are consistent with previous reports in other species. [Work supported by NIH DC01372 to RJD and DC05450 to AML.]

4p THU. PM

4pABa4. Automatic type classification and speaker identification of african elephant (*Loxodonta africana*) vocalizations. Patrick J. Clemins and Michael T. Johnson (Elec. and Computer Eng. Dept., Marquette Univ., P.O. Box 1881, Milwaukee, WI 53201-1881, Patrick.Clemins@marquette.edu)

This paper presents a system for automatically classifying African elephant vocalizations based on systems used for human speech recognition and speaker identification. The experiments are performed on vocalizations collected from captive elephants in a naturalistic environment. Features used for classification include Mel-Frequency Cepstral Coefficients (MFCCs) and log energy which are the most common features used in human speech processing. Since African elephants use lower frequencies than humans in their vocalizations, the MFCCs are computed using a shifted Mel-Frequency filter bank to emphasize the infrasound range of the frequency spectrum. In addition to these features, the use of less traditional features such as those based on fundamental frequency and the phase of the frequency spectrum is also considered. A Hidden Markov Model with Gaussian mixture state probabilities is used to model each type of vocalization. Vocalizations are classified based on type, speaker and estrous cycle. Experiments on continuous call type recognition, which can classify multiple vocalizations in the same utterance, are also performed. The long-term goal of this research is to develop a universal analysis framework and robust feature set for animal vocalizations that can be applied to many species.

4pABa5. Sound localization of aerial broadband noise in pinnipeds. Marla M. Holt, Ronald J. Schusterman, David Kastak, and Brandon L. Southall (Long Marine Lab., Univ. of California, Santa Cruz, 100 Shaffer Rd., Santa Cruz, CA 95060)

Pinnipeds (seals, sea lions, and walruses) emit broadband calls on land as part of their communication system in order to coordinate their reproductive activities. How well do they localize these types of signals? In this study, the aerial sound localization acuities of a harbor seal (*Phoca vitulina*), a California sea lion (*Zalophus californianus*), and a northern el-

ephant seal (*Mirounga angustirostris*) were measured in the horizontal plane with a broadband white noise stimulus. Testing was conducted in a hemi-anechoic chamber using a left/right forced choice procedure to measure the minimum audible angle (MAA) for each subject. MAAs were defined as half the angular separation of two sound sources relative to a subject's midline that corresponded to 75% correct discrimination. MAAs were 3.6, 4.2, and 4.7 deg for the harbor seal, California sea lion, and northern elephant seal, respectively. These results demonstrate that these pinniped species had sound localization abilities comparable to the domestic cat and rhesus macaques. The acuity differences between our subjects were small, were not predicted by head size, and therefore likely reflect the relatively acute abilities of other pinniped species to localize aerial broadband signals.

4pABa6. Evoked potential measurement of the masked hearing threshold of a Pacific white-sided dolphin (*Lagenorhynchus obliquidens*). Whitlow W. L. Au (Marine Mammal Res. Prog., Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), Thomas Jeanette, A. Western (Western Illinois Univ./RIRUC, Moline, IL 61265), and Kenneth M. Rameriz (John G. Shedd Aquarium, Chicago, IL 60605)

The masked hearing threshold of a Pacific white-sided dolphin (*Lagenorhynchus obliquidens*) was determined by measuring the animal's auditory brainstem response (ABR). The dolphin was trained to wear surface-contact electrodes embedded in suction cups and to swim into a hoop centered at 1 m below the water surface facing a sound projector 5 m away. Broadband transient signals with center frequencies of 8, 16, 32, 64, 80, and 100 kHz were used as the stimuli. ABR signals were measured by digitizing the electrode signals in 32 point blocks at a sampling rate of 20 kHz. Five hundred blocks were averaged in order to obtain an ABR. The response latency for suprathreshold threshold signals was approximately 1.9 ms with the highest peak-to-peak ABR amplitude of approximately 2.8 uV occurring for a signal frequency of 64 kHz. The spectrum of the ABR signal was similar to that of *Tursiops truncatus*, with a major peak at 1120 Hz and a secondary peak at 664 Hz. Threshold was determined by progressively reducing the amplitude of the stimulus until an evoked potential could not be detected. The energy signal-to-noise ratio within an integration window at threshold varied between 1 and 8 dB.

THURSDAY AFTERNOON, 1 MAY 2003

ROOM 201, 3:00 TO 5:05 P.M.

Session 4pABb

Animal Bioacoustics: General Topics in Animal Bioacoustics

Elizabeth Brittan-Powell, Chair

Psychology Department, University of Maryland, Biology-Psychology Building, College Park, Maryland 20742

Chair's Introduction—3:00

Contributed Papers

4pABb1. Toward standardization of noise exposure in animal bioacoustics. Robert Burkard (Ctr. for Hearing & Deafness, Univ. at Buffalo, 215 Parker Hall, Buffalo, NY 14214)

In the United States, the American National Standards Institute (ANSI) Accredited Standards Committee on Bioacoustics S3 develops voluntary standards related to psychological and physiological acoustics, including areas such as general acoustics, vibration, and shock. Several years ago, an S3 working group (WG) was formed, WG90 Animal Bioacoustics, with

the goal of developing guidelines and standards in the area of animal bioacoustics, including relevant terminology, procedures to quantify noise exposure, and the effects of noise exposure on animals. The scope of this working group involves both terrestrial and marine animals, and their environments. Although ANSI Standards are voluntary, the development of accepted sound measurement procedures is needed before any legislation can be implemented that is intended to protect these animals from sound exposures detrimental to their long-term survival. To date, no standards have emerged from the Animal Bioacoustics working group. The purpose of this presentation is to provide an overview of the ANSI stan-

dards development process, briefly review some of standards that have been developed that are intended to quantify human exposure to sound, as well as describe the standards that need to be developed in order to protect both terrestrial and marine animals from excessive sound exposure.

3:20

4pABb2. Ocean noise and marine mammals: A summary report of the U.S. National Research Council Committee on Potential Impacts of Ambient Noise in the Ocean on Marine Mammals. George V. Frisk (Appl. Oceanogr. Phys. and Eng. Dept., M.S. #11, Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, gfrisk@whoi.edu)

In 2000, the U.S. National Ocean Partnership Program, with leadership by the Office of Naval Research, requested that the U.S. National Academies examine the potential effects of noise in the sea on marine mammals. A committee of 11 international experts was convened to study this controversial issue. The committee included marine mammals specialists, acousticians, bioacousticians, a geophysical exploration expert, and an expert in vessel engineering. Specifically, the committee was asked to evaluate the human and natural contributions to ocean noise and describe the long-term trends in noise levels, especially from human activities. In the report, the committee outlines the research needed to evaluate the impacts of marine noise from various sources (natural, commercial, naval, and acoustic-based ocean research) on marine mammal species, especially in biologically sensitive areas. The study reviews and identifies gaps in existing marine noise databases and recommends research needed to develop a model of ocean noise that incorporates temporal, spatial, and frequency-dependent variables. The findings and research recommendations of the committee will be presented.

3:35

4pABb3. Acoustic signal detection of manatee calls. Christopher Niezrecki, Richard Phillips, Michael Meyer (Dept. of Mech. and Aerosp. Eng., Univ. of Florida, Gainesville, FL 32611-6250), and Diedrich O. Beusse (College of Veterinary Medicine, Univ. of Florida, Gainesville, FL 32610-0126)

The West Indian manatee (*Trichechus manatus latirostris*) has become endangered partly because of a growing number of collisions with boats. A system to warn boaters of the presence of manatees, that can signal to boaters that manatees are present in the immediate vicinity, could potentially reduce these boat collisions. In order to identify the presence of manatees, acoustic methods are employed. Within this paper, three different detection algorithms are used to detect the calls of the West Indian manatee. The detection systems are tested in the laboratory using simulated manatee vocalizations from an audio compact disc. The detection method that provides the best overall performance is able to correctly identify $\approx 96\%$ of the manatee vocalizations. However the system also results in a false positive rate of $\approx 16\%$. The results of this work may ultimately lead to the development of a manatee warning system that can warn boaters of the presence of manatees.

3:50

4pABb4. Localization of dolphin whistles through frequency domain beamforming using a narrow aperture audio/video array. Keenan R. Ball (Woods Hole Oceanogr. Inst., Dept. 4., M.S. 18, Woods Hole, MA 02543) and John R. Buck (UMass Dartmouth, North Dartmouth, MA 02747)

Correlating the acoustic and physical behavior of marine mammals is an ongoing challenge for scientists studying the links between acoustic communication and social behavior of these animals. This talk describes a system to record and correlate the physical and acoustical behavior of

dolphins. A sparse, short baseline audio/video array consisting of 16 hydrophones and an underwater camera was constructed in a cross configuration to measure the acoustic signals of vocalizing dolphins. The bearings of vocalizing dolphins were estimated using the broadband frequency domain beamforming algorithm for sparse arrays to suppress grating lobes of Thode *et al.* [J. Acoust. Soc. Am. **107** (2000)]. The estimated bearings from the acoustic signals were then converted to video image coordinates and a marker was placed on the video image. The system was calibrated both at an indoor tank and from an outdoor dock at UMass Dartmouth prior to field tests in a natural lagoon at the Dolphin Connection on Duck Key, FL. These tests confirmed that the system worked well within the limits of underwater visibility by consistently placing the marker on or near the whistling or echolocating dolphin. [Work supported by NSF Ocean Sciences.]

4:05

4pABb5. Automated detection of sperm whale sounds as a function of abrupt changes in sound intensity. Christopher D. Walker (Dept. of Computer Sci., Univ. of Southern Mississippi, Hattiesburg, MS 39406, cdwalker@ocean.otr.usm.edu), Grayson H. Rayborn, Benjamin A. Brack, Stan A. Kuczaj, and Robin L. Paulos (Univ. of Southern Mississippi, Hattiesburg, MS 39406)

An algorithm designed to detect abrupt changes in sound intensity was developed and used to identify and count sperm whale vocalizations and to measure boat noise. The algorithm is a MATLAB routine that counts the number of occurrences for which the change in intensity level exceeds a threshold. The algorithm also permits the setting of a "dead time" interval to prevent the counting of multiple pulses within a single sperm whale click. This algorithm was used to analyze digitally sampled recordings of ambient noise obtained from the Gulf of Mexico using near bottom mounted EARS buoys deployed as part of the Littoral Acoustic Demonstration Center experiment. Because the background in these data varied slowly, the result of the application of the algorithm was automated detection of sperm whale clicks and creaks with results that agreed well with those obtained by trained human listeners. [Research supported by ONR.]

4:20

4pABb6. Behavior of dusky dolphins foraging on the deep-scattering layer in Kaikoura Canyon, New Zealand. Kelly Benoit-Bird (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), Bernd Wursig, and Cynthia McFadden (Texas A&M Univ., College Station, TX 77843-2258)

Little is known of foraging habits of sound-scattering layer consumers. A 200-kHz echosounder was used to survey dusky dolphins and the sound-scattering layer in winter 2002, in Kaikoura Canyon, New Zealand. Visual observations of dolphin surfacings occurred 84% of the time that dolphins were acoustically detected, confirming identifications from the acoustic data. Dusky dolphins were within the layer at 2000 h (about 1.5 h after dusk), within 125 m of the surface. As the layer rose to within 30 m of the surface at 0100 h, the observed depth of dolphins decreased presumably as the dolphins followed the vertical migration of their prey. The mean depth of dolphins was within the scattering layer except when the top of the layer was deeper than 125 m. Dusky dolphins often forage within large groups. Acoustically identified subgroups of coordinated animals ranged from 1 to 5 dolphins. Subgroup size varied with time of night,

minimum depth of the scattering layer, and the variance of the food resource. The largest subgroups occurred when the scattering layer was closest to the surface, and when the layer was most heterogeneous. Time, depth of layer, and layer variance contributed significantly to predicting foraging dusky dolphin subgroup size.

4:35

4pABb7. Head sinuses, melon, and jaws of bottlenose dolphins, *Tursiops truncatus*, observed with computed tomography structural and single photon emission computed tomography functional imaging. Sam Ridgway, Dorian Houser, James J. Finneran, Don Carder, William Van Bonn, Cynthia Smith (Navy Marine Mammal Prog., 235 PLBS, SSC SD, 53560 Hull, San Diego, CA 92152), Carl Hoh, Jacqueline Corbeil, and Robert Mattrey (School of Medicine, Univ. of California, San Diego, CA 92093)

The head sinuses, melon, and lower jaws of dolphins have been studied extensively with various methods including radiography, chemical analysis, and imaging of dead specimens. Here we report the first structural and functional imaging of live dolphins. Two animals were imaged, one male and one female. Computed tomography (CT) revealed extensive air cavities posterior and medial to the ear as well as between the ear and sound-producing nasal structures. Single photon emission computed tomography (SPECT) employing 50 mCi of the intravenously injected ligand technetium [Tc-99m] bismisate (Neurolite) revealed extensive and uptake in the core of the melon as well as near the pan bone area of the lower jaw. Count density on SPECT images was four times greater in melon as in the surrounding tissue and blubber layer suggesting that the

melon is an active rather than a passive tissue. Since the dolphin temporal bone is not attached to the skull except by fibrous suspensions, the air cavities medial and posterior to the ear as well as the abutment of the temporal bone, to the acoustic fat bodies of each lower jaw, should be considered in modeling the mechanism of sound transmission from the environment to the dolphin ear.

4:50

4pABb8. Concentric scheme of monkey auditory cortex. Hiroko Kosaki, Richard C. Saunders, and Mortimer Mishkin (Neuropsychology Lab., NIMH, 49/1B80, 49 Convent Dr., Bethesda, MD 20892)

The cytoarchitecture of the rhesus monkey's auditory cortex was examined using immunocytochemical staining with parvalbumin, calbindin-D28K, and SMI32, as well as staining for cytochrome oxidase (CO). The results suggest that Kaas and Hackett's scheme of the auditory cortices can be extended to include five concentric rings surrounding an inner core. The inner core, containing areas A1 and R, is the most densely stained with parvalbumin and CO and can be separated on the basis of laminar patterns of SMI32 staining into lateral and medial subdivisions. From the inner core to the fifth (outermost) ring, parvalbumin staining gradually decreases and calbindin staining gradually increases. The first ring corresponds to Kaas and Hackett's auditory belt, and the second, to their parabelt. SMI32 staining revealed a clear border between these two. Rings 2 through 5 extend laterally into the dorsal bank of the superior temporal sulcus. The results also suggest that the rostral tip of the outermost ring adjoins the rostroventral part of the insula (area Pro) and the temporal pole, while the caudal tip adjoins the ventral part of area 7a.

THURSDAY AFTERNOON, 1 MAY 2003

ROOM 208, 1:00 TO 4:00 P.M.

Session 4pBB

Biomedical Ultrasound/Bioresponse to Vibration, Signal Processing in Acoustics and Physical Acoustics: High-Intensity Focused Ultrasound (HIFU) and Imaged-Guided Therapy II

J. Brian Fowlkes, Chair

Department of Radiology, University of Michigan Medical Center, Kresge III R3315, 200 Zina Pitcher Place, Ann Arbor, Michigan 48109-0553

Invited Papers

1:00

4pBB1. Detection theory applied to high intensity focused ultrasound (HIFU) treatment evaluation. Narendra Sanghvi, Adam Wunderlich, Ralf Seip, Jahangir Tavakkoli (Focus Surgery, Inc., Indianapolis, IN 46226), Kris Dines (XDATA Corp., Indianapolis, IN 46226), Michael Bailly, and Lawrence Crum (Univ. of Washington, Seattle, WA)

The aim of this work is to develop a HIFU treatment evaluation algorithm based on 1-D pulse/echo (P/E) ultrasound data taken during HIFU exposures. The algorithm is applicable to large treatment volumes resulting from several overlapping elementary exposures. Treatments consisted of multiple HIFU exposures with an on-time of 3 seconds each, spaced 3 mm apart, and an off-time of 6 seconds in between HIFU exposures. The HIFU was paused for approximately 70 milliseconds every 0.5 seconds, while P/E data was acquired along the beam axis, using a confocal imaging transducer. Data was collected from multiple *in vitro* and *in vivo* tissue treatments, including shams. The cumulative energy change in the P/E data was found for every HIFU exposure, as a function of depth. Subsequently, a likelihood ratio test with a fixed false alarm rate was used to derive a positive or negative lesion creation decision for that position. For false alarm rates less than 5%, positive treatment outcomes were consistently detected for better than 90% of the HIFU exposures. In addition, the algorithm outcome correlated to the applied HIFU intensity level. Lesion formation was therefore successfully detected as a function of dosage. [Work supported by NIH SBIR Grant 2 R 44 CA 83244-02.]

4pBB2. Nonlinear pulse-echo imaging methods for HIFU-induced lesion visualization. Emad Ebbini (ECE Dept., Univ. of Minnesota, Minneapolis, MN 55455)

We have recently investigated the use of two different nonlinear imaging methods for visualization of HIFU-induced lesions. The hypothesis is that bubble activity at the location of the HIFU beam can be observed using diagnostic ultrasound with proper post-beamforming signal processing of the rf data. In particular, pulse inversion (PI) and quadratic imaging techniques have been investigated for this purpose. Results from over 100 *ex vivo* tissue experiments clearly demonstrate the superiority of the nonlinear imaging techniques over conventional B-scan imaging in terms of accurate mapping of lesion boundaries. We are currently examining imaging data to determine the use of nonlinear imaging methods for quantitative assessment of tissue damage. In this paper, we will describe a new experimental procedure for determining the existence of bubble activity associated with different levels of exposure to HIFU, from underexposure to overexposure conditions. Images from conventional B-scan will be compared and contrasted with quadratic and pulse inversion images for various levels of exposure. These results continue to support the hypothesis that bubble activity is extremely important in both lesion formation and lesion visualization using ultrasonic techniques.

1:50

4pBB3. Factors affecting radiation force method for monitoring therapeutic ultrasonic lesions. Frederic L. Lizzi, Robert Muratore, Samuel Mikaelian, Paul Lee, and S. Kaisar Alam (Riverside Res. Inst., 158 William St., New York, NY 10038)

Investigations are being conducted to develop a method using motion induced by radiation force to monitor HIFU lesions, by virtue of their increased stiffness. A therapeutic transducer, periodically excited at subthreshold levels, generates the radiation force: a collinear diagnostic transducer monitors the degree and time-course of induced motion. Lesions are detected by changes in pre- and post-treatment motion patterns. *In vitro* experiments and computer simulations have been performed to clarify the roles of several phenomena so that optimized systems can be designed for practical on-line applications. Lesion attenuation has been found to be a key factor since it directly affects radiation force. Lesions with near-normal attenuation exhibit less motion than normal, because of increased stiffness. Lesions with significantly higher attenuation can exhibit increased motion, due to increased force, but they demonstrate spatial patterns differing from normal. The profile of the push beam was found to affect the magnitude, spatial pattern, and time course of induced motion. The diagnostic transducer's bandwidth and beam profile were shown to affect the precision and sensitivity of motion characterization. Quantitative findings from these studies are being combined with thermal studies to design monitoring systems for specific applications.

2:15–2:30 Break

Contributed Papers

2:30

4pBB4. Study of a scanning HIFU therapy protocol, Part I: Theory and simulations. Steven G. Kargl and Marilee A. Andrew (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Over the last several years, many researchers have compared model predictions of isolated thermal lesions, caused by acoustic fields generated from (spherically) focused transducers, to thermal lesions created in tissue phantoms, *in vitro* soft tissue samples, and *in vivo* soft tissue samples. The models typically couple a nonlinear acoustic field from a focused transducer to the bio-heat transfer equation (BHTE). Recent experiments in polyacrylamide gel phantoms and excised bovine liver samples have demonstrated the possible deposition of thermal lesions in a scanning mode. An initial modeling effort to predict scanned thermal lesions will be discussed. The nonlinear acoustic field from a spherically focused transducer is predicted by a time-domain solution of the Khokhlov-Zabolotskaya-Kuznetsov equation. This field is then used as a heat source in the BHTE. The importance of scan rate, acoustic frequency, and time-averaged intensity will be investigated. [Work supported by USAMRMC.]

2:45

4pBB5. Fast calculations of the exact nearfield produced by a circular piston. Robert J. McGough (Dept. of Elec. and Computer Eng., Michigan State Univ., 2120 Eng. Bldg., East Lansing, MI 48824-1226) and Thaddeus V. Samulski (Duke Univ. Medical Ctr.)

The computation time and the numerical error are both reduced substantially with an improved technique that calculates the nearfield pressure produced by a circular piston. A comparison of simulation results shows that, once the errors for all results are normalized, this new technique is at least 3 to 4 times faster than the impulse response approach of Oberhettinger and Stepanishen. Further improvements are obtained when the computational grid is divided into sectors defined by the angle between the coordinates of each field point and the normal evaluated at the center of

the circular piston. The sector number is then converted through a linear function to a value that specifies the sampling of the integrand. When continuous wave excitations are applied to a simulated piston with a radius of five wavelengths, the resulting computations verify that this sectoring scheme reduces the computation time by an additional factor of 2 without increasing the maximum error value. These fast calculations are especially useful in numerical simulations for thermal therapy, where pressure fields are computed in large volumes for arrays with multiple sources.

3:00

4pBB6. Study of a scanning HIFU therapy protocol, Part II: Experiment and results. Marilee A. Andrew, Peter Kaczowski, Bryan W. Cunitz, Andrew A. Brayman, and Steven G. Kargl (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Instrumentation and protocols for creating scanned HIFU lesions in freshly excised bovine liver were developed in order to study the *in vitro* HIFU dose response and validate models. Computer-control of the HIFU transducer and 3-axis positioning system provided precise spatial placement of the thermal lesions. Scan speeds were selected in the range of 1 to 8 mm/s, and the applied electrical power was varied from 20 to 60 W. These parameters were chosen to hold the thermal dose constant. A total of six valid scans of 15 mm length were created in each sample; a 3.5 MHz single-element, spherically focused transducer was used. Treated samples were frozen, then sliced in 1.27 mm increments. Digital photographs of slices were downloaded to computer for image processing and analysis. Lesion characteristics, including the depth within the tissue, axial length, and radial width, were computed. Results were compared with those generated from modified KZK and BHTE models, and include a comparison of the statistical variation in the across-scan lesion radial width. [Work supported by USAMRMC.]

4pBB7. Estimating “true” HIFU-induced temperature changes using thermocouples. Xinmai Yang, Ronald A. Roy, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummings St., Boston, MA 02215, xmyang@bu.edu)

An easy and economical way to measure temperature rises in tissue and tissue-mimicking phantoms is with imbedded thermocouples. However, a thermocouple undergoing HIFU exposure is subject to the formation of a viscous boundary layer, leading to an additional local temperature rise not experienced by the rest of the focal region. The resulting temperature measurement is then not indicative of the average undisturbed temperature in the region. This is the so-called thermocouple artifact (we will argue against the use of the word artifact). Motivated by our bubble-enhanced heating experiments, we present a simple method for estimating HIFU-induced temperature changes. The influence of the thermocouple is modeled by an effective increase in the local sound absorption coefficient. The effective absorption coefficient is estimated by measuring the increased rate of heating brought on by the thermocouple. The temperature rise in the medium adjacent to the thermocouple is then predicted by incorporating the effective local absorption into the acoustic source term in the 3D heat conduction equation. Comparisons between simulation and experiment and between the reported method and other methods for thermocouple-artifact correction are presented. [Work supported by the U.S. Army.]

3:30

4pBB8. Monitoring evolution of HIFU-induced lesions with backscattered ultrasound. Ajay Anand and Peter J. Kaczkowski (Ctr. for Industrial and Medical Ultrasound, Univ. of Washington, 1013 40th St. NE, Seattle, WA 98105, ajaya@apl.washington.edu)

Backscattered radio frequency (rf) data from a modified commercial ultrasound scanner were collected in a series of *in vitro* experiments in which high intensity focused ultrasound (HIFU) was used to create lesions in freshly excised bovine liver tissue. Two signal processing approaches were used to visualize the temporal evolution of lesion formation. First,

apparent tissue motion due to temperature rise was detected using cross-correlation techniques. Results indicate that differential processing of travel time can provide temperature change information throughout the therapy delivery phase and after HIFU has been turned off, over a relatively large spatial region. Second, changes in the frequency spectrum of rf echoes due to changes in the scattering properties of the heated region were observed well before the appearance of hyper-echogenic spots in the focal zone. Furthermore, the increase in attenuation in the lesion zone changes the measured backscatter spectrum from regions distal to it along the imaging beam. Both effects were visualized using spectral processing and display techniques that provide a color spatial map of these features for the clinician. Our results demonstrate potential for these ultrasound-based techniques in targeting and monitoring of HIFU therapy, and perhaps post-treatment visualization of HIFU-induced lesions.

3:45

4pBB9. Acoustic field of a spherically curved wedge transducer. Jeffrey A. Ketterling (Riverside Res. Inst., 156 William St., New York, NY 10038)

High-intensity focused ultrasound (HIFU) applications typically utilize spherically curved transducers, often with modifications to the surface electrode. These modifications, such as rectangular cut-outs, can normally be modeled as a collection of annular rings and wedges. The acoustic field of an annular ring has been well characterized, but a wedge shape has not. We present a model of the acoustic field of a wedge transducer based on the spatial impulse response (SIR) method. A wedge transducer is created by making a cut normal to a spherical cap and keeping the section with the smaller surface area. Analytic expressions for the SIR are derived for an arbitrary point P in space. To find the time domain pressure at P , the SIR is convolved with the time derivative of the wedge surface velocity. Numerical examples of SIRs and the acoustic pressure field are given for a wedge formed from a spherical cap with a geometric focus of 9 cm and an outer diameter of 8 cm. For a single frequency drive of 4.7 MHz, calculations of pressure are compared to hydrophone measurements of a HIFU transducer with equivalent geometry. The results show an excellent agreement between theory and experiment.

THURSDAY AFTERNOON, 1 MAY 2003

ROOMS 110/111, 1:30 TO 3:15 P.M.

Session 4pEA

Engineering Acoustics, Signal Processing in Acoustics, Physical Acoustics, Structural Acoustics and Vibration: Nondestructive Evaluation and Material Characterization

Sally J. Pardue, Chair

Mechanical Engineering, Tennessee Technological University, 115 West 10th Street, Box 5014, Cookeville, Tennessee 38505

Contributed Papers

1:30

4pEA1. Towards a laser-vibrometry technique for measuring railroad rail stress. Vesna Damljanić and Richard L. Weaver (Dept. of Theoret. and Appl. Mech., Univ. of Illinois, 104 S. Wright St., Urbana, IL 61801, damljano@uiuc.edu)

There is a broad consensus on the need for cost-effective and reliable methods for measurement of axial stress in railroad rails. Constrained thermal expansions and contractions induce large forces which in turn lead to bucklings and rail breaks. We are developing a new method for such measurements, based on the decrease (increase) of effective dynamic flexural rigidity under the influence of compressive (tensile) loads. Scanned laser vibrometry measurements of vibration fields at a prescribed frequency, followed by a comparison with guided wave theory for the complex cross section of the rail, allows the extraction of flexural wavenum-

bers. These wavenumbers are in turn related to contained load and to the rail's intrinsic rigidity. Here we report the ongoing status of this work, describing the results of extensive laboratory measurements on unloaded rail, and test-bay measurements on rails subjected to compressive loads to 100 kip.

1:45

4pEA2. Analysis of digitized on-site ultrasonic state measurements. Sissay Hailu, Dov Hazony (Dept. of EECS, CWRU, Cleveland, OH 44106, dxh2@po.cwru.edu), and Yehonathan Hazony (Boston Univ., Boston, MA 02115)

Of concern is noninvasive *gage-less* monitoring fatigue experiments as well as detection and tracking of cracks, surface, and bulk. Accordingly, a digital signal decomposition method is developed for the analysis of sharp,

time-limited ultrasonic pulses traveling through a tensile specimen on-site. The method extracts from the propagating pulses, respective elongation, and diameter reduction due to stress, associated with signal delays and changes in amplitude and shape. The analytical process is based on a *nonlinear multivariate regression analysis*. When applied to the current experiments, the method extracts delay variations accurate to the order of ± 1 ps and quantified signal-deformation data. These results relate to a tensile-stress experiment under a five-cycle strain-cycle experiment, extending some of the cycles well into the plastic regime of a stainless steel specimen. The three-section cylindrical specimen consists of two gripper heads and a slender cylindrical section 6 mm in diameter and 20 mm long. Transducers are implanted on both specimen faces along the main axis [Mostafa *et al.*, Int. J. Fract. **85**, 99–109 (1997)]. [Work supported by NASA.]

2:00

4pEA3. Nonlinear acoustic spectroscopy applied to damage detection of elastic structures. Gerard Vanderborck (Thales Underwater Systems, 525 route des Dolines, BP 157, 06903 Sophia Antipolis Cedex, France, gerard.vanderborck@fr.thalesgroup.com) and Michel Lagier (Actea, 06183 Juan les Pins, France)

The strong linkage between damage and nonlinear elasticity of materials leads to much research in the field of nonlinear response of mechanical structures. The aim of this work is to demonstrate the relevance of nonlinear acoustic spectroscopy for damage detection in elastic structures. The proposed method consists of the observation of linear and nonlinear responses to harmonic excitations of the structure: the damage detection and localization are obtained from an array of sensors. The nonlinear response amplitude is proportional to the local nonlinearity (failure nonlinear elastic response). A dual frequency method has been demonstrated for cracked beam excited by two force sources located at different points; the theoretical analysis comes from the Green function method applied to each overtone displacement field, the crack elastic response is handled by a nonlinear force/displacement behavior. Small scale experiments are presented for aluminum and concrete beams and the experimental results confirm the capability of nonlinear acoustic spectroscopy for detection and monitoring of damaged structures. [Work supported by European Community.]

2:15

4pEA4. Linear and nonlinear angle beam ultrasonic spectroscopy to evaluate adhesive layers. Laszlo Adler, Bin Xie, Vadim Iakovlev (Adler Consultants, Inc., 1275 Kinnear Rd., Columbus, OH 43212, ladler1@aol.com), and Stanislav Rokhlin (Ohio State Univ., Columbus, OH 43210)

A new technique for quantitative evaluation of adhesive bond integrity of metals and composites called angle beam ultrasonic spectroscopy was developed. The novelty of this approach is that it utilizes both normal and obliquely incident ultrasonic beams on the bond line simultaneously and measures the frequency response of the reflected ultrasonic signals. From the frequency dependence of the reflection coefficients, elastic moduli, attenuation coefficients, thickness and density of the adhesive layer can be measured. The method was further enhanced by adding a low-frequency dynamic load acting as a nonlinear (parametric) enhancement. The low-frequency excitation parameters are set to make the maximum stress distribution coincide with the bond line. In addition to the bond layer properties, interfacial spring constants are also obtained by using the nonlinear method. Experimental results on adhesive samples will be presented and discussed with model prediction.

2:30

4pEA5. Influence of acoustic memory on ultrasonic attenuation. Michael McPherson, Igor Ostrovskii, and M. A. Breazeale (Natl. Ctr. for Physical Acoust., Coliseum Dr., University, MS 38677)

Recently a new physical phenomenon of acoustic memory has been described [Phys. Rev. Lett., **89**, 155506-1/3 (2002)]. In this lecture, we show for the first time the connection between acoustic memory and ultrasonic attenuation. Different results are given showing attenuation as a function of temperature, of ultrasound amplitude, and of frequency. We used [001] LiNbO₃ single-crystal samples and longitudinal acoustic waves over the frequency range from 10 to 32 MHz. Strong correlation exists between acoustic memory amplitude and the measured coefficient of ultrasonic attenuation. Acoustic memory causes strong variation in attenuation measurements. In this light, we discuss the known difficulties in practical attenuation measurements in ferroelectrics. We mention an unusual temperature dependence, including a strong hysteresis effect. The frequency and amplitude dependence of attenuation are also discussed.

2:45

4pEA6. A comparison of single crystal versus ceramic piezoelectric materials for acoustic applications. James F. Tressler (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375, tressler@pa.nrl.navy.mil)

For nearly fifty years, piezoelectric ceramics (primarily from the PZT family) have been the materials of choice as the active elements in sound projectors and receivers, medical ultrasound probes, etc. There is currently great interest in the materials community in the use of newly discovered single crystal relaxor ferroelectric materials as a replacement in applications that currently utilize piezoelectric ceramics. The salient features of single crystal piezoelectrics are strains an order of magnitude larger than are achievable in PZT and electromechanical coupling coefficients on the order of 90 percent. This presentation will provide a more thorough comparison of the physical and piezoelectric properties of single crystal piezoelectrics versus conventional PZT piezoceramics. This will hopefully provide the transducer designer with a better understanding of the pros and cons in the use of single crystals. In addition, a brief review of single crystal synthesis procedures will be described. Finally, some acoustic devices utilizing both single crystals and piezoceramics will be compared. [Work supported by DARPA and ONR Code 321-TS.]

3:00

4pEA7. Acoustical investigations of borate glasses containing oxides of some transition elements and ferric oxide dopants. Surjit Singh Bhatti and Kanwar Jit Singh (Appl. Phys. Dept., Guru Nanak Dev Univ., Amritsar-143005, India, kanwarjitsinghpad@yahoo.com)

Glass samples of manganese oxide borate and zinc oxide borate (with and without ferric oxide doping) have been prepared to study their acoustical, mechanical, and thermal behavior as function of composition. Sound velocities and attenuation measurements in these glass systems at 1, 2, and 5 MHz give elastic moduli, Poissons ratio, micro-hardness, acoustic impedance, internal friction, thermal expansion coefficient and Debye and softening temperatures. Structural changes involve boron anomaly, field strengths of cations, difference in ionic radii, and charge state of iron. Makishima-Mackenzie (theoretical model) and IR and NGR techniques confirm the conclusions arrived at. The network modifier (NWM) is varied from 25 to 45 mol % for manganese oxide borate and from 15 to 40 mol % with 10 mol % doping of ferric oxide. For zinc oxide borate glasses, it varies from 26 to 34 mol % and with 10 mol % of ferric oxide, its variation is from 15 to 35 mol %. Impact of doping by ferric oxide on the properties of these glass systems have been investigated.

4p THU. PM

Session 4pPA

Physical Acoustics: Atmospheric and Seismic Propagation

Vladimir E. Ostashev, Chair

Environmental Technology Laboratory, NOAA, 325 Broadway, Boulder, Colorado 80303

Contributed Papers

1:30

4pPA1. Predicting sound propagation in the atmosphere using an artificial neural network. Michael Mungiole (Army Res. Lab., Attn: AMSRL-CI-EE, 2800 Powder Mill Rd., Adelphi, MD 20783-1197, mmungiole@arl.army.mil) and D. Keith Wilson (Eng. R&D Ctr., U.S. Army CRREL, Hanover, NH 03755-1290)

The spectral content of acoustic signals is dramatically altered as sound propagates through the atmosphere due to refraction from vertical gradients in wind and temperature, ground interactions, and other effects. Parabolic equation (PE) techniques have been successfully used to calculate numerical solutions for many of these effects. The PE models generally produce accurate attenuation values, but execution time is excessive for many applications where near real-time results are required. To obtain sound level attenuation predictions more quickly, we developed an artificial neural network for a range of heights and source/receiver horizontal separations. The PE and boundary conditions were modified to obtain a nondimensional representation, which resulted in seven parameters required to specify all input combinations. This model version was then used to train the neural network, using a range of values for each parameter. The standard deviation of the errors (propagation model minus neural network simulations) was generally within 2 dB for the training data set containing approximately 15 000 cases. The smaller test data sets resulted in errors having standard deviations of approximately 1 dB. The neural network was a good predictor of the sound propagation model results except for small values of the nondimensional ground impedance parameter.

1:45

4pPA2. On a connection between amplitude fluctuations, phase fluctuations, and processing gain. Ronald A. Wagstaff (1 Coliseum Dr., NCPA, Univ. of Mississippi, Univ., MS 38677)

Weston *et al.* [Philos. Trans. R. Soc. London, Ser. A **265**, 595 (1969)] show the remarkable agreement between the spectra for the coefficients of variation of amplitude and phase, i.e., amplitude and phase fluctuations, for a hydrophone in shallow water. Wagstaff [J. Acoust. Soc. Am. **112**, 2422 (2002)] showed the functional similarity between phase angles and pseudo-phase angles for an outdoor microphone. Pseudo-phase angles are amplitude fluctuations that have been scaled appropriately to have a similar functionality in temporally coherent signal processing as phase angles. The concepts of the two previously mentioned references are merged to exploit the similarity of phase fluctuations and amplitude fluctuations to achieve multiplicative pseudo-coherent gain. Gains in excess of $20 \log(N)$ have been achieved (N is the number of samples averaged). $10 \log(N)$ is considered ideal for vector averaging. It is seldom achieved, because of coherent attenuation and cancellation associated with the use of real phase angles. Results are included for wind noise in outdoor measurements. [Work supported by U.S. Army Armament Research Development and Engineering Center.]

2:00

4pPA3. Cramer–Rao lower bounds on the angle-of-arrival estimates for a wave propagating in a random medium: Geometric acoustics regime. Sandra L. Collier (U.S. Army Res. Lab., AMRSL-CI-EE, 2800 Powder Mill Rd., Adelphi, MD 20783-1197, scollier@arl.army.mil) and D. Keith Wilson (Eng. Res. and Development Ctr., U.S. Army Cold Regions Res. and Eng. Lab, Hanover, NH 03755-1290)

In the geometric acoustics regime, a propagating wave is weakly diffracted and weakly scattered by the medium. The variance of the real component of the signal is much less than the variance of the imaginary component, thus the signal may not be modeled as a complex Gaussian random variable (whose real and imaginary components have equal variance), as is often done in the Rytov extension region, where both scattering and diffraction are strong. A statistical model for a signal in the geometric acoustics regime has been previously developed [S. L. Collier and D. K. Wilson, J. Acoust. Soc. Am. **111**, 2379 (2002)] and its properties have been further investigated [S. L. Collier and D. K. Wilson, ASA April 2003 Meeting on Signal Processing (submitted)]. This statistical model is applied here to an acoustic wave propagating in a random medium with fluctuations described by von Kármán's spectrum. Additive white Gaussian noise is also considered. The correlation functions of the phase and log-amplitude fluctuations for a von Kármán spectrum are derived in the geometric acoustics limit. The Cramer–Rao lower bounds (CRLBs) on the angle-of-arrival estimates are calculated assuming multiple unknown parameters. The range dependence of the CRLBs is studied in detail.

2:15

4pPA4. Equations for direct numerical simulation of sound propagation in a moving atmosphere. Vladimir E. Ostashev (Environ. Technol. Lab., 325 Broadway, Boulder, CO 80305), Lanbo Liu, D. Keith Wilson, Mark L. Moran (U.S. Army Cold Regions Res. and Eng. Lab., Hanover, NH 03755), David F. Aldridge (Sandia Natl. Labs., Albuquerque, NM 87185), and David Marlin (U.S. Army Res. Lab., White Sands Missile Range, NM 88002)

Most previous analytical and numerical studies of sound propagation in a moving atmosphere have been based on wave equations for the sound pressure and on various parabolic approximations to the wave equations. However, these equations cannot be used as starting equations for recently proposed direct numerical simulation (DNS) of sound propagation outdoor since such starting equations should be first-order differential equations with respect to time. In the present paper, we derive two closed sets of the first-order differential equations for the sound pressure and fluctuations in medium velocity and density due to a propagating sound wave. These sets can be used as starting equations for DNS of sound propagation in a moving atmosphere. The ranges of applicability of these sets are studied by comparing them with the equations for the sound pressure used previously. Note that both sets can also be employed for analytical studies of sound propagation in a moving atmosphere. Examples of the use of these sets for DNS and analytical studies of sound propagation in a moving

atmosphere are presented. [Work partially supported by a DoD High-Performance Computing Modernization Office grant and U.S. Army Research Office Grant No. DAAG19-01-1-0640.]

2:30

4pPA5. On the use of modal expansions to model broadband propagation in the nighttime boundary layer and other downward refracting atmospheres over lossy ground planes. Roger Waxler (N.C.P.A., Univ. of Mississippi, 1 Coliseum Dr., P.O. Box 1848, University, MS 38677, rwax@olemiss.edu)

A modal expansion for long-range, low-frequency propagation in downward refracting atmospheres over complex impedance planes has been developed. This modal expansion decomposes the sound field in the frequency domain into a ducted part (the modal sum) and an upwardly propagating part (expressed as an integral over a continuum). One may simulate the propagation of a pulse released from a point source by multiplying the Fourier transform of the pulse shape by the frequency domain Greens function and inverting the Fourier transform. At long distances from the source one expects that the sound field is adequately described by the modal sum alone. For sound speeds which are asymptotically (with height) constant this is indeed the case. However, the nighttime boundary layer is characterized by a sound speed which is downward refracting near the ground, increasing up to an inversion point, above which it decreases slowly becoming slightly upward refracting. In this case there are certain exceptional frequencies for which the continuum integral contains a resonant component which propagates to large horizontal distances and must be added to the modal sum.

2:45–3:00 Break

3:00

4pPA6. An analytic model for acoustic scattering from an impedance cylinder placed normal to an impedance plane. Michelle E. Swearingen and David C. Swanson (Penn State Univ., P.O. Box 30, State College, PA 16802, michelle.swearingen@pobox.com)

An analytic model for the scattering of a spherical wave off an infinite right cylinder embedded perpendicularly in a ground surface is developed in cylindrical coordinates. The model is developed to simulate a single tree and is developed as a first piece to creating a model for estimating attenuation in a forest based on scattering from individual tree trunks. Comparisons are made to the plane wave case, the transparent cylinder case, the transparent ground case, and the rigid ground case as a method of theoretically verifying the model. Agreement is excellent for all benchmark cases. Model sensitivity to five parameters is determined, which aids in error analysis, particularly when comparing the model results to experimental data, and offers insight into the inner workings of the model. An experiment was performed to collect real-world data on scattering from a cylinder normal to a ground surface. The data from the experiment is analyzed with a transfer function method into frequency and impulse responses. The model results are compared to the experimental data. Agreement is good below 1500 Hz, and poor for higher frequencies.

3:15

4pPA7. Examining surface sealing and crusting using the acoustic to seismic transfer function. Del Leary, Craig J. Hickey, James M. Sabatier (NCPA—Univ. of Mississippi, Oxford, MS 38677), and David A. DiCarlo (USDA—ARS Natl. Sedimentation Lab.)

Soil sealing is examined by measuring the acoustic to seismic (A/S) transfer function. An A/S transfer function is a swept sine measurement using a suspended loud speaker to impinge acoustic energy from the air

onto a soil sample. A laser Doppler vibrometer (LDV) is used to obtain the surface particle velocity as a measure of the seismic energy that has been transferred into the soil. This technique is noncontact and therefore allows successive measurements to be taken in time as the surface crust is formed. Soil samples are rained upon then allowed to dry forming a crust or seal that changes both the stiffness and hydraulic properties of the surface layer. Neshoba soils tested show a quantifiable decrease in the seismic energy transferred, as well as an increase in correlation between successive trials, as the crust forms. Additional measurements done with a submerged transducer show an even greater decrease in surface velocity. Current studies also include an acoustic reflection technique to measure changes in the hydraulic properties. [Work supported by USDA ARS.]

3:30

4pPA8. The measurements of the compressional wave velocity of soils during unconsolidated-undrained triaxial testing. Zhiqiu Lu, C. J. Hickey, and J. M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Coliseum Dr., University, MS 38677)

In this study, a conventional triaxial cell was modified to measure the compressional wave velocity during a triaxial test. Two air-dry remolded soils taken from the counties in Sharkey and Neshoba, MS, were chosen for the study. Unconsolidated-undrained triaxial tests with pore pressure measurement were carried out. Soil samples were isotropically consolidated up to three different cell pressures, axially compressed with the axial strain up to 22%, and subjected to the unload–reload stress path cycles before and after soil failures. The velocity of the compressional wave in the axial direction as a function of the axial strain was measured along with the measurement of the stress–strain response. A comparison of the load-deformation behavior with the load-acoustic velocity behavior was made. The variation of the acoustic velocity with the effective stresses during the isotropic loading, normally consolidated compression and unload–reload stress path cycles were examined. Several empirical expressions of the compressional wave velocity in terms of the effective stresses and the over-consolidation ratio were proposed and examined with the measured data.

3:45

4pPA9. Coordinate determination of the low-flying object by the acoustical method. Igor L. Obozenko (Kyiv Polytechnic Univ., Kyiv 03057, Ukraine) and Tran H. Dat (Nagoya Univ., Nagoya 464-8603, Japan)

The acoustical method of passive sonar is proposed for determination of the object flying at low altitudes. The method is based on sound refraction into the lower relatively to the object medium. The Rayleigh waves are used with the speed far higher than the low-flying noisy target. Besides the spherical wave, the side wave with attenuation in depth is reflected to the receiver. The latter can be clearly detected on the depth of the surface half-wave that in reality for the infralow frequency amounts up to hundreds of meters. The location finding of the object is obtained via creation of the static fan of the polar pattern with the specified angular resolution of each radar view. At the horizontal range (100 km), depth of immersion multiple of N receivers (horizontal collinear equidistant antenna) and the height of the flying object 3–10 m, in the wave field the correction to the ray acoustics, which is very significant in its amplitude, is done. To count horizontal range to the object, a vectorial (three components) displacement (or vibrating speed) receiver, placed into the center of the radar array is used. The model experiments in water match the calculations.

4p THU. PM

Session 4pSA**Structural Acoustics and Vibration and Musical Acoustics: Structural Acoustics of Musical Instruments**

Jeffrey S. Vipperman, Chair

*Department of Mechanical Engineering, University of Pittsburgh, 648 Benedum Hall, 3700 O'Hara, Pittsburgh, Pennsylvania 15261***Chair's Introduction—1:30*****Invited Papers*****1:40****4pSA1. Structural acoustics for the violin.** George Bissinger (Phys. Dept., East Carolina Univ., Greenville, NC 27858) and Earl G. Williams (Naval Res. Lab., Washington, DC 20375-5350)

The violin is a long-standing problem in acoustics due to the complexity of its shape, the interactivity of its various modes of response, and the inherent complications of coupling these to the acoustic field, independent of any quality aspects. By combining zero-mass loading modal analysis of the violin with simultaneous acoustic measurements over a sphere in an anechoic chamber, major acoustic energy sources were identified among the coterie of corpus (top+back+ribs) modes and interior cavity (via two ports, or f-holes) modes up to 4 kHz. Direct radiation from substructures such as the neck-fingerboard, bridge, and tailpiece was not significant. For each mode the total damping and the radiation efficiency and damping were computed, along with a simple top-back radiation directivity ratio using rms radiativities. The fraction-of-vibrational-energy-radiated was estimated from the radiation-to-total-damping ratio. A new addition to the modal violin acoustics investigations was the application of near-field patch holography to help determine the relative contributions of corpus and port radiation to overall violin radiativity. [Work supported by NSF and ONR.]

2:00**4pSA2. Analysis of the radiation from the violin f-holes using patch near field acoustical holography.** Earl G. Williams (Naval Res. Lab., Washington, DC 20375, williams@pa.nrl.navy.mil) and George Bissinger (East Carolina Univ., Greenville, NC 27858)

Although differential pressure measurements offer a direct means to understand the energy flow from the f-holes of the violin they have been performed only at discrete frequencies over relatively limited portions of the acoustic field, and none have ever covered an entire f-hole over a broad frequency region. Application of recently developed near field acoustical holography (NAH) patch processing techniques to 108-node planar rectangular grid microphone data provides a powerful tool to understand the flow of acoustic energy from the f-holes up to 4 kHz. The grid covered each f-hole as well as a small portion of the violin top-plate and provided the necessary spatial resolution to allow isolation of only f-hole aperture radiation in the NAH processing. The projected radiativity in the far field at 1.2 m from just the f-holes was compared with prior microphone measurements in an anechoic chamber over an entire sphere around the violin. As expected the lowest cavity mode A0 was the major radiator at the frequencies below all the corpus modes. Surprisingly the first corpus bending modes appear to radiate strongly through the f-holes also. [Work supported by ONR and NSF.]

2:20**4pSA3. A structural dynamics and experimental investigation of the American five-string banjo.** Joe Dickey (Ctr. for Non Destructive Evaluation, Johns Hopkins Univ., Baltimore, MD 21218) and Ray Wakland (Penn State Univ., University Park, PA 16804)

The American five-string banjo is unique among musical instruments in that many significant parameters that effect tone are easily adjusted. This is probably why so many banjo players fiddle with their banjo. The instrument is a combination of canonical vibrating systems, i.e., plucked strings that drive a circular, radiating, membrane. A structural dynamics model and experiment are used to characterize the sound and relate changes in sound to setup parameters. Three figures-of-merit, FOMs, are defined; they are loudness, brightness, and decay rate of the sound. The effects of a number of parameters on the FOMs are investigated analytically and experimentally. Among these are the loss factor and tension of the membrane, mass of the bridge, and the location on the string of the excitation. It is noted that the calculated effects of the changes agree with generally accepted setup practices.

4pSA4. The application of smart structures toward feedback suppression in amplified acoustic guitars. Steven F. Griffin (3112 Dakota St. NE, Albuquerque, NM 87110), Steven A. Lane (Airforce Res. Lab., Kirtland AFB, NM), and Robert L. Clark (Duke Univ., Durham, NC)

Smart structures technology can be applied to amplified acoustic guitars to prevent instability resulting from acoustic feedback. This work presents a coupled model of the guitar dynamics and the acoustic feedback mechanism, and explains how a simple control loop using a piezoceramic actuator can be used to reduce the effects of acoustic feedback. In addition to model simulations, experimental results using a real system and a simple controller are presented. The results show that a significantly higher (7-dB) guitar output can be achieved before instability, without detrimentally affecting the amplified and un-amplified guitar response.

3:00

4pSA5. SEA applications to wind instruments. Peter L. Hoekje (Dept. of Phys., Baldwin-Wallace College, Berea, OH 44017, phoekje@bw.edu)

The behavior of wind instruments, including brass instruments, is primarily determined by the shapes of their air columns, and their interaction with the sound generation mechanism. However, the influence of the surrounding body of the instrument has been a matter of some debate, and papers exploring this question have been published since the early years of the J. Acoust. Soc. Am. An apparent correlation between instrument material and playing behavior is disputed by arguments that the structure is stiff and massive compared to the air inside, and that many of the apparent effects are linked to machining differences among materials. The complexity of the instrument body makes this problem well suited for Statistical Energy Analysis (SEA), which treats the air column and the external structure as coupled statistical subsystems that share energy. For trumpets and trombones, the power radiated from the structural vibrations is about 40 dB lower than the energy radiated directly from the air column, with an enhancement at high frequencies due in part to the increasing modal density of the three dimensional structure. The coupling to the structural vibrations themselves from the player's lips and from the air vibrations are similar to each other in magnitude.

3:20–3:35 Break

3:35

4pSA6. Modal analysis and sound radiation from Caribbean steelpan. Andrew Morrison (Dept. of Phys., Northern Illinois Univ., DeKalb, IL 60115)

The Caribbean steelpan is one of the most interesting acoustic musical instruments invented in the last century. Although simple in design, the acoustic properties of the steelpan are surprisingly complicated. Holographic interferometry was used to determine the modes of vibration of a double second and a low tenor steelpan. Sound intensity measurements were taken to explore the relationship between the modes of vibration and the radiated sound field. The pan was placed in an anechoic chamber, and selected notes were excited electromagnetically with a swept sinusoid signal. A two-microphone probe was used to gather sound intensity measurements. Sound intensity maps and animations were constructed for the first three harmonics.

Contributed Papers

3:55

4pSA7. The effects of bell vibrations on the acoustic spectrum of the trumpet. Thomas R. Moore (Dept. of Phys., Rollins College, Winter Park, FL 32789, tmoore@rollins.edu)

The acoustic spectrum of a modern trumpet with the bell section heavily damped has been compared to the spectrum of the same instrument with the bell section left free to vibrate. The amplitude of vibration of the metal was measured in both cases and was shown to be significantly different between the two sets of measurements. Artificial lips were used to ensure consistency between trials. A significant change in the acoustic spectrum between the two cases is found, with the variation being largest in the lower harmonics where the relative power may change by as much as a factor of 2. It is shown that the changes can be explained by a variation in the viscous boundary layer that is attributable to the vibrating walls of the bell. [Work supported by a grant from the Jessie Ball duPont Fund.]

4:10

4pSA8. Structural dynamics of concert harps with wooden and composite material soundboards. Melinda Carney and Thomas J. Royston (Univ. of Illinois at Chicago, 842 W. Taylor St. MC 251, Chicago, IL 60607, troyston@uic.edu)

The replacement of a Sitka spruce grand concert harp soundboard with a carbon fiber-reinforced plastic soundboard could provide improved durability and long-term stability. Experimental vibratory studies on concert

harps with wooden soundboards are reviewed. A computational finite-element model is used to identify critical material properties by matching its predictions to the experimental data. With the material properties identified in the finite-element model, the lay-up of the composite soundboard is created using matching criteria based on research of wood replacements for violin top plates. The composite lay-up is then incorporated into the finite-element model, verifying that the dynamic response closely approximates that of the wooden soundboard. The identification technique and composite replacement design process may be applicable to other musical instruments, as well as other nonmusical, wooden plate structures.

4:25

4pSA9. An experimental study of changes in the impulse response of a wood plate that is subject to vibrational stimulus. Jared Grogan, Michael Braunstein, and Andrew Piacsek (Dept. of Phys., Central Washington Univ., Ellensburg, WA 98926)

It is a well-known dictum among players of stringed instruments that the tone of a new instrument improves with playing and that a fine instrument needs to be played if it is to maintain its optimum sound quality. This process is sometimes referred to as "playing in" an instrument. There is scant mention in the scientific literature, however, of a quantitative analysis of this phenomenon. As a first step in rigorously testing this hypothesis, measurements were made of tap tones of rectangular pieces of thin spruce before and after they were subjected to vibrational stimulus. Four spruce rectangles (20x28 cm) were cut from a single sheet obtained from a luthier supplier; three of these were stimulated at different amplitudes, while the fourth was a control plate. The stimulus (provided by a harmonically

driven guitar string connected to the plate via a bridge) lasted approximately 10 weeks, during which time tap tones of all four plates were periodically recorded. Spectrograms of the tap tones are compared among the plates and over time. A preliminary analysis of the data does not reveal any significant changes in the acoustic response of the plates.

4:40

4pSA10. Effect of body shape on vibration of electric guitars. Daniel A. Russell, Wesley S. Haveman, Willis Broden, and N. Pontus Weibull (Sci. & Math. Dept., Kettering Univ., 1700 W. Third Ave., Flint, MI 48504, drussell@kettering.edu)

The body vibrations of an electric guitar are typically ignored since the string vibrations are converted to sound through the use of a magnetic pickup. However, vibrations in the neck have been shown to cause dead spots at certain fret positions [H. Fleischer, *J. Acoust. Soc. Am.* **105**, 1330 (1999)]. In this paper we compare the vibrational mode shapes and frequencies of three electric guitars with different body shapes. Two guitars are solid-body electrics: one with a body shape which is symmetric about the neck axis (Epiphone Coronet) and the other which is not (Gibson Explorer). Mode shapes and frequencies are considerably different for the body, though neck vibrations are more closely related. The third guitar is an arched top hollow-body electric (Gibson ES-335). For this guitar, the top and back plates and the air cavities may also contribute to the guitar sound quality. Mode shapes and frequencies are determined from experimental modal analysis using an impact hammer and accelerometer.

4:55

4pSA11. On the scaling of free-reed organ pipes: A comparison of 19th-century theories to modern nonlinear models. Jonas Braasch (Institut für Kommunikationsakustik, Ruhr-Universität Bochum, 44780 Bochum, Germany)

After their introduction in the 19th century, the scaling of free reed organ pipes was often done on the basis of investigations by W. Weber, J. G. Töpfer, and F. Haas. Especially, Töpfer is often named as the founder of mathematical/physical oriented scaling. Unfortunately, it was impossible in the 19th century to derive a number of measures (e.g., optimal lengths of a conical pipe and the boot) theoretically. In this presentation it is demonstrated how recent theories can be used to derive some of those measures. For this purpose, a one-dimensional physical model was developed. The model uses the reed generator developed by Tarnopolsky *et al.* [*J. Acoust. Soc. Am.* **108**, 400–406 (2001)] and a wave-guide model for the resonator so as to determine the optimal resonator and boot dimensions. Wrong assumptions that were made in the 19th century will be also addressed. For example, organ builders frequently used the solution of Weber to determine the optimal resonator length. The solution of Weber was aimed to make the reed pipe's frequency independent from temperature and blowing pressure, but what the organ builders would have needed is a pipe whose frequency is mainly determined by the resonator length, so that the pipes would detune with the flue pipes, when the temperature changes.

THURSDAY AFTERNOON, 1 MAY 2003

ROOM 205, 2:00 TO 4:40 P.M.

Session 4pSC

Speech Communication and Psychological and Physiological Acoustics: The Funding Game—Strategies for Different Federal Agencies for Behavioral Sciences

Astrid Schmidt-Nielsen, Cochair

Office of Naval Research, Code 342, 800 North Quincy Street, Arlington, Virginia 22217-5660

Patrice S. Beddor, Cochair

University of Michigan, 1076 Frieze Building, Ann Arbor, Michigan 48109-1285

Chair's Introduction—2:00

Invited Papers

2:05

4pSC1. Merit-based funding at the National Science Foundation. Mary P. Harper (Natl. Sci. Foundation, 4201 Wilson Blvd., Arlington, VA 22230, mharper@nsf.gov)

National Science Foundation (NSF) invests in the best ideas from the most capable people as determined by competitive merit review. This talk will first summarize how NSF evaluates proposals for research and education projects in speech communication. The review timeline, as well as the merit review criteria of intellectual merit and broader impacts, will be discussed. Finally, a summary of current and future NSF funding opportunities will be provided.

2:20

4pSC2. NSF funding opportunities in the behavioral and cognitive sciences. Philip Rubin (Natl. Sci. Foundation, 4201 Wilson Blvd., Ste. 995, Arlington, VA 22230, prubin@nsf.gov)

This presentation will provide a brief overview of existing and emerging funding opportunities at the National Science Foundation (NSF). The Division of Behavioral and Cognitive Sciences (BCS) at the NSF supports research to develop and advance scientific knowledge focusing on human cognition, language, social behavior and culture, as well as research on the interactions between human societies and the physical environment. BCS programs consider proposals that fall squarely within disciplines, but they also encourage and support interdisciplinary projects, which are evaluated through joint review among Programs in BCS, as well as a joint review with programs in other Divisions, and NSF-wide multi-disciplinary panels, as appropriate. All programs in BCS consider proposals for

research projects, conferences, and workshops. Some programs also consider proposals for doctoral dissertation improvement assistance, the acquisition of specialized research and computing equipment, group international travel, and large-scale data collection. BCS participates in special initiatives and competitions on a number of topics, including cognitive neuroscience, children's research, human origins, and the development of infrastructure to improve data resources, data archives, collaboratories, and centers.

2:30

4pSC3. Seeking NIH funding: Defining the process. Lana Shekim (Natl. Institutes of Health, Bethesda, MD 20892)

The presentation will provide a brief introduction to the National Institute on Deafness and other Communication Disorders (NIDCD) with emphasis on the Voice and Speech program in the Division of Extramural Research. The process of seeking NIH funding will be outlined and a number of funding mechanisms will be described. The peer review process and the time course of a grant application will be highlighted.

2:40

4pSC4. Training camp: The quest to become a new National Institutes of Health (NIH)-funded independent investigator. Daniel A. Sklare (SPB/DER, NIDCD, NIH, EPS-400C, 6120 Executive Blvd., MSC-7180, Rockville, MD 20892-7180)

This presentation will provide information on the research training and career development programs of the National Institute on Deafness and Other Communication Disorders (NIDCD). The predoctoral and postdoctoral fellowship (F30, F31, F32) programs and the research career development awards for clinically trained individuals (K08/K23) and for individuals trained in the quantitative sciences and in engineering (K25) will be highlighted. In addition, the role of the NIDCD Small Grant (R03) in transitioning postdoctoral-level investigators to research independence will be underscored.

2:50

4pSC5. Office of Naval Research funding strategies. Astrid Schmidt-Nielsen (ONR, Code 342, 800 N. Quincy St., Arlington, VA 22217-5660, schmidat@onr.navy.mil)

The best strategies for obtaining military research funding can be quite different from NIH and NSF. Funding opportunities at the Office of Naval Research will be discussed. Different types of Navy/military funding will be described, and strategies for obtaining various types of funding, including the Young Investigator Program, will be discussed.

3:00

4pSC6. Air Force program in hearing research. Willard D. Larkin (Air Force Office of Sci. Res., AFOSR/NL, Rm. 713, 4015 Wilson Blvd., Arlington, VA 22203-1954, willard.larkin@afosr.af.mil)

The Air Force Office of Scientific Research (AFOSR) manages all the basic science activity for the U.S. Air Force. Research is supported in the Air Force Laboratory, and grants are provided to universities and research institutions. This presentation will outline the scope of AFOSR's support for basic research in areas germane to human hearing. Topics of current interest include the complex of problems associated with spatial hearing, the psychophysical and perceptual foundations for the design of auditory displays, problems of auditory information overload, and neurologically informed computational models of auditory perceptual processing. To develop better methods of hearing protection, AFOSR also supports efforts to understand the biophysical basis of cochlear excitation via bone conduction. Opportunities to contribute to these and related topics will be described.

3:10–3:40

Panel Discussion

3:40–4:40

Breakout Sessions

4p THU. PM

Session 4pUW

Underwater Acoustics and Acoustical Oceanography: High Frequency Sediment Acoustics II and Ambient Noise

Eric I. Thorsos, Chair

Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105-5598

Contributed Papers

1:00

4pUW1. *In situ* measurement of geoacoustic sediment properties: An example from the ONR Mine Burial Program, Martha's Vineyard Coastal Observatory. Barbara J. Kraft, Larry A. Mayer (Ctr. for Coastal and Ocean Mapping, 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)

In support of the Office of Naval Research's Mine Burial Program (MBP), *in situ* acoustic and resistivity measurements were obtained using ISSAP, a device developed and built by the Center for Coastal and Ocean Mapping. One of the field areas selected for the MBP experiments is the WHOI coastal observatory based off Martha's Vineyard. This area is an active natural laboratory that will provide an ideal environment for testing and observing mine migration and burial patterns due to temporal seabed processes. Seawater and surficial sediment measurements of compressional wave sound speed, attenuation, and resistivity were obtained at 87 stations. The ISSAP instrument used four transducer probes arranged in a square pattern giving acoustic path lengths of 30 and 20 cm with a maximum insertion depth of 15 cm. Transducers operated at a frequency of 65 kHz. The received acoustic signal was sampled at a frequency of 5 MHz. A measurement cycle was completed by transmitting 10 pulses on each of the five paths and repeating three times for a total 150 measurements. Resistivity measurements were obtained from two probes mounted on ISSAP following completion of the acoustic measurements. [Research supported by ONR Grant Nos. N00014-00-1-0821 and N00014-02-1-0138.]

1:15

4pUW2. Development of an impedance tube technique for *in-situ* classification of marine sediments. Preston S. Wilson, Ronald A. Roy, and William M. Carey (Dept. of Aersp. and Mech. Eng., Boston Univ., Boston, MA 02215)

The removal of samples or the insertion of measuring devices into the ocean bottom can alter the acoustic behavior of these sedimentary materials. A less invasive method for the acoustic characterization of marine sediments in the 1 to 10 kHz frequency range has been investigated. A water-filled impedance tube has been used to measure the complex reflection coefficient at the surface of a sediment layer at the bottom of a water tank. The acoustic properties of the sediment were obtained as a function of frequency by an inversion process. The system is modeled numerically and an effective fluid description of the sediment is used. The difference between the measured and predicted complex reflection coefficient is minimized through variation of the sediment sound speed and attenuation in the model. Results will be presented for different sediments. [Work supported by U. S. Navy ONR.]

1:30

4pUW3. High-frequency seafloor studies using buried directional receivers. Anthony P. Lyons (Appl. Res. Lab., The Penn State Univ., P.O. Box 30, State College, PA 16804-0030, apl2@psu.edu) and John C. Osler (Defence R&D Canada Atlantic, Dartmouth, NS B2Y 3Z7, Canada)

Knowledge of the acoustic arrival angle can be useful for studying penetration mechanisms and for estimating sediment sound speed dispersion. The arrival angle of acoustic energy penetrating the sediment water interface, however, is difficult to measure using sparsely distributed pressure sensors. The arrival angle, as well as amplitude information, can be unambiguously obtained by measuring particle motion with directional receivers. An experiment was conducted off of Elba Island in 1999 to assess the feasibility of using of high-frequency accelerometers to measure directionality. Measurements of acoustic penetration onto a sandy seafloor were obtained over wide frequency range (2.5 to 29 kHz) using the accelerometers. Results from the experiment demonstrating this novel measurement technique will be presented, as will issues concerning accelerometer design and calibration.

1:45

4pUW4. Use of a buried array to characterize sediment volume scattering. Darrell R. Jackson, Kevin L. Williams, Anatoliy N. Ivakin, and Eric I. Thorsos (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

During the SAX99 experiment, a buried array was used to measure the sound field due to sources placed in the water column. Measurements were made over the frequency range 11–50 kHz. The sound field at the array showed a "coda" that was longer than suggested by simulations of rough-interface scattering. Simulations have been used to show that scattering by sediment volume heterogeneities is the likely cause of the coda. The volume scattering strength has been estimated by fitting a point scatterer model to the data, and the strongest discrete scatterers have been located and characterized by means of backpropagation. When the measured properties of the volume scatterers are inserted in backscattering models, the predicted backscattering strengths fall somewhat below measured values, indicating that roughness scattering is dominant at this (sandy) site. [Work supported by ONR.]

2:00

4pUW5. Acoustic imaging of small water-lain sand deposits. Max Deffenbaugh, Neal L. Adair, David C. J. D. Hoyal, and David E. Giffin (ExxonMobil Upstream Res. Co., P.O. Box 2189, Houston, TX 77252)

Reduced-scale physical modeling of depositional systems, like submarine fans, river deltas, and point bars, provides insight into the formation and internal structure of the full-scale systems which may become economic hydrocarbon reservoirs. Turbid water with controlled sediment concentration and flow velocity is discharged into a 3 m×5 m tank of still water to create deposits up to typically 10 cm thick. These deposits are imaged by a pencil-beam high-frequency (7 MHz) acoustic system to cap-

ture the evolution of deposit topography and by a broad-beam lower-frequency system (150 kHz) to image an internal structure. An x-y positioning system moves the transducer to create detailed 3D images. At 7 MHz, the deposit surface is "rough," so significant backscattered energy is detected even for non-normal incidence. This, together with the narrow beamwidth, allows the deposit elevation directly below the sensor to be measured independent of the local slope. The deposit surface is "smooth" to the 150 kHz system, so reflections come only from points of normal incidence. This makes imaging more complicated, but the lower frequencies penetrate the deposit and reveal some internal structure. Images from both systems will be shown and compared.

2:15

4pUW6. Porosity and grain size dependence of the longitudinal wave velocity of water-saturated beach sand. Masao Kimura and Masahiro Noguchi (Dept. of Geo-Environ. Technol., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka 424-8610, Japan, mkimura@scc.u-tokai.ac.jp)

The longitudinal wave velocity of water-saturated sand is dependent on the porosity. The data which show the relationship between the velocity and the porosity are dispersed [E. L. Hamilton and R. T. Bachman, J. Acoust. Soc. Am. **72**, 1891–1904 (1982)]. It seems that this dispersion is due to the grain size, the standard deviation of the grain size, and the grain shape. In this study, to investigate the dispersion, the longitudinal wave velocities, the porosities, and the grain sizes of many kinds of water-saturated beach sands are measured. The relationships between the velocity, the porosity, and the grain size are obtained. From these results, it is seen that the velocity of the water-saturated beach sand with the same porosity varies with the grain size. That is, the velocity of the water-saturated beach sand with the same porosity increases, as the grain size increases. It is considered that the frame bulk modulus of the water-saturated beach sand with the same porosity varies with the grain size.

2:30

4pUW7. Tomographic measurements of sandy sediment sound speed and attenuation. Dajun Tang (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

During the SAX99, a tomographic system called the Acoustic Imager (AI) was developed to measure sound speed, attenuation and their variability in sandy sediments. It consists of 60 transducers arranged in a ring with a diameter of 29 cm placed in the horizontal plane. When the ring is pressed into the sediment at a certain depth, a set of sound speed and attenuation measurements is made. Taking data at different depths, a 3-D view of the sediment sound speed and attenuation can be obtained. In this paper, two kinds of measurements are reported. The first is measurements of sound speed at different locations over a depth of 10 cm with measurements taken every two cm. The second is a measurement at one location and one depth over a period of 48 hours to examine temporal variation of sediments. [Work supported by ONR.]

2:45

4pUW8. High frequency scattering, attenuation and permeability in sands. Tokuo Yamamoto (Geoacoustic Lab., Appl. Marine Phys. Div. RSMAS, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149), Junichi Sakakibara (Geoacoustic Lab., Taito-ku, Tokyo 111-0051, Japan), and Yoshiyuki Mohri (Natl. Inst. for Rural Eng., Tsukuba, Japan)

Imaging the permeability within sediments is one of the important problems that is not well investigated in geoacoustics. Acoustic wave energy scattered from velocity and density fluctuations within sediments increases the attenuation. However, only the intrinsic attenuation due to the Biot mechanism is related to permeability. We investigate this problem using the crosswell acoustic tomography and simple propagation measurements in water saturated sands in a large tank and a pond. The range of frequency is from 10 to 100 kHz. The sands used in the experiments are of

uniform grain sizes of 0.05 to 0.5 mm. We find that the wave propagation data follows the Biot theory. The apparent attenuation due to scattering is proportional to the velocity fluctuations and can be significant. For accurate imaging of permeability in sand volume, the removal of the apparent attenuation due to volume scattering from total attenuation is very important. [Work supported by ONR Code 3210A, NIRE, and MEST.]

3:00–3:15 Break

3:15

4pUW9. High frequency bottom scattering: Theory and experiment. Tokuo Yamamoto, Keiichi Ohkawa, and Haruhiko Yamaoka (Geoacoustic Lab., Appl. Marine Phys. Div. RSMAS, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

The exact solution of the Biot theory for the transmission of an acoustic wave at the fluid-porous seabed interface is obtained in closed form. For grazing angles larger than 5 degrees, the transmission effect on volume scattering is *ca.* 20 dB larger than that of a fluid-fluid bottom model. If grain size and porosity is known, the Biot model calculates the attenuation. A Biot scattering model is constructed by replacing the transmission coefficients and attenuation of the fluid-fluid bottom model by Yamamoto (1996). During the Shallow Water Acoustic Technology (SWAT) 2000 experiments, the 5.5 kHz 3D backscattering data from the sandy bottom at the SS site were obtained using a self-contained UM 32 channel bi-linear hydrophone array rigidly mounted on a 5 m high rigid metal tower. The Biot model extracted, sound velocity, aspect ratio, and backscatter model parameters. The Biot model is also applied to the 15 kHz 3D backscattering data from a clay bottom. For this case, the difference between the Biot model and the fluid-fluid bottom was small. It is concluded that the Exact Biot model should be used when dealing with the high frequency bottom backscattering or forward scattering from the sandy bottom. [Work supported by ONR Code 3210A.]

3:30

4pUW10. Assessing the sediment volume contribution to scattering in SAX99 sediment: Sound-speed fluctuations. Kevin B. Briggs (Naval Res. Lab., Seafloor Sci. Branch, Stennis Space Center, MS 39529-5004) and Dajun Tang (Univ. of Washington, Seattle, WA 98105)

A simulation experiment using Monte Carlo realizations has been undertaken based on real variations in sound-speed and bulk density data from SAX99 sand sediment. In this investigation, sound-speed profiles measured from cores are used to estimate the sound-speed power spectra and sound-speed correlation lengths. Due to the manner in which sound speed is measured from 6.1-cm-diam cores at 1-cm intervals, measured sound speed is an average value across the core. Thus, the values of the measurements as well as the estimates of the correlation lengths may be a function of the sampling interval and estimation of correlation lengths may be distorted. After realizations are generated from the sound-speed power spectrum and correlation length, the virtual core is sampled to obtain a sound-speed profile and then the power spectrum is re-estimated. Because actual sound speed is known from the original data in the simulation, the difference between the parameters used to generate the simulation and the virtual parameters is a measure of the distortion by averaging or smoothing. In addition, the cross-correlation between sediment bulk density and sound-speed fluctuations in the same cores is shown. [Work sponsored by ONR.]

4p THU. PM

4pUW11. Models of discrete scattering in marine sediments. Anatoliy N. Ivakin (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, ivakin@apl.washington.edu)

Acoustic scattering in marine sediments with discrete inclusions, such as shell fragments and rocks, is described in terms of the scattering cross section per unit sediment volume and the scattering cross section per unit seabed area. These cross sections are expressed through the individual scattering functions of discrete targets and statistical distributions of their parameters (size, shape, orientation, material, etc). Some simplified approaches for description of the individual scattering from irregular inclusions are considered. Rigorously, individual scattering functions should take into account the influence of the water–sediment interface. Effects which result from this influence are discussed. The role of the size distribution function is analyzed. Two types of distributions are considered. The first one is a narrow distribution with one dominating scale. The second type is a wide or multiscale distribution, for example, described by a power-law function. The frequency-angular dependencies of the seabed scattering strength for various models of discrete scattering are calculated and discussed. [Work supported by ONR, Ocean Acoustics.]

4:00

4pUW12. Modeling the sensitivity of sediment backscattering measurements to the underlying variability of the sediment bulk environmental characterizations and inhomogeneity spectra. Kevin D. LePage (Naval Res. Lab., Code 7144, 4555 Overlook Ave. SW, Washington, DC 20375), Charles W. Holland (The Penn State Univ., State College, PA 16804), and Henrik Schmidt (MIT, Cambridge, MA 02139)

A significant part of the uncertainty in reverberation predictions for shallow-water environments lies in the variability of the scattering strength associated with the bottom sediments. In this paper we evaluate how variability in the background properties of these sediments, and the characterizations of the inhomogeneity distributions within these sediments is propagated through to uncertainty in the sediment scattering strength. We also address the uncertainty in the measurement problem, i.e., how precise is the determination of the sediment bulk and inhomogeneity properties from local measurements of the time-angle evolution of bottom backscatter. We take as an example an environmental characterization obtained for a site on the Malta Plateau, south of Sicily. Here both the reflection properties of the seafloor and the time-angle evolution of the backscatter have been used to generate a notional environmental characterization of the bulk and inhomogeneity spectra of the bottom. We use the OASES volume scattering model to address the sensitivity and uncertainty issues surrounding the interpretation of this data set and to address the implications on the resulting reverberation uncertainty. [Work supported by the ONR Capturing Uncertainty DRI.]

4:15

4pUW13. Scattering of plane evanescent waves by buried cylinders: Modeling the coupling to guided waves and resonances. Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814, marston@wsu.edu)

The coupling of sound to buried targets can be associated with acoustic evanescent waves when the sea bottom is smooth. To understand the excitation of guided waves on buried fluid cylinders and shells by acoustic evanescent waves and the associated target resonances, the two-dimensional partial wave series for the scattering is found for normal incidence in an unbounded medium. The shell formulation uses the simplifications of thin-shell dynamics. The expansion of the incident wave becomes a double summation with products of modified and ordinary Bessel functions [P. L. Marston, *J. Acoust. Soc. Am.* **111**, 2378 (2002)]. Unlike the case of an ordinary incident wave, the counterpropagating par-

tial waves of the same angular order have unequal magnitudes when the incident wave is evanescent. This is a consequence of the exponential dependence of the incident wave amplitude on depth. Some consequences of this imbalance of partial-wave amplitudes are given by modifying previous ray theory for the scattering [P. L. Marston and N. H. Sun, *J. Acoust. Soc. Am.* **97**, 777–783 (1995)]. The exponential dependence of the scattering on the location of a scatterer was previously demonstrated in air [T. J. Matula and P. L. Marston, *J. Acoust. Soc. Am.* **93**, 1192–1195 (1993)].

4:30

4pUW14. Ambient noise analysis of underwater acoustic data. Mark A. Snyder, Pete Orlin, Annette Schulte (Naval Oceanogr. Office, Stennis Space Ctr., MS 39522), and Joal Newcomb (Naval Res. Lab., Stennis Space Ctr., MS 39529)

The Littoral Acoustic Demonstration Center (LADC) deployed three Environmental Acoustic Recording System (EARS) buoys in the northern Gulf of Mexico during the summers of 2001 and 2002. The buoys recorded frequencies up to 5859 Hz continuously for 36 days in 2001 and for 72 days in 2002. The acoustic signals recorded include sperm whale vocalizations, seismic airguns, and shipping traffic. The variability of the ambient noise is analyzed using spectrograms, time series, and statistical measurements. Variations in ambient noise before, during, and after tropical storm/hurricane passage are also investigated.

4:45

4pUW15. Long-term noise statistics from the Gulf of Mexico. Anthony I. Eller (Sci. Applications Intl., Inc., 1710 Saic Dr., McLean, VA 22102, anthony.i.eller@saic.com), George E. Ioup, Juliette W. Ioup, and James P. Larue (Univ. of New Orleans, New Orleans, LA 70148)

Long-term, omnidirectional acoustic noise measurements were conducted in the northeastern Gulf of Mexico during the summer of 2001. These efforts were a part of the Littoral Acoustic Demonstration Center project, Phase I. Initial looks at the noise time series, processed in standard one-third-octave bands from 10 to 5000 Hz, show noise levels that differ substantially from customary deep-water noise spectra. Contributing factors to this highly dynamic noise environment are an abundance of marine mammal emissions and various industrial noises. Results presented here address long-term temporal variability, temporal coherence times, the fluctuation spectrum, and coherence of fluctuations across the frequency spectrum. [Research supported by ONR.]

5:00

4pUW16. Underwater sound from the whale's point of view. Paul T. Arveson (Balanced Scorecard Inst., 6902 Breezewood Terrace, Rockville, MD 20852)

There have been numerous reports in the recent literature of apparently stressful effects on marine mammals due to sonar experiments. But another man-made source—the radiated noise from ships—contributes significantly to the ocean ambient, nearly everywhere and all the time. The technical basis for this talk is a set of accurate and detailed measurements of the radiated noise of a typical cargo ship [P. Arveson and D. Vendittis, “Radiated noise characteristics of a large cargo ship,” *J. Acoust. Soc. Am.* (2000)]. However, the talk will be a popular-level demonstration and a (necessarily) fictitious narrative of acoustical experiences from a humpback whale's point of view. Room acoustics permitting, the audience should be able to gain an experiential insight into the environmental impact of shipping noise on the life and habits of these creatures.

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 0R6, 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, 1 May 2003.*

Scope of S1: Standards, specifications, methods of measurement and test, and terminology in the field of physical acoustics, including architectural acoustics, electroacoustics, sonics and ultrasonics, and underwater sound, but excluding those aspects which pertain to biological safety, tolerance, and comfort.

Meeting of Accredited Standards Committee (ASC) S3 Bioacoustics

to be held jointly with the

**ANSI-Accredited U.S. Technical Advisory Group (TAG) Meeting for:
ISO/TC 108/SC 4 Human Exposure to Mechanical Vibration and Shock**

R. F. Burkard, Chair S3

Hearing Research Laboratory, State University of New York at Buffalo, 215 Parker Hall, Buffalo, New York 14214

C. Champlin, Vice Chair S3

University of Texas, Department of Communication Sciences & Disorders, CMA 2-200, Austin, Texas 78712

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

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, 1 May 2003.*

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.

Session 5aAA**Architectural Acoustics: Architectural and Electroacoustic Design of Interactive Museums and Halls of Fame**

Steven M. Haas, Chair

*Jaffe Holden Acoustics, 114A Washington Street, Norwalk, Connecticut 06854***Chair's Introduction—8:30*****Invited Papers*****8:35****5aAA1. Base-building architectural and engineering acoustic concerns unique to museums and exhibit environments.** Mark Reber (Jaffe Holden Acoust., 114A Washington St., Norwalk, CT 06854, mreber@jhacoustics.com)

Museums, especially those with intensive multimedia programs, require special attention by the base building design team to ensure that the final aural experience throughout the exhibit galleries is optimized. Often, exhibit galleries and public spaces are designed with minimal attention placed on the acoustic environment. There are countless new and old museums that suffer from excessive liveness, extremely noisy mechanical/electrical systems, and significant bleed between multiple sound sources. The author will describe the importance of addressing the above issues, how the base building design team interfaces with the exhibit design team, and present some case study examples of how improperly designed museum buildings could have a detrimental impact on the acoustic function of the exhibit experience.

8:55**5aAA2. Shape optimization: Good looks and acoustics too!** Peter D'Antonio (RPG Diffusor Systems, Inc., 651-C Commerce Dr., Upper Marlboro, MD 20772, pdantonio@rpginc.com), Trevor J. Cox (Univ. of Salford, Salford M5 4WT, UK), and Steve Haas (Jaffe Holden Acoust., Norwalk, CT 06854)

One of the challenges in the architectural acoustic design of museums and other public spaces is to develop contemporary scattering surfaces that complement contemporary architecture in the way that statuary, coffered ceilings, columns, and relief ornamentation complemented classic architecture. Often acoustic surfaces satisfy the acoustics, but may or may not satisfy the aesthetics. One approach that has been successful employs a combination of boundary element and multidimensional optimization techniques. The architect supplies the desired shape motif and the acoustician supplies the acoustical performance requirements. The optimization program then provides an Arcousthetic surface, which simultaneously satisfies the architecture, the acoustics, and the aesthetics. The program can be used with diffusive or difforsorptive surfaces. Photos of installations using these acoustic tools and a description of the design of the National Museum of the American Indian will also be presented to illustrate the usefulness of these devices and their impact on architectural acoustics.

9:15**5aAA3. Acoustic and electro-acoustic design of the Country Music Hall of Fame and Museum in Nashville.** Steve Haas (Jaffe Holden Acoust., 114A Washington St., Norwalk, CT 06854, shaas@jhacoustics.com)

The Country Music Hall of Fame and Museum in Nashville is the newest major music museum to open in the United States. It represents the latest in acoustic and audio delivery technology to provide excellent sound quality and sound control for museum public, performance and exhibit spaces. The author will present some of the unique approaches and technologies that were first implemented into this facility and explain how the acoustic/electroacoustic interface allowed the Hall of Fame to maximize their use of sound in the exhibits as well as hold simultaneous special functions in many of the public spaces.

9:35**5aAA4. A new seamless, smooth, interior, absorptive finishing system.** Peter D'Antonio (RPG Diffusor Systems, Inc., 651-C Commerce Dr., Upper Marlboro, MD 20772, pdantonio@rpginc.com)

Architects and acousticians have sought a field-applied, absorptive finishing system that resembles a smooth plaster or painted drywall surface, since the dawn of architectural acoustics. Some success has been achieved using sprayed cellulose or cementitious materials, but surface smoothness has been a challenge. A new approach utilizing a thin microporous layer of mineral particles applied over a mineral wool panel will be described. This material can be applied to almost any shape surface, internally pigmented to match almost any color and renovated. It is currently finding application in many architectural applications, including museums. A recent installation in the New Pinakothek Museum in Munich will be illustrated.

9:55

5aAA5. Coupled auralization and virtual video for immersive multimedia displays. Paul D. Henderson, Rendell R. Torres, Yasushi Shimizu (Prog. in Architectural Acoust., Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, hendep2@rpi.edu), Richard Radke, and Brian Lonsway (Rensselaer Polytechnic Inst., Troy, NY 12180)

The implementation of maximally-immersive interactive multimedia in exhibit spaces requires not only the presentation of realistic visual imagery but also the creation of a perceptually accurate aural experience. While conventional implementations treat audio and video problems as essentially independent, this research seeks to couple the visual sensory information with dynamic auralization in order to enhance perceptual ac-

curacy. An implemented system has been developed for integrating accurate auralizations with virtual video techniques for both interactive presentation and multi-way communication. The current system utilizes a multi-channel loudspeaker array and real-time signal processing techniques for synthesizing the direct sound, early reflections, and reverberant field excited by a moving sound source whose path may be interactively defined in real-time or derived from coupled video tracking data. In this implementation, any virtual acoustic environment may be synthesized and presented in a perceptually-accurate fashion to many participants over a large listening and viewing area. Subject tests support the hypothesis that the cross-modal coupling of aural and visual displays significantly affects perceptual localization accuracy.

FRIDAY MORNING, 2 MAY 2003

ROOM 208, 9:00 TO 11:00 A.M.

Session 5aBB

Biomedical Ultrasound/Bioresponse to Vibration: General Topics in Biomedical Ultrasound

Charles C. Church, Chair

National Center for Physical Acoustics, University of Mississippi, 1 Coliseum Drive, University, Mississippi 38677

Contributed Papers

9:00

5aBB1. Comparison of ultrasound-induced bioeffects in glass catfish after injection with optison and liquid perfluorocarbon droplets. Subha Maruvada and Kullervo Hynynen (Focused Ultrasound Group, Dept. of Radiol., Harvard Med. School and Brigham and Women's Hospital, 221 Longwood Ave., Boston, MA 02115)

This work is an investigation of ultrasound-induced bioeffects *in vivo*. Glass catfish were used for these experiments because they are optically transparent. Anaesthetized fish were injected with either optison (OPT) or liquid perfluorocarbon droplets (LPD), using microinjection techniques. Shortly after injection, the fish were insonified with one of two single element focused transducers (1.091 MHz 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 200x magnification. The entire interaction was videotaped and digitized. The fish were insonified at power levels between 1–80 W, which translated into acoustic pressures from 0.45–15 MPa. Sonications were pulsed with burst lengths of 10 ms and 100 ms and a repetition frequency of 1 Hz. The entire length of one sonication at a specific pressure was 20 seconds. The effects of the sonication were analyzed at each pressure level. The ultrasound-induced bioeffects due to OPT and LPD were compared. Threshold values for damage were lower after OPT injection than after LPD injection, especially at lower frequencies.

9:15

5aBB2. Acoustic characterization of liposome suspensions and Optison® at 3.5 MHz. Constantin-C. Coussios^{a)}, Christy K. Holland, Ludwika Jakubowska (Dept. of Biomed. Eng., Univ. of Cincinnati, 231 A. Sabin Way, Cincinnati, OH 45267-0586, coussios@bu.edu), Shao-Ling Huang, Robert C. MacDonald (Northwestern Univ., Evanston, IL), Ashwin Nagaraj, and David D. McPherson (Northwestern Univ., Chicago, IL)

Liposomes—phospholipid vesicles that can encapsulate both gas and fluid—are currently being developed as contrast agents and vehicles for targeted drug delivery. The backscattering coefficient and attenuation of liposome suspensions were measured at 3.5 MHz as a function of lipo-

some concentration and compared to those of Optison®. The backscattering coefficient of liposomes at concentrations in excess of $1.15 \times 10^8/\text{ml}$ exceeds that of Optison® at the clinical dose ($5.0\text{--}8.0 \times 10^4/\text{ml}$). On the other hand, the attenuation of ultrasound through liposome suspensions is negligible relative to the attenuation through Optison® (0.25 dB/cm), suggesting that liposomes have a much higher scatter-to-attenuation ratio and could be more efficient as contrast agents. A theoretical model that treats the gas within the liposomes as a free air bubble was used to estimate the backscattering coefficient and correlated well with experimental results. Church's model [J. Acoust. Soc. Am. **97**, 1510–1521 (1995)] for Albunex® was also found to provide an accurate estimate of the backscattering coefficient of Optison®. These models suggest that the marked difference in attenuation between the two agents can be attributed to the fact that liposomes are sub-resonant in size at 3.5 MHz, but Optison® microspheres act as resonant scatterers. ^{a)}Currently at Boston University.

9:30

5aBB3. Parametric mechanism of subharmonic generation by the microbubble contrast agents. Vladimir Genis (Drexel Univ., 3141 Chestnut St., Philadelphia, PA 19104), Igor Oboznenko (Kyiv Polytechnic Univ., Kyiv 03057, Ukraine), and Dat Tran (Nagoya Univ., 464-8603 Nagoya, Japan)

Contrast agents, consisting of gas-filled microbubbles, are widely used in biomedical ultrasound for enhancing the capabilities of ultrasound imaging. The behavior of contrast agents has been shown to lead to the generation of harmonic and subharmonic components in the backscattered signal. The mechanism of subharmonic generation is not quite understood. It is known that the oscillation of gas-filled bubbles is described by nonlinear equations. In this work, the parametric approach to solving this problem is introduced, where the nonlinear equation is substituted by the linear equation with time-varying parameters, such as flexibility of the bubble. The equation describing the oscillation of the bubble is reduced to the Mathieu equation, and conditions for unstable solutions are presented. It has been demonstrated that the regions of instability occur in the vicinity of the frequencies $\omega = 1/2 \omega_0$, $\omega = 3/2 \omega_0$, etc., which correspond to gen-

eration of the first and higher subharmonics. Criteria for determining the relationship between the acoustic pressure, required for subharmonic generation, and the radius of the air-filled bubble are also determined.

9:45

5aBB4. Experimental and numerical investigation of tissue harmonic imaging (THI). Yuan Jing, Xinmai Yang, and Robin O. Cleveland (Dept. of Aersp. and Mech. Eng., Boston Univ., Boston, MA 02215, yuanjing@bu.edu)

In THI the probing ultrasonic pulse has enough amplitude that it undergoes nonlinear distortion and energy shifts from the fundamental frequency of the pulse into its higher harmonics. Images generated from the second harmonic (SH) have superior quality to the images formed from the fundamental frequency. Experiments with a single element focused ultrasound transducer were used to compare a line target embedded in a tissue phantom using either fundamental or SH imaging. SH imaging showed an improvement in both the axial resolution (0.70 mm vs 0.92 mm) and the lateral resolution (1.02 mm vs 2.70 mm) of the target. In addition, the contrast-to-tissue ratio of the target was 2 dB higher with SH imaging. A three-dimensional model of the forward propagation has been developed to simulate the experimental system. The model is based on a time-domain code for solving the KZK equation and accounts for arbitrary spatial variations in all tissue properties. The code was used to determine the impact of a nearfield layer of fat on the fundamental and second harmonic signals. For a 15 mm thick layer the SH side-lobes remained the same but the fundamental side-lobes increased by 2 dB. [Work supported by the NSF through the Center for Subsurface Sensing and Imaging Systems.]

10:00

5aBB5. Simulation and experiment of nonlinear scattering mechanisms producing difference frequencies from two distinct primary frequencies incident on a target. Jonathan S. Thierman (MIT-Harvard HST Prog., Cambridge, MA 02139, thierman@mit.edu), Nicholas C. Makris, Purnima Ratilal (MIT, Cambridge, MA 02139), and Kullervo Hynynen (Harvard Med. School, Brigham and Women's Hospital, Boston, MA 02115)

Our theoretical derivation of nonlinear wave propagation with scatterers is compared with simulated and experimental results regarding the production of difference frequency signals from targets insonified by sources radiating at two distinct primary frequencies. The previously presented nonlinear wave equation derived from first principles using coherent wave theory is applied to the problem of focused ultrasound transducers and various target types and geometries. First, with computer simulation, we explore the relative contribution of each term of the theoretical derivation as applied to the geometry of focused sources. The terms of this equation represent different physical mechanisms responsible for the production of a difference frequency. Their relative magnitude reveals which mechanisms are more likely to produce the measured difference frequency signal in experiments. Next, we measure the difference frequency field in experiments to demonstrate the correlation of theory with simulation and experiment. The experiments reveal the profound effect various scattering phenomenon have on the produced difference frequency field.

10:15

5aBB6. High-frequency characterization of tissue phantoms using a scanning acoustic microscope (SAM). Caleb H. Farny, Tara C. Chklovski, T. L. Szabo, Ron A. Roy, and Robin O. Cleveland (Dept. of Aersp. and Mech. Eng., 110 Cummington St., Boston, MA 02048)

A SAM has been developed to measure the acoustic properties of samples. The SAM consists of a wideband (35–75 MHz) ultrasound transducer operating in pulse-echo mode. Waveforms measured from the front

and back surfaces of a uniform sample are used to determine the impedance, sound-speed, and attenuation of the sample as a function of frequency. The principle of causality is used to ensure that the sound-speed dispersion and attenuation measurements are consistent. The transducer is scanned over the surface of the sample in a raster scan and a two-dimensional map of the acoustic properties averaged over depth can be obtained. Scans of high-density polyethylene show agreement with previously published properties. The SAM was used to evaluate the high-frequency acoustic properties of a tissue phantom composed of agar, gelatin and silica. The silica content was varied from 2% to 10% and the impact of the silica scattering particles on the acoustic properties was evaluated. The sound-speed showed almost no dependence on silica content but the attenuation at 30 MHz increased from 8.4 to 23.8 dB/cm. This phantom has the potential to be used to mimic the walls of blood vessels. [Work supported by the NSF through the Center for Subsurface Sensing and Imaging Systems.]

10:30

5aBB7. Ultrasound phase-contrast transmission imaging of localized thermal variation in a breast tissue model. Greg Clement and Kullervo Hynynen (Dept. of Radiol., Harvard Med. School, Brigham and Women's Hospital, Boston, MA 02115)

The present numeric study examines an ultrasound phase contrast method for imaging localized temperature rises in tissue, such as heating experienced in thermal ablation treatments. We tested the method's ability to produce thermal images by simulating the signal from a planar ultrasound source directed through tissue containing skin, fat, and muscle layers. Phase-contrast imaging only requires phase shifts a fraction of a wavelength for detection, thus we tested for the feasibility of imaging at low (submegahertz) frequencies, allowing greater depth penetration and reduced scattering away from the image plane. The tradeoff for lower frequencies was a reduced spatial resolution. It was therefore necessary to determine the lowest possible frequency that could still provide information about the tissue structure and information about the temperature rise in the heated volume. After studying a tomographically reconstructed 4-mm cylindrical heated region in a 100-mm-thick section breast tissue model, thermal effects were detected at frequencies as low as 0.4 MHz, which allowed reconstruction resolution of about 2 mm² over the image plane. This method may have applications for example in controlling focused ultrasound surgery of breast.

10:45

5aBB8. A voltage-doubling (2-kV) pulser for biomedical ultrasound measurements. Paul M. Gammell (Gammell Appl. Technologies, 6139 Pleasant Cove Dr., Exmore, VA 23350, pgammell@ieee.org) and Gerald R. Harris (Food and Drug Administration, Rockville, MD 20850)

High-voltage pulsers offer advantages in some biomedical ultrasound measurement applications, such as driving focused transducers for bioeffects studies, or transiently exciting plane, thick transducers for broadband measurements of attenuation or frequency response of receiving transducers. Because the duty cycle is low (typically less than 0.1%), bulky, expensive pulsers having large temporal-average power capability are not required. In the present design integrated gate bipolar transistors (IGBTs) are used to energize the ultrasound transducer. This solid-state switching device can handle large currents, has a low on-resistance, and can be turned off as well as on by its gate signal. However, widely available devices are limited to 1200-V peak. Therefore, the present design charges two capacitors in parallel and then uses two IGBTs to connect them in series to produce the output pulse. With a 1200-V supply this approach produces a pulse of 2100 V into a 50- Ω load at a 200-Hz repetition rate. The length of the output pulse is determined by the controlling input TTL pulse unless small capacitors are used, in which case a decaying exponential pulse is obtained. This design is readily reproduced and the components, exclusive of the power supply, cost less than \$100.

Session 5aPP**Psychological and Physiological Acoustics and Animal Bioacoustics: Ecological Psychoacoustics**

John G. Neuhoff, Chair

*Psychology Department, The College of Wooster, Wooster, Ohio 44691***Chair's Introduction—8:30*****Invited Papers*****8:35**

5aPP1. An assessment tool for auditory realism. Stephen Lakatos (Washington State Univ., 14204 NE Salmon Creek Ave., Vancouver, WA 98686), Perry R. Cook (Dept. of Computer Sci., Princeton Univ., 35 Olden St., Princeton, NJ 08544, prc@cs.princeton.edu), and Colin R. Harbke (Washington State Univ., Vancouver, WA 98686)

An item-based survey was developed to evaluate how effectively physically-based models can yield sounds that appear realistic to the ear. Based on past research and current insights, realism as a construct was hypothesized to comprise eight orthogonal factors: detail, physical plausibility (biological and nonbiological), temporal consistency, vividness, presence, and whether a sound evokes sensory images and associated memories. Sets of 8–9 declarative statements were selected by committee for each factor from large initial lists of such statements. Another committee selected weak, moderate and strong sound exemplars for each factor from digital libraries that were felt to vary uniquely along that factor. In an exploratory phase, 82 participants rated how representative each sound exemplar was of each statement. Based on factor analyses, the three highest loading statements for each factor were compared across sound exemplars; statistically significant differences were found between exemplars for all factors. A subsequent confirmatory phase used new sounds, coupled with synthetic versions from physically-based simulations, to validate the proposed realism factors. Data from an additional 200 participants ratings for these stimuli supported the surveys construct validity. The survey should benefit sound engineers wishing to evaluate the degree to which their synthesis algorithms generate realistic output.

9:00

5aPP2. Listening to sound patterns as a dynamic activity. Mari Riess Jones (Dept. of Psych., The Ohio State Univ., Columbus, OH 43210)

The act of listening to a series of sounds created by some natural event is described as involving an entrainmentlike process that transpires in real time. Some aspects of this dynamic process are suggested. In particular, real-time attending is described in terms of an adaptive synchronization activity that permits a listener to target attending energy to forthcoming elements within an acoustical pattern (e.g., music, speech, etc.). Also described are several experiments that illustrate features of this approach as it applies to attending to musiclike patterns. These involve listeners' responses to changes in either the timing or the pitch structure (or both) of various acoustical sequences.

9:25

5aPP3. Sex, acceleration, brain imaging, and rhesus monkeys: Converging evidence for an evolutionary bias for looming auditory motion. John G. Neuhoff (Dept. of Psych., The College of Wooster, Wooster, OH 44691, jneuhoff@wooster.edu)

Increasing acoustic intensity is a primary cue to looming auditory motion. Perceptual overestimation of increasing intensity could provide an evolutionary selective advantage by specifying that an approaching sound source is closer than actual, thus affording advanced warning and more time than expected to prepare for the arrival of the source. Here, multiple lines of converging evidence for this evolutionary hypothesis are presented. First, it is shown that intensity change specifying accelerating source approach changes in loudness more than equivalent intensity change specifying decelerating source approach. Second, consistent with evolutionary hunter-gatherer theories of sex-specific spatial abilities, it is shown that females have a significantly larger bias for rising intensity than males. Third, using functional magnetic resonance imaging in conjunction with approaching and receding auditory motion, it is shown that approaching sources preferentially activate a specific neural network responsible for attention allocation, motor planning, and translating perception into action. Finally, it is shown that rhesus monkeys also exhibit a rising intensity bias by orienting longer to looming tones than to receding tones. Together these results illustrate an adaptive perceptual bias that has evolved because it provides a selective advantage in processing looming acoustic sources. [Work supported by NSF and CDC.]

10:05–10:20 Break

9:50

5aPP4. Statistical evidence that musical universals derive from the acoustic characteristics of human speech. David Schwartz, Catharine Howe, and Dale Purves (Dept. of Neurobiology, Duke Univ., 101 Bryan Res. Bldg., Durham, NC 27710)

Listeners of all ages and societies produce a similar consonance ordering of chromatic scale tone combinations. Despite intense interest in this perceptual phenomenon over several millennia, it has no generally accepted explanation in physical, psychological, or physiological terms. Here we show that the musical universal of consonance ordering can be understood in terms of the statistical relationship between a pattern of sound pressure at the ear and the possible generative sources of the acoustic energy pattern. Since human speech is the principal naturally occurring source of tone-evoking (i.e., periodic) sound energy for human listeners, we obtained normalized spectra from more than 100 000 recorded speech segments. The probability distribution of amplitude/frequency combinations derived from these spectra predicts both the fundamental frequency ratios that define the chromatic scale intervals and the consonance ordering of chromatic scale tone combinations. We suggest that these observations reveal the statistical character of the perceptual process by which the auditory system guides biologically successful behavior in response to inherently ambiguous sound stimuli.

10:20

5aPP5. An avoidance behavior model for migrating whale populations. John R. Buck (ECE Dept. and SMAST, UMass Dartmouth, 285 Old Westport Rd., North Dartmouth, MA 02747) and Peter L. Tyack (Woods Hole Ocean. Inst., Woods Hole, MA 02543)

A new model is presented for the avoidance behavior of migrating marine mammals in the presence of a noise stimulus. This model assumes that each whale will adjust its movement pattern near a sound source to maintain its exposure below its own individually specific maximum received sound-pressure level, called its avoidance threshold. The probability distribution function (PDF) of this avoidance threshold across individuals characterizes the migrating population. The avoidance threshold PDF may be estimated by comparing the distribution of migrating whales during playback and control conditions at their closest point of approach to the sound source. The proposed model was applied to the January 1998 experiment which placed a single acoustic source from the U.S. Navy SURTASS-LFA system in the migration corridor of grey whales off the California coast. This analysis found that the median avoidance threshold for this migrating grey whale population was 135 dB, with 90% confidence that the median threshold was within ± 3 dB of this value. This value is less than the 141 dB value for 50% avoidance obtained when the 1984 "Probability of Avoidance" model of Malme *et al.*'s was applied to the same data. [Work supported by ONR.]

10:35

5aPP6. From acoustics to perception: How to listen to meaningful sounds in a meaningful way. Brian Gygi (East Bay Inst. for Res. and Education, 150 Muir Rd., Martinez, CA 94553)

Making sense of the sounds we encounter everyday requires a coupling between what we sense (acoustics) and what we "hear" (perception). Data will be presented suggesting that the psychological organization of familiar nonspeech, nonmusical sounds ("environmental sounds") is closely aligned to the acoustic similarity of these sounds, as well as to the physical events that produced the sounds. Results of identification studies indicate that the particular acoustic features listened for include pitch salience, rhythmicity, and amplitude modulation. Far from being

impooverished, the acoustic information available in the world is sufficient to quickly and reliably determine the source of a sound and allow for appropriate responses.

10:50

5aPP7. Phylogenetic perspectives on noise-induced fear and annoyance. Ann Bowles (2595 Ingraham St., San Diego, CA 92109, annb1@san.rr.com)

Negative human responses to noise are typically interpreted in terms of human psychological, cognitive, or social processes. However, it may be useful to frame hypotheses about human responses in terms of evolutionary history, during which negative responses have been part of a suite of adaptations to a variable sound environment. By comparing the responses of a range of nonhuman animals to various types of noise, it is possible to develop hypotheses about the ecology of human responses. Examples of noise-related phenomena that could be explained usefully from this perspective include the Schulz curve, noise-induced physical stress, acute fear responses induced by transient noise, and the relationship between temperament and noise-induced annoyance. Responses of animals from a range of taxa will be described and their behavior interpreted in terms of their life-history strategies. With this perspective, some testable hypotheses about noise-induced fear and annoyance will be suggested.

11:05

5aPP8. Perturbation analysis of acoustic cues for sound source identification. Robert A. Lutfi (Waisman Ctr., Univ. of Wisconsin, Madison, WI 53706)

Perturbation analysis has enjoyed recent success in vision as a means of revealing decision processes underlying object identification [J. Vision 2, 1–131 (2002)]. Here it is applied to auditory identification of elementary sound sources and their attributes. The analysis proceeds in two stages: First, listener decision strategy is determined from regression weights relating listener judgments to lawful perturbations in acoustic parameters from theoretical acoustics. Second, limits in processing are identified by comparing the obtained weights to those of a hypothetical observer who bases decisions on acoustic information intrinsic to the source attribute(s) being judged. The approach represents a significant advance over past methods that infer decision strategy from performance accuracy or from the effect of placing acoustic cues in unlawful opposition. Initial studies are reviewed revealing a complex pattern of results in which decision strategy is based on intrinsic acoustic information in some conditions but not in others, and in which there are sometimes individual differences in decision strategy that yield the same level of identification performance. [Work supported by NIDCD.]

11:20

5aPP9. Identifying the sources of environmental sounds with a varying number of spectral channels. Valeriy Shafiro, James J. Jenkins, and Winifred Strange (Speech & Hearing Dept., CUNY Grad. Ctr., 365 Fifth Ave., New York, NY 10016, vshafiro@gc.cuny.edu)

Normally hearing adult listeners were asked to identify the sources of 60 familiar environmental sounds processed through a vocoder simulation of a cochlear implant with a varying number of frequency channels. Using a Latin square design, listeners heard 10 different sounds in one of the six channel conditions (2, 4, 8, 16, 24, 32) followed by all 60 unprocessed sounds. For each sound, listeners selected one of 60 response options that best described the sound's source. Preliminary results indicate that (1) source identification accuracy markedly improves up to 16 channels, but does not continue to improve with further increases in the number of channels; (2) individual sounds vary in the number of channels needed for accurate source perception; (3) increasing the number of channels positively affects source identification of some sounds, while it negatively

affects others; (4) listeners show different response biases when listening to low versus high number-of-channels stimuli; (5) systematic confusion can be found in listeners' erroneous responses to stimuli. These findings highlight differences among spectral and temporal factors important for source identification and suggest ways of classifying environmental sounds based on spectral resolution required for source identification. [Work supported by NIDCD.]

11:35

5aPP10. From sound to size: A research in ecological acoustics. Massimo Grassi (Dept. of Gen. Psych., Univ. of Padova, Via Venezia 8, Padova 35133, Italy, grassi@psy.unipd.it)

Two experiments investigated the perception of the size of a solid ball from the sound it produces when dropped upon a plate (i.e., object's vibration contributing only slightly to the resulting pressure wave). Experi-

ment one: listeners estimated metrically the size of seven balls when dropped upon a 215-mm-diam plate. Experiment two: a new group of listeners performed the same scaling, however, balls could be dropped either upon the 215-mm-diam plate or upon a 185-mm-diam plate. In both experiments listeners had no foregoing information about the sound-producing event. Overall, listeners' estimations matched veridically the actual size of the balls. However, in the second experiment, listeners' estimations were dependent on the dimension of the plate: balls were judged slightly larger when dropped on the largest plate. Results were analyzed in order to understand whether physical or acoustical features of the sound-producing event could predict the performance: the actual size of the ball resulted in the best predictor for all estimations in both experiments. Nonetheless, the difference in spectral centroid of the sound produced by the two plates could explain the different estimations obtained in the second experiment. [Work supported by European Commission, Contract No. IST-2000-25287.]

FRIDAY MORNING, 2 MAY 2003

ROOM 206, 8:00 A.M. TO 12:00 NOON

Session 5aSC

Speech Communication: From Phonemes to Models (Poster Session)

Jody Kreiman, Chair

Head and Neck Surgery, University of California, Los Angeles, School of Medicine, 1000 Veteran Avenue, Los Angeles, California 90095-1794

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

5aSC1. Identifying intrinsic constituents of focus through "imitation via restoration." Yi Xu (The Univ. of Chicago, 1126 E. 59th St. SS4, Chicago, IL 60637, xuyi@uchicago.edu), Ching X. Xu, and Xuejing Sun (Northwestern Univ., Evanston, IL 60208)

In this study we test the hypothesis that although certain parts of an observed intonation may seem dispensable in perception tests, they nevertheless are consistently produced by speakers. We refer to all consistently produced parts of an intonation as its "intrinsic constituents." To identify the intrinsic constituents, we developed an experimental paradigm called "imitation via restoration." In this paradigm, the intonation under scrutiny is first recorded by a native speaker. Then words carrying a potential constituent of the intonation are replaced by a loud noise. During the experiment, the sentence containing the replacement noise is presented to the subjects together with the text. The subjects' task is to repeat the sentence in exactly the same way as they hear it. The consistency with which subjects restore the missing parts of the target intonation would therefore provide a reasonable indication as to which of them are truly intrinsic to the intonation. Our first such experiment was conducted on determining whether focus consists of only on-focus pitch range expansion or it also involves obligatory post-focus pitch range suppression. Eight native speakers of Beijing Mandarin participated as subjects. Preliminary results have provided supporting evidence for the dual-component hypothesis.

5aSC2. Recognizing focus in noise filled sentences. Ching X. Xu (Dept. of Commun. Sci. and Disord., Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208) and Yi Xu (Univ. of Chicago, Chicago, IL 60637)

This study is designed to help identify the intrinsic constituents of focus. Twelve four-word Mandarin sentences were recorded by a native speaker five times, each time either with focus on one of the words, or without any focus. Then, one, two or three words in each sentence produced by the speaker were replaced by pink noise. The noise-filled sentences were presented to subjects along with the text. The subjects' task was to determine if the sentence had a focus, and if yes, on which word. Ten native Mandarin speakers participated as subjects. Their performance was compared across noise replacement conditions. It was found that, when both on-focus and post-focus words were present, focus could be recognized consistently. When only the focused word was present, focus could be recognized fairly well unless the focus position was sentence final, in which case it was not very distinct from no focus. When post-focus word(s) was(were) left intact while focused words were replaced by noise, focus could still be detected successfully, but its exact localization was sometimes judged wrong. These results seem to support the dual-component hypothesis about focus. Further implications of the findings will be discussed.

5aSC3. VOT of American English stops with prosodic correlates.

Mark Van Dam (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405)

It is known that in a word-medial bisyllabic sequence, a stop beginning a stressed syllable (e.g., /p/ in *póodle*) has a much longer voice-onset time (VOT) than following an unstressed syllable (e.g., /p/ in *stripper*). The current study investigates VOT of stops in unstressed syllables not immediately following the primary stressed syllable in the word, e.g., before the primary word stress (the /p/ in *Winnepegósis*, the /k/ in *baccaláureate*), and after the primary word stress (the /p/ in *Óedipus*, the /k/ in *américan*). A set of 42 test words were recorded in carrier phrases spoken by 4 native speakers of American English. In bisyllabic words, word initial or stressed-syllable initial stops (e.g., *póodle*, *appéal*) averaged 63 ms while following unstressed stops (e.g., *stripper*) averaged 25 ms. Stops in unstressed syllables preceding the primary stressed syllable (e.g., *Winnepegósis*) averaged 32 ms while those stops in following unstressed syllables (e.g., *Óedipus*) averaged 41 ms. To account for these results, we need at least four VOT levels. Also interesting is the interaction between stop place /p/ vs /k/ and position: *Óedipus* has the same VOT as *américan* while *Winnepegósis* and *baccaláureate* differ significantly.

5aSC4. The influence of phonetic context and formant measurement location on acoustic vowel space.

Greg S. Turner, David T. Hutchings, Betsy Sylvester (Dept. of Commun. Disord., Central Missouri State Univ., Warrensburg, MO 64093, turner@cmsu1.cmsu.edu), and Gary Weismer (Univ. of Wisconsin–Madison, Madison, WI 53705-2280)

One way of depicting vowel production is by describing vowels within an $F1/F2$ acoustic vowel space. This acoustic measure illustrates the dispersion of $F1$ and $F2$ values at a specific moment in time (e.g., the temporal midpoint of a vowel) for the vowels of a given language. This measure has recently been used to portray vowel production in individuals with communication disorders such as dysarthria and is moderately related to the severity of the speech disorder. Studies aimed at identifying influential factors effecting measurement stability of vowel space have yet to be completed. The focus of the present study is to evaluate the influence of phonetic context and spectral measurement location on vowel space in a group of neurologically normal American English speakers. For this study, vowel space was defined in terms of the dispersion of the four corner vowels produced within a CVC syllable frame, where C includes six stop consonants in all possible combinations with each vowel. Spectral measures were made at the midpoint and formant extremes of the vowels. A discussion will focus on individual and group variation in vowel space as a function of phonetic context and temporal measurement location.

5aSC5. On the relation of apparent naturalness to phonetic perceptual identification.

Robert E. Remez, Cynthia Y. Yang (Dept. of Psych., Barnard College, 3009 Broadway, New York, NY 10027), Rebecca L. Piorowski (Johns Hopkins Univ., Baltimore, MD 21218), Stephanie Wissig, Abigail Batchelder, Hedy Nam (Barnard College, 3009 Broadway, New York, NY 10027), and Anne E. Timberlake (Oberlin College, Oberlin, OH 44074)

A set of synthetic test syllables was created varying in the attributes of apparent naturalness and phonetic place-of-articulation. These acoustic items were used in a new test of the relation of naturalness to phonetic perceptual resolution. An earlier study had found that variation in the form of the excitation of a formant pattern greatly affected impressions of naturalness while minimally affecting the resolution of the phonetic place. The present study assessed the relation of naturalness and intelligibility when the bandwidths of formant centers were varied. We conducted a naturalness tournament composed of items drawn from the test series; and, we assayed the sensitivity of perceivers to place contrasts by estimating the cumulative d' across the series in identification tests. Unlike our prior observations, these new outcomes reveal both a narrower range of naturalness variation and a close relation in findings of the naturalness tournament and the measures of phonetic sensitivity. Together, this pair of stud-

ies shows that intelligibility and naturalness can be either orthogonal or contingent aspects of speech perception. These measures offer a tool to understand rule-based and exemplar-based components of phonetic perception. [Research supported by NIDCD.]

5aSC6. Difference limens for vocal aperiodicities.

Jody Kreiman and Bruce R. Gerratt (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab. Ctr., 1000 Veteran, Los Angeles, CA 90095-1794, jkreiman@ucla.edu)

Although jitter, shimmer, and noise characterize all voice signals, recent data from a method of adjustment task suggest that these acoustic attributes do not all have equal perceptual importance. Listeners did not agree well with one another when asked to match the amounts of jitter or shimmer in a voice, although they did agree well in their matching responses to noise. Several explanations for this finding are possible. Listeners may be insensitive to differences in amounts of jitter and shimmer in a voice, so that difference limens are large relative to customary measurement resolution; or listeners may have difficulty separating jitter and shimmer perceptually from the composite noise component. To investigate these issues, we synthesized several series of stimuli with and without noise, and differing by steps in the amounts of jitter or shimmer present. Listeners heard pairs of these stimuli and were asked to respond whether stimuli were the same or different. Receiver operating characteristic analysis was applied to determine how listener sensitivity to jitter and shimmer varies. [Research supported by NIDCD.]

5aSC7. Perceptual differences in source–filter characteristics of racially affiliated dialects of American English.

Thomas Purnell and Laura Kopplin (Dept. of Linguist., Univ. of Wisconsin, Madison, 1168 Van Hise, 1220 Linden Dr., Madison, WI 53706, tpurnell@facstaff.wisc.edu)

Three forced-choice experiments are described investigating perceptual cues of tokens simulating three racially-affiliated varieties of American English, namely African American English (AE), Chicano English (CE), and General American English (GE). These experiments test whether the locus of the conceptual assignment of pseudo-tokens to specific socially constructed groups resides primarily in either source or filter characteristics of the tokens. In the first experiment, responses to real tokens matched the implied dialect better than chance. The subsequent experiments used synthetic versions of the token with the greatest percent match for each group. For the second experiment the source and filter were separated using LPC filtering, then swapped across groups. Two mirror sets of tokens were produced where $F0$ was fixed and $F0$ was absent (whispered). Findings provide evidence that cues for token identification are primarily filter oriented, although for CE and GE source cues also play a minor perceptual role. To test the finding that the filter plays a greater perceptual role than the source, the same three sets of tokens were pass band masked by noise. Tokens pass band filtered for the 400 Hz to 2,000 Hz frequency range were perceived less well than other filtered tokens.

5aSC8. Appalachia meets the mid-south in Tennessee.

Richard S. McGowan and Michel T-T. Jackson (CRess LLC, 1 Seaborn Pl., Lexington, MA 02420, rsmcgowan@earthlink.net)

The father of bluegrass and long-time member of the Opry in Nashville, Bill Monroe, had a number of musical influences, including blues, gospel, and the American versions of Anglo-Scots-Irish folk music endemic to Appalachia. (Although Bill Monroe was from western Kentucky, the Appalachian influence is apparent.) In 1946 he recorded a song for Columbia Records that he had written. This song was recorded again in 1954 at Sun Studios as a “B side” by the future king of rock and roll, Elvis Presley. Raised in East Tupelo, MS and Memphis, Elvis’ music derived from the mid-south’s blues and gospel music, but with a peculiar

honkytonk-informed, Memphis style. We will compare these artists through this common song and their spoken interviews in terms of voice quality and English dialect. Some copy-synthesis of these artists will be attempted in our search for the high, lonesome and rock and roll sounds.

5aSC9. Talker and accent variability effects on spoken word recognition. Edna E. Nyang, Catherine L. Rogers (Dept. of Commun. Sci. & Dis., Univ. of South Florida, 4202 E. Fowler Ave., PCD1017, Tampa, FL 33620), and Kanae Nishi (Speech and Hearing Sci., City Univ. of New York, New York, NY 10016)

A number of studies have shown that words in a list are recognized less accurately in noise and with longer response latencies when they are spoken by multiple talkers, rather than a single talker. These results have been interpreted as support for an exemplar-based model of speech perception, in which it is assumed that detailed information regarding the speaker's voice is preserved in memory and used in recognition, rather than being eliminated via normalization. In the present study, the effects of varying both accent and talker are investigated using lists of words spoken by (a) a single native English speaker, (b) six native English speakers, (c) three native English speakers and three Japanese-accented English speakers. Twelve /hVd/ words were mixed with multi-speaker babble at three signal-to-noise ratios (+10, +5, and 0 dB) to create the word lists. Native English-speaking listeners' percent-correct recognition for words produced by native English speakers across the three talker conditions (single talker native, multi-talker native, and multi-talker mixed native and non-native) and three signal-to-noise ratios will be compared to determine whether sources of speaker variability other than voice alone add to the processing demands imposed by simple (i.e., single accent) speaker variability in spoken word recognition.

5aSC10. Acoustic correlates of Georgian stops. Tamra M. Wysocki (Dept. of Linguist., Univ. of Chicago, 1050 E. 59th St., Chicago, IL 60637)

This paper presents results from an acoustic analysis of Georgian stops. Georgian is a South Caucasian language and has a three-way opposition among voiced, voiceless aspirated, and ejective stops for three places of articulation: bilabial, dental, and velar. Tokens consist of initial and medial stops in isolated words produced by two male and two female native Georgian speakers. Closure duration, VOT, and burst amplitude were measured using waveforms, spectrograms, and FFTs. The voice quality of adjacent vowels was examined for possible consonantal effects (e.g., creaky voicing adjacent to ejectives). Additional observations included noise characteristics during closure and following oral release. Results show much variation in the realizations of stops both within and across speakers. While a general trend exists for VOT and burst amplitude to differentiate initial voiced and voiceless stops, values of voiceless aspirated and ejective stops overlap. In intervocalic stops, VOT and burst amplitude are more variable. Closure duration remains fairly stable across the three stop categories. Noise quality following oral release and, to some extent, voice quality of a following vowel more consistently distinguish the three stop types in both positions than do the quantitative measures. These characteristics are further discussed in relation to similar cross-linguistic studies.

5aSC11. Effects of speaking rate on the vowel length distinction in Japanese nonsense words. Yukari Hirata (Speech and Hearing Sci., CH 20, Rm. 121, Univ. of Alabama at Birmingham, 1530 3rd Ave. S., Birmingham, AL 35294, hirata@shrp.uab.edu)

This study examined how speaking rate affected the durations of phonemic short and long vowels in Japanese. Four native Japanese speakers produced five triplets of disyllabic nonsense words: CVCV, CVCVV, and CVCVV (C=/m/; V=/i/, /e/, /a/, /o/, and /u/), e.g., /máma/, /má:ma/, and /máma:/. Speakers read these words in a carrier sentence at slow, normal,

and fast rates three times each. Durations of accented and unaccented short vowels in CVCV, and contrasting accented long vowels in CVCVV and unaccented long vowels in CVCVV, as well as word durations, were measured from spectrograms. There was a significant amount of overlap between the durations of short and long vowels across the three rates. Rate changes affected the duration of long vowels more than short vowels, showing asymmetry of distribution, consistent with Port *et al.* [Phonetica 37, 235–252 (1980)]. In contrast with the absolute durations, the ratios of long to short vowels, three-mora (CVCVV or CVCVV) to two-mora (CVCV) words, and vowel (short or long) to word durations were little affected by rate changes. These results are consistent with the view that a relational invariance exists in temporal dimensions of speech [Pickett *et al.*, Phonetica 56, 135–157 (1999)].

5aSC12. An acoustic study of strident fricatives in Mandarin Chinese. Chao-Yang Lee (Speech Commun. Group, Res. Lab. of Electron., MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Mandarin has a distinction between flat and palatalized post-alveolar fricatives in addition to the alveolar–palatoalveolar contrast, which has been shown to have well-defined acoustic correlates. The goal of the study is to identify the acoustic cues to the place distinction in these strident fricatives and to evaluate how the acoustic properties map onto relevant phonetic features. Acoustic analyses were conducted of 216 fricative-vowel syllables produced by six speakers in three vowel and four tone contexts. Measures were taken of the properties of the frication noise and comparisons between fricative and vowel. The lower frequency limit of the frication noise, spectral peak location, spectral moments, spectral tilt, normalized amplitude, relative amplitude in specific frequency regions, and *F2* onset frequency and movement in the adjacent vowel were examined. The results indicate the lower frequency limit of the frication noise is associated with distinct formant regions: *F4* or *F5* for the alveolar, *F3* for the palatalized post-alveolar, and *F4* for the flat post-alveolar fricative. Measures that successfully classified all three fricatives include spectral peak location, spectral mean, variance and skewness, amplitude difference between fricative and vowel in the *F3* and *F4* region, and *F2* frequency at vowel onset. [Work supported by NIH.]

5aSC13. Sensitivity of spectral moments to Swedish dental versus American English alveolar word-initial /t/ bursts. Eugene H. Buder (School of Audiol. & Speech-Lang. Pathol., The Univ. of Memphis, 807 Jefferson Ave., Memphis, TN 38105, ehbuder@memphis.edu), Douglas F. Parham (The Univ. of Memphis, Memphis, TN 38105), and Carol Stoel-Gammon (Univ. of Washington, Seattle, WA 98105-6246)

Previous investigations comparing Swedish dental versus American English alveolar /t/ production found many acoustic differences in both adults and 30-month-old children [C. Stoel-Gammon, K. Williams, and E. H. Buder, "Cross-language differences in phonological acquisition: Swedish and American /t/," Phonetica 51, 146–158 (1994)]. The primary differences in the word initial consonants included the second spectral moment ("spectral SD"), VOT, and burst intensity. However, subsequent studies demonstrated that spectral SD and other moments measures were sensitive to recording environment and correlated with other production characteristics. Moreover, measurements of spectral SD differed across these languages in other stop bursts such as /p/ and /k/. These observations raised questions regarding the true sensitivity of this measure in place of articulation for /t/. To examine these questions, the current work examines stop burst characteristics of two adult male native speakers of Swedish and American English recorded in an anechoic chamber with digital equipment. Results suggest that the primary variation of spectral SD in /t/ bursts is attributable to the dental versus the alveolar place of articulation. However, the results indicate that spectral moment measures are also generally sensitive to burst intensity, which may be unrelated to place but still language-specific. [Work supported by NICHD R01-HD32065.]

5aSC14. Stress effects on coarticulation in English and Greek. Laura Koenig (Haskins Labs, 270 Crown St., New Haven, CT 06511 and Long Island Univ., Brooklyn, koenig@haskins.yale.edu) and Areti Okalidou (Univ. of Macedonia, Thessaloniki 540 06, Greece)

The effects of stress on production variability and V-to-V coarticulation were compared in American English and Greek, two languages which differ in vowel inventory size and in the magnitude of V-to-V coarticulation (Okalidou and Koenig, 1999). Six speakers, one male and two females from each language, were recorded producing nonsense VCV utterances in a carrier phrase, with randomly alternating stress. The Greek stimuli included all five vowels of the language; the English stimuli included the closest counterparts of the Greek vowels. The medial plosive consonants alternated between the bilabial and alveolar place of articulation, yielding different degrees of coarticulatory resistance (Recasens, 1985, 1989). Plosives were chosen to have similar VOT values across languages in order to minimize duration effects. A comparison of stressed versus unstressed vowel areas reveals significant differences across languages. Specific vowel effects and language effects are also noted with respect to changes in production variability under the stressed versus unstressed conditions. The magnitude of V-to-V influences as a function of stress and position is discussed in terms of the above crosslinguistic comparison in order to obtain further insight on the coarticulatory behavior of languages which differ in the size of vowel inventory (Manuel and Krakow, 1984).

5aSC15. Cross-language and speaker differences in the durational patterns of different topic transition types. Caroline Smith and Lisa Hogan (Dept. of Linguist., Univ. of New Mexico, Albuquerque, NM 87131, caroline@unm.edu)

Previous work [Smith and Hogan (2001), (2002)] has shown that the organization of topics in a text affects durational patterning when the text is read aloud. Comparable texts in English, French, and Japanese were selected, and native speakers of each language were recorded reading them aloud. For all three texts, the transition in topic from each sentence to the next was labeled as a Shift, Continuation, or Elaboration. Acoustic measurements were made of lengthening of the sentence-final word, the duration of pauses between sentences and changes in speaking rate from one sentence to the next, and in all languages, these measures were significantly different for different topic transition types. The direction and magnitude of the effects differed considerably among the single speakers of the three languages examined so far. However, with just one speaker per language, it is impossible to distinguish language-specific from speaker-specific behavior. Additional data from more speakers will be reported, suggesting that differences among languages can be attributed to known prosodic differences, whereas differences among speakers of the same language are less systematic. [Work supported by NSF Grant No. BCS-9983106.]

5aSC16. Affricate gemination in the English of Polish speakers: A study in second language variability. Ela Thurgood (Dept. of English, California State Univ. Chico, Chico, CA 95920-0830)

This study investigates the nature of the acoustic variation in sequences of identical affricates produced by Polish learners of English. In both English and Polish sequences of identical affricates occur across word boundaries, but only in Polish do such sequences also occur root internally and across morpheme boundaries. In Polish sequences of identical affricates are manifested variably both by rearticulation of both affricates and by articulation of a single affricate but with lengthened duration of either the stop or the fricative. To investigate their English, the subjects performed two tasks: repetition of 12 English sentences and orally responding to 17 multiple choice questions. The task produced significant cross-speaker differences in the phonetics of the geminates, differences correlated with differences in their proficiency levels in English. The more Polish-like singly articulated long affricates were produced by 22% of the intermediate speakers but by 48% of the advanced speakers, the opposite

of what one might expect. The intermediate speakers appear to have paid more attention to the phonetics of the English cues, thus producing more fully rearticulated affricates; the more advanced speakers appear to have paid less attention to the phonetics of the cues, thus reverting more to the norms of Polish pronunciation.

5aSC17. Perception of English palatal codas by Korean speakers of English. Sang-Hee Yeon (Prog. in Linguist., Univ. of Florida, P.O. Box 115454, FL)

This study aimed at looking at perception of English palatal codas by Korean speakers of English to determine if perception problems are the source of production problems. In particular, first, this study looked at the possible first language effect on the perception of English palatal codas. Second, a possible perceptual source of vowel epenthesis after English palatal codas was investigated. In addition, individual factors, such as length of residence, TOEFL score, gender and academic status, were compared to determine if those affected the varying degree of the perception accuracy. Eleven adult Korean speakers of English as well as three native speakers of English participated in the study. Three sets of a perception test including identification of minimally different English pseudo- or real words were carried out. The results showed that, first, the Korean speakers perceived the English codas significantly worse than the Americans. Second, the study supported the idea that Koreans perceived an extra /i/ after the final affricates due to final release. Finally, none of the individual factors explained the varying degree of the perceptual accuracy. In particular, TOEFL scores and the perception test scores did not have any statistically significant association.

5aSC18. Effects of consonantal contexts on the perception of English vowels by experienced and inexperienced Japanese learners of English. Takeshi Nozawa (Kansai Univ. of Intl. Studies, 1-18 Aoyama, Shijimi-cho Miki Hyogo 673-0521, Japan), Ratree Wayland (Univ. of Florida at Gainesville, Gainesville, FL 32611-5454), and Elaina M. Frieda (Auburn Univ., AL 36849)

The perception of English vowels by experienced and inexperienced Japanese learners of English in different consonantal context were compared. Experienced Japanese learners of English were recruited in Gainesville, FL and inexperienced Japanese learners of English were recruited in Kobe, Japan and vicinity. In one experiment, subjects' ability was tested to discriminate English vowels in different consonantal contexts. What was expected to affect the discriminability was (1) voicing of the preceding and the following stops, (2) places of articulation of the preceding and the following stops, and (3) nasality of the following consonant. In another experiment, the subjects identified English vowels in terms of Japanese vowel categories. Generally, English vowels identified with the same Japanese vowels were more difficult to discriminate than those identified with different vowels. In both experiments, the inexperienced learners were more susceptible to talker and consonantal context differences. [Work supported by Grant-in-Aid for Scientific Research (C)(1410635).]

5aSC19. Sentence durations and accentedness judgments. Z. S. Bond, Verna Stockmal (Ohio Univ., Gordy Hall, Athens, OH 45701, Bond@ohio.edu), and Dace Markus (Univ. of Latvia, Riga, Latvia)

Talkers in a second language can frequently be identified as speaking with a foreign accent. It is not clear to what degree a foreign accent represents specific deviations from a target language versus more general characteristics. We examined the identifications of native and non-native talkers by listeners with various amount of knowledge of the target language. Native and non-native speakers of Latvian provided materials. All the non-native talkers spoke Russian as their first language and were long-term residents of Latvia. A listening test, containing sentences excerpted from a short recorded passage, was presented to three groups of listeners: native speakers of Latvian, Russians for whom Latvian was a second

language, and Americans with no knowledge of either of the two languages. The listeners were asked to judge whether each utterance was produced by a native or non-native talker. The Latvians identified the non-native talkers very accurately, 88%. The Russians were somewhat less accurate, 83%. The American listeners were least accurate, but still identified the non-native talkers at above chance levels, 62%. Sentence durations correlated with the judgments provided by the American listeners but not with the judgments provided by native or L2 listeners.

5aSC20. Evaluating causes of foreign accent in English sentences spoken by native speakers of Italian differing in age of arrival (AOA) in Canada. James Flege (Div. of Speech and Hearing Sci., Univ. of Alabama at Birmingham, Birmingham, AL 35294, jeflege@uab.edu), Ian MacKay (Univ. of Ottawa, Ottawa, ON, Canada), and Satomi Imai (Univ. of Alabama at Birmingham, Birmingham, AL 35294)

This study evaluated potential causes of foreign accent (FA) by including native Italian (NI) speakers with a later age of arrival (AOA) in Canada than in previous studies. Three NI groups ($n=18$ each) differing in AOA (means=10, 18, and 26 years) participated. Listeners used a 9-point scale to rate sentences produced by the three NI groups and native English controls. The ratings obtained for all four groups differed significantly. The stronger foreign accents of the AOA-18 than AOA-10 group might be attributed to the passing of a critical period, or to stronger cross-language interference by more robust Italian phonetic categories. The difference might also be attributed to differences in language use. This is because the AOA-10 and AOA-18 groups (but not the AOA-18 and AOA-26 groups) differed significantly in percentage of English and Italian use, length of residence in Canada, and years of education in Canada. None of these explanations will apparently explain the stronger FAs of the AOA-26 than AOA-18 group. The difference between these groups might be attributed to cognitive aging [Hakuta *et al.*, Appl. Psycholinguistics (in press)], which results in gradually less successful second-language acquisition across the adult life span. [Work supported by NIH.]

5aSC21. Vowel spacing in four-vowel systems. Ian Maddieson (Univ. of California, Berkeley, CA 94720-2650)

Languages with fewer than average vowel quality contrasts provide a testbed for proposed universals of vowel systems. This paper examines the positioning in two- and three-formant acoustic spaces of the vowels of several languages with four contrastive vowel qualities. All have vowels like /i/ and /a/, with /i/ seeming the most similar across the languages. Most have a back rounded vowel with a higher $F1$ than that of /i/, resulting in the transcription /o/ rather than /u/. The position of the fourth vowel is highly variable. In San Carlos Apache it is a front mid vowel, in Shipibo a high central-back vowel, and in Paiwan a mid central vowel. The overall maximal separation of vowels in the acoustic space is therefore not the major organizing principle of these vowel systems. Rather, their construction is anchored by the polarity of /i/ vs /a/ (dispersed spectral energy versus compact mid-frequency energy). A high back vowel with close rounding is the next most distinct vowel in terms of spectral energy distribution. But this is avoided, perhaps because vowels of this type have attenuated amplitude, and a lower vowel is preferred. The fourth vowel is free to take a variety of characteristics.

5aSC22. Padé approximations for boundary-layer losses in articulatory synthesis. Reiner Wilhelms-Tricarico and Richard S. McGowan (CRESS LLC, 1 Seaborn Pl., Lexington, MA 02420, rsmcgowan@earthlink.net)

To build an articulatory synthesizer it is necessary to model acoustic propagation in tubes with variable area. Acoustic propagation entails viscous and thermal losses, which are strongest at the vocal tract walls. In the

standard boundary-layer approximation irrational frequency laws containing the square root of frequency best represent these losses. Our immediate goal is, given a sound source distribution in the vocal tract, to efficiently calculate the sound output from the mouth. In order to use digital filter theory in this process, a rational approximation to the square root of frequency is sought in the form of Padé approximations. One implementation is obtained by modification of the Kelly–Lochbaum algorithm for calculating wave propagation in a tube, using a high over-sampling rate. However, frequency-dependent loss means that both the reflection coefficients and the time delays through a tube section are non-constant functions of frequency—an assumption used in the Kelly–Lochbaum algorithm. The reflection coefficients are replaced by digital filters and the delay elements by filters with frequency-dependent group velocities. We discuss implementation of a Kelly–Lochbaum algorithm in a digital filter design using a Padé approximation of the viscous and thermal loss terms. [Work supported by NIH Grant No. DC-01247.]

5aSC23. Speaker recognition using dynamic synapse based neural networks with wavelet processing. Sageev George, Theodore Berger, Alireza Dibazar, and Walter Yamada (3650 McClintock Ave., OHE 500, Los Angeles, CA 90089-1451, sageev@usc.edu)

We have designed systems utilizing a dynamic synapse neural network (DSNN) to perform speech recognition tasks. The DSNN architecture has demonstrated noise-resistive properties in previous research. The speech signal is first passed through a bank of bandpass filters. The output of each filter then undergoes wavelet transformation to extract speaker-dependent features. Transformed signals are used as input for a DSNN, which is trained using a genetic algorithm training method. We are seeking to explore the capabilities of a biologically based neural network in solving a problem which humans are still better at solving than conventional solutions. The dynamic synapse neural networks were trained to perform several speech recognition tasks, including speaker verification, speaker identification, and word recognition. In the speaker verification task, the objective was to distinguish one speaker from a group of up to 22 imposters based on the utterance of a codeword. In the speaker identification task, the objective was to identify which of 6 speakers has uttered the codeword. In the word recognition task, the objective was to perform speaker-dependent word recognition on a set of 30 words. Neural networks were capable of 90% correct identifications in the presence of white noise (0 dB SNR).

5aSC24. A new approach for isolated word recognition using dynamic synapse neural networks. Alireza A. Dibazar, Hassan Heidari Namarvar, and Theodore W. Berger (BME Dept., Univ. of Southern California, OHE-500, University Park, Los Angeles, CA 90089-1451, dibazar@usc.edu)

In this paper we introduce a hybrid system for robust automatic speech recognition. We first present a new architecture for dynamic synapse neural networks (DSNNs) for a speech recognition task and then we extend the quasi-linearization algorithm to estimate the DSNNs parameters. This algorithm converges quadratically to the external solution of a given set of nonlinear differential equations. In our application, the DSNN parameters and or the signals spectrums estimated by DSNN are classified using a hidden Markov model (HMM) based classifier. Our results indicate that the features obtained from DSNN are robust with respect to state of the art frequency features in the presence of a high level of noise. [Work supported in part by DARPA, NASA, and ONR.]

Session 5aUW

Underwater Acoustics: Environmental Variability and Scattering

David King, Chair

Naval Research Laboratory, Stennis Space Center, Mississippi 39529-5004

Contributed Papers

8:00

5aUW1. Spatial and temporal coherence of broadband acoustic transmissions in the Straits of Florida. Harry A. DeFerrari, Neil Williams, and Hien Nguyen (RSMAS, Univ. of Miami, Miami, FL 33149)

Month-long time series of broadband coherent measurements of channel pulse responses in the Florida Straits allow for estimation of signal coherence under a great variety of signal parameters and environmental conditions. Two 32-element arrays, one vertical and another horizontal (bottomed) along the path of propagation allow for comparison of spatial coherency and single phone temporal coherencies. The transmitted signals cover 5 octaves from 100 through 3200 Hz. Coherencies for single resolved SRBR arrivals are compared with those for unresolved multipath BRB focused arrivals. Many factors are at play including the complication of coherent reception from nearby shipping and multipath interference. However, the time series are long enough to sort out and explain most relations to the environmental variability. Vertical and horizontal coherence lengths are compared over a wide range of frequencies. Generally, SBRB paths are found to be far more stable and coherent than RBR paths especially at higher frequencies suggesting that sound-speed variability near turning RBR rays/modes is more destructive to coherency. The loss of signal coherency for RBR paths is accompanied by a significant loss of signal intensity—as much as 10 to 15 dB.

8:15

5aUW2. Acoustic effects of near-bottom soliton packets. Stanley A. Chin-Bing, Alex C. Warn-Varnas, David B. King (Naval Res. Lab., Stennis Space Center, MS 39529-5004), Kevin G. Lamb (Univ. of Waterloo, Waterloo, ON N2L 3G1, Canada), James A. Hawkins (Planning Systems, Inc., Slidell, LA 70458), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The winter 1997 Primer4 experiment was conducted on the shelfbreak and continental slope south of Cape Code in the Middle Atlantic Bight. Internal solitary waves were surveyed with rapid-sampling thermistor chains, current meters, and an upward looking ADCP. Using the available environmental data we have simulated the generation of these near-bottom solitary internal waves in this region using a forcing tidal velocity of 0.3 m/s to initiate the primitive equation Lamb model. The simulated soliton packets compare favorably in period and amplitude with the measured data taken in the winter 1997 Primer4 experiment. Previously, resonance effects leading to anomalous signal losses have been observed in numerous simulations of acoustic propagation through near-surface soliton packets. The objective of this work is to determine if similar resonance effects are caused by the interaction of the acoustic field with these near-bottom soliton packets. In contrast to our previous acoustic simulations, it was necessary to place the acoustic source and receivers above the thermocline for these near-bottom soliton packet studies. Simulation results will be presented that illustrate the similarities and differences between the acoustic effects produced by the near-surface and near-bottom soliton packets. [Work supported by ONR/NRL.]

8:30

5aUW3. Physical oceanography and acoustic propagation during LADC experiment in the Gulf of Mexico in 2001. Sergey Vinogradov, Jerald W. Caruthers, Grayson H. Rayborn (The Univ. of Southern Mississippi, Stennis Space Center, MS 39529), Ilya A. Udovydchenkov, Natalia A. Sidorovskaia, Irina I. Rypina (The Univ. of Louisiana, Lafayette, LA 70504-4210), Joal J. Newcomb, Robert A. Fisher (The Naval Res. Lab., Stennis Space Center, MS 39529), George E. Ioup, and Juliette W. Ioup (The Univ. of New Orleans, New Orleans, LA 70148)

The Littoral Acoustic Demonstration Center (LADC) deployed three environmental and acoustic moorings in a downslope line just off the Mississippi River Delta in the northern Gulf of Mexico in an area of a large concentration of sperm whales in July 2001. The measurement of whale vocalizations and, more generally, ambient noise, were the objectives of the experiment. Each mooring had a single hydrophone autonomously recording Environmental Acoustic Recording System (EARS) obtained from the U.S. Naval Oceanographic Office and modified to recorded signals up to 5859 Hz continuously for 36 days. Also, self-recording, environmental sensors were attached to the moorings to obtain profiles of time series data of temperature and salinity. Satellite imagery and NOAA mooring data were gathered for an analysis of eddy formations and movement in the Gulf. This paper will discuss the possible environmental impact of two events that occurred during the experiment: the passage of Tropical Storm Barry and the movement of the remnants of an eddy in the area. Discussed also will be the expected effects of these events on acoustic propagation based on modeling, which are carried out for long range and low frequency (300 km and 500 Hz) using the normal-mode acoustic model SWAMP (Shallow Water Acoustic Modal Propagation by M. F. Werby and N. A. Sidorovskaia) and for short range and high frequency (10 km and 5000 Hz) using the parabolic-equation acoustic model RAM (Range-dependent Acoustic model by M. Collins). [Work supported by ONR.]

8:45

5aUW4. Temporal coherent and total intensity fluctuation of mid-frequency acoustic wave propagation during SWAT 2000 experiment. Haruhiko Yamaoka and Tokuo Yamamoto (Geoacoustic Lab., Appl. Marine Phys. Div. RSMAS, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149, hyamaoka@rsmas.miami.edu)

Temporal coherent and total intensity fluctuations of acoustic wave propagation have been measured during SWAT (Shallow Water Acoustic Technology) 2000 Experiment data. The objective of this paper is to investigate the effect of the temporal fluctuation of an acoustic signal in shallow water. The acoustic propagation experiment was carried out among the three bottom mounted sonar towers, that were located in a triangle formed by 3, 4 and 5 km ranges. The center frequency of the acoustic signal is 5.5 kHz. A total of 32 hours acoustic data was sampled by a 2 min or 20 min sampling interval. In addition, oceanographic data including current velocity profile data and sound speed data is used to estimate the internal wave activity on the observation triangle. The time series of coherent intensity quickly attenuates, while the total intensity is maintained with weak fluctuation. According to the reciprocal transmission data, the reciprocity of coherent intensity is significantly distorted by

the existence of a large internal wave, while the total intensity retains the reciprocity. The characteristics of these quantities are discussed by using the PE (Parabolic Equation) model. [Work supported by ONR 321OA, NIRE, and MEST.]

9:00

5aUW5. Fluctuations in acoustic transmissions when propagation is at an oblique angle to the direction of oceanic tides. Jacob George and Robert L. Field (NRL Code 7185, Stennis Space Center, MS 39529)

Ocean tides such as the lunar M2 tide in coastal regions cause time variations (also range variations) in water depth as well as sound speed profile. When acoustic transmissions are perpendicular to the direction of the tide the received signal exhibits the fundamental oceanographic frequency of the tide and overtones of the fundamental. When the transmissions are along the direction of the tide or at an oblique angle over a range spanning several tidal wavelengths, some reduction in fluctuations can be expected from general theoretical considerations. This phenomenon is studied using the adiabatic mode/WKB model, and the implications for coastal regions are discussed. For example, in SWAT data, the tidal phase remains constant over a 10 km range span along the tidal wave-front, but undergoes change over a similar range span along the tidal direction. Several illustrative examples are presented to show both the magnitude and complexity (relevant to overtone structure) of the acoustic fluctuations. [Work supported by ONR (PE-62435N), and administered by NRL.]

9:15

5aUW6. Studies of horizontal refraction and scattering of low-frequency acoustic signals using a modal approach in signal processing of NPAL data. Alexander G. Voronovich, Vladimir E. Ostashev (Environ. Technol. Lab./NOAA, 325 Broadway, Boulder, CO 80305), and The NPAL Group^{a)}

In our previous paper [J. Acoust. Soc. Am. **112**, 2232], we obtained a time dependence of the horizontal refraction angle (HRA) of acoustic signals propagating over a range of about 4000 km in the ocean. This dependence was computed by processing of acoustic signals recorded during the North Pacific Acoustic Laboratory (NPAL) experiment using a ray-type approach. In the present paper, we consider the results obtained in signal processing of the same data using a modal approach. In this approach, the acoustic field is represented as a sum of local acoustic modes with amplitudes depending on a frequency and arrival angle. We obtained a time dependence of HRA for a time interval of about a year. Time evolution of HRA exhibits long-period variations which could be associated with seasonal trends in the sound speed profiles. The results are consistent with those obtained by the ray approach. Different horizontal angles within arrivals were impossible to resolve due to sound scattering by internal waves. A theoretical estimate of the angular width of the acoustic signals in a horizontal plane was obtained. It appears to be consistent with the observed variance of HRA 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.

9:30

5aUW7. Effects of the propagation environment on coherence of underwater broadband acoustic signals. Rachel Romond, David L. Bradley, and R. Lee Culver (Appl. Res. Lab., Penn State Univ., P.O. Box 30, N. Atherton St., State College, PA 16804-0030, romond@psu.edu)

Acoustic and environmental measurements were performed in 550 m of water approximately 2500 m east of San Clemente Island, California over a four-day period in August 2002. Continuous wave and frequency modulated pulses with center frequencies of 20 kHz to 40 kHz were transmitted over a 700 m distance underwater. Four projectors were located between 10 m and 70 m deep, and five receivers were positioned between 45 m and 220 m deep. Data from an anemometer, wave rider buoy, acoustic Doppler current profiler and towed conductivity-temperature-depth array were recorded simultaneously over the test period to characterize the marine environment. The extensive set of environmental data provided a

basis for boundary condition and water column characterization. These conditions were input into a propagation model to predict the acoustic signal structure and propagation behavior in the measured environment. Predictions of acoustic propagation parameters will be compared to results from on-site acoustic measurements. [Work supported by ONR under Award No. N00014-02-1-0156.]

9:45

5aUW8. The two-frequency third moment with internal wave induced fluctuations. Frank S. Henyey and Terry E. Ewart (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105-6698)

Many moments of the pressure field that results from sound propagation in the random ocean, for example the first moment, fall rapidly with range due to phase fluctuations associated with travel time "wander." As a result, such moments are not of much experimental interest. However, just as with intensity moments, the third moment $\langle p^2(\omega)p^*(2\omega) \rangle$ is free from wander effects. In the Mid-ocean Acoustic Transmission Experiment (MATE), data was taken from which this moment can be extracted, both for $\omega=2$ kHz and $\omega=4$ kHz. The standard theory (assuming the Markov approximation) is developed for this moment and compared to the MATE results. The imaginary part of this moment describes the frequency dependence of travel time, while its real part can be thought of as an intensity moment. [Work supported by ONR.]

10:00–10:15 Break

10:15

5aUW9. Amplitude and time fluctuations and their relationship to temperature variations. Steven D. Lutz, R. Lee Culver, and David L. Bradley (Appl. Res. Lab., The Penn State Univ., P.O. Box 30, State College, PA 16804)

Amplitude and time fluctuations of direct and refracted paths measured during a field test near San Clemente Island in August 2002 are presented. Signals used include 20 and 40 kHz CW pulses with 0.14 ms and 1.0 ms durations. The transmitters were suspended from a buoy with a bungee section in the mooring to minimize the watch circle. The receive hydrophone array was suspended from a research vessel in a three point moor. The source to receiver separation was 714 m. Towed CTD data, also collected during the experiment, is presented. A relationship between the amplitude and phase fluctuation in the acoustic arrivals and temperature variations in the water column is developed and compared to existing theory. Acoustic signals will be analyzed using a matched filter, fluctuations will be computed in normalized level and variance of time of arrival. Temperature variations will be computed as mean-square fluctuations of the index of refraction. Other physical mechanisms contributing to acoustic fluctuations including source/receiver motion, turbulence, and internal waves are also investigated.

10:30

5aUW10. Bistatic reverberation benchmarking exercise: BiStaR versus analytic formulas. Kevin D. LePage (Naval Res. Lab., Code 7144, 4555 Overlook Ave. SW, Washington, DC 20375) and Chris H. Harrison (SACLANT Undersea Res. Ctr., 19138 La Spezia (SP), Italy)

The newly developed BiStaR range-dependent bistatic reverberation model is benchmarked against analytic formulas for bistatic reverberation intensity in range-dependent environments. The BiStaR model is based on the C-SNAP range-dependent normal mode propagation model with a narrow-band approximation time domain extension and a coherent patch interaction model. The code is also capable of producing incoherent reverberation predictions that here are directly compared to the analytic formulas. The analytic formulas have been derived using ray-based intensity formulas for propagation in range-dependent waveguides and a generalized class of Lambert law type scattering kernels. Comparison between the BiStaR and the analytic predictions show good agreement for a variety of range-dependent bistatic scenarios. Such agreement generates confidence

in the closed form formulas, which are particularly valuable for generating insight into the simple mechanisms which control the dominant characteristics of range-dependent bistatic reverberation, and in the BiStaR model, which has been developed to model bistatic reverberation in 3-D environments with high resolution environmental characterizations. [Work supported by ONR and NATO SACLANT Undersea Research Centre.]

10:45

5aUW11. Doubly-focused echos from spheres unfold into a hyperbolic umbilic diffraction catastrophe. Ben Dzikowicz and Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99163, dzikowicz@wsunix.wsu.edu)

An underwater spherical target resides in an Airy field formed by reflection off a curved surface representing the sea floor or sea surface. In prior work [B. Dzikowicz and P. L. Marston, *J. Acoust. Soc. Am.* **110**, 2778 (2001)] direct returns of a tone burst from the surface reflection focused toward the target were shown to have a dependence on the target position described by an Airy function. The return echo can also be focused again by the surface onto the source and receive transducer. This gives the square of an Airy function for the case of a point target. With a finite sized target (as in the experiment) this goes over to a hyperbolic umbilic catastrophe with symmetric arguments. The arguments of the hyperbolic umbilic function are derived from only the relative return times of a transient pulse. Experiment confirms the predicted merging of transient echoes in the time domain, as well as the hyperbolic umbilic diffraction integral amplitudes for a tone burst. This method would allow for the observation of a target at a greater distance in the presence of a focusing surface. [Research supported by ONR.]

11:00

5aUW12. Acoustical scattering from N spheres using a multilevel fast multipole method. Nail A. Gumerov and Ramani Duraiswami (Perceptual Interface and Reality Lab., Inst. for Adv. Computer Studies, Univ. of Maryland, College Park, MD 20742)

We develop an efficient computational method for wave scattering from a large number of spherical objects that are characterized by their radii, locations, and complex acoustical impedances of the surfaces. The direct T-matrix method for solution of the problem that was developed and tested in [Gumerov and Duraiswami, *J. Acoust. Soc. Am.* **112**, 2688–2701 (2002)] is inefficient for computations involving a large number of scatterers. Here, we implement and test a multilevel fast multipole method for speeding up the solution and achieving better memory complexity. The method is based on hierarchical space subdivision with oct-trees using optimal space-partitioning, on a theory for fast translation and re-expansion of multipole solutions of the Helmholtz equation, and employs an iterative technique for the solution of large dense systems of linear equations with reflection-based iterations. For N scatterers the method provides $O(N \log N)$ asymptotic complexity opposed to $O(N^3)$ complexity of the direct T-matrix approach. The results of computations were tested against solutions obtained by other methods, such as the boundary element method and the direct T-matrix method tested in our early study, and show the computational efficiency and accuracy of the solution technique. [Work supported by NSF Awards 0086075 and 0219681 is gratefully acknowledged.]

11:15

5aUW13. Tilt angle dependence of backscattering enhancements from organ pipe modes of open water-filled cylinders: Measurements and models. Curtis F. Osterhoudt and Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814)

A simple target for simulating narrow low-frequency resonances of cylinders is an open metal pipe completely filled with water. We have previously described how the high- Q organ-pipe modes having a pressure node near each end are easily observed in backscattering experiments with

small cylinders [C. F. Osterhoudt and P. L. Marston, *J. Acoust. Soc. Am.* **110**, 2773 (2001)]. The resonance occurs because of the strong reflection of internal acoustic waves from the open ends of the pipe [H. Levine and J. Schwinger, *Phys. Rev.* **73**, 383–406 (1948)]. In the present research, the dependence of the backscattering amplitude on the orientation of the cylinder is measured and modeled. The tilt angle dependence is affected by the symmetry of the organ pipe mode. An approximation was also developed for the backscattering amplitude at high Q resonances based on energy conservation, reciprocity, and the optical theorem. While this analysis applies to cylinders suspended in water away from boundaries, the organ-pipe modes studied may be useful for investigating scattering processes for buried or partially buried cylinders. [Research supported in part by ONR.]

11:30

5aUW14. Local coupled modes and volume scattering in heterogeneous anisotropic shallow water environments. Darin J. Soukup (Dept. of Earth and Space Sci., Univ. of Washington, 63 Johnson Hall, Box 351310, Seattle, WA 98195) and Robert I. Odom (Univ. of Washington, Seattle, WA 98195)

Seafloor bottom/subbottom interactions are significant for low frequency acoustic wave propagation in shallow water environments. We investigate the significance of deterministic and stochastic volume scattering for generally anisotropic seafloor bottoms. The following assumptions are made: (i) the model is laterally heterogeneous with an elastic bottom/subbottom, (ii) anisotropy is hexagonally symmetric, (iii) the bottom/subbottom structure is dominated by deterministic elastic moduli with additional small-scale stochastic variations. The local coupled mode rough interface scattering theory of Park and Odom [*Geophys. J. Int.* **136**, 123 (1999)] is extended by including volume scattering terms. New volume scattering terms are derived by applying perturbation theory to the elastic equations of motion. The acoustic wavefield is expressed as a sum of a mean deterministic wavefield and a small-scale scattered wavefield. Perturbations in the elastic moduli contribute to energy loss and the loss of signal coherence. The local mode formalism brings insight into the efficiency of volume scattering by indicating directly how much energy of coherent mode q is redistributed into scattered mode r . Numerical calculations indicating the effects of volume scattering are presented. [Work supported by ONR.]

11:45

5aUW15. On the relationship between signal bandwidth and coherence for ocean surface forward scattered signals. R. Lee Culver, David L. Bradley, and Joe G. Keranen (Appl. Res. Lab., The Penn State Univ., P.O. Box 30, State College, PA 16804)

The relationship between the signal bandwidth and the coherence of high frequency, surface forward scattered underwater acoustic signals has been investigated experimentally and compared with available theory. Acoustic measurements were made under ONR sponsorship (the ARL Program) during 2000 [J. G. Keranen *et al.*, *J. Acoust. Soc. Am.* **109** (2001)] and again in 2002, both times in deep water off the coast of San Diego. Transmitted signals consisted of 0.1 to 1.0 ms CW pulses and 500 Hz to 22 kHz wide linear frequency modulated sweeps using center frequencies of 18 to 46 kHz. Concurrent environmental measurements included water temperature, current speed and direction, directional wave height and wind speed. An analysis of the data is focused on how the coherence of the surface forward scattered path depends upon the signal bandwidth. We evaluate the usefulness of the frequency coherence bandwidth parameter [P. H. Dahl, *IEEE J. Oceanic Eng.* **26**(1) (2001)], and compare measurements with time spread-based theoretical predictions [J. C. Reeves, Ph.D. thesis, UCLA, 1973; M. H. Brill *et al.*, *J. Acoust. Soc. Am.* **75** (1984)]. [Work supported by ONR under Award No. N00014-02-1-0156.]